

Learning and Becoming in Practice

THE INTERNATIONAL CONFERENCE OF THE LEARNING SCIENCES (ICLS) 2014

PROCEEDINGS VOLUME 2



www.isls.org/icls2014

HOSTED BY:



University of Colorado
Boulder

SPONSORED BY:



International Society of
the Learning Sciences

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

Volume 2

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 978-0-9903550-3-8 (Volume 2, 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 2*. 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

CAMBRIDGE

Forthcoming Scholarship from Cambridge University Press

The Cambridge Handbook of the Learning Sciences, 2nd Edition

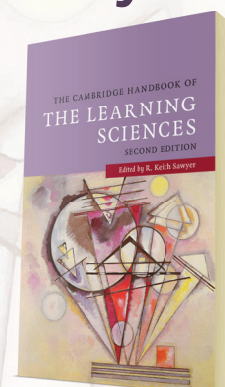
Edited by R. Keith Sawyer
University of North Carolina, Chapel Hill

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

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
 Cathryn 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

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 Ierche 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

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

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

Table of Contents

Volume 1

Keynotes

Changing Practice	3
<i>Jean Lave, Rogers Hall</i>	
Research-Practice Partnerships	4
<i>Megan Bang, Paul Cobb, Kara Jackson, Michael Sorum, Kris D. Gutiérrez</i>	
Designing with Communities: Transforming Historically Powered Relations in Teaching and Learning	4
<i>Megan Bang</i>	
Partnering with School and Districts to Support All Students' Learning	4
<i>Paul Cobb, Kara Jackson, Michael Sorum</i>	
I, Thou, and Them: Distributed Memory and Learning.....	6
<i>Geoffrey C. Bowker, Kevin O'Connor</i>	
Approaches to Studying and Modeling Learning Across Setting and Time	7
<i>Jeremy Roschelle, Anna Sfard, Reed Stevens, Leona Schauble, Beth Warren, Marianne Wiser</i>	

Invited Sessions

Presidential Session: Learning and Assessment of 21st Century Skills	11
<i>Eleni A. Kyza, Jan van Aalst, Daniel T. Hickey, Alina von Davier, Jan-Willem Strijbos, Cindy E. Hmelo-Silver</i>	
Research-Practice Partnerships in Communities	12
<i>Daniel J. Gallagher, Nichole Pinkard, Jasmine Alfonso-Gurneau, Megan Bang, Mary Dempsey, Lori Faber, Ananda Marin, Tim S. Truitt</i>	
Children Becoming Collaborators	13
<i>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</i>	
Teachers as Designers.....	14
<i>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</i>	
Where are the Learning Sciences in the MOOC Debate?.....	15
<i>George Siemens, Pierre Dillenbourg, Gerhard Fischer, Danielle S. McNamara, Nikol Rummel</i>	
Learning and Becoming through Making and Participatory Media	18
<i>Kristiina Kumpulainen, Julian Sefton-Green, Karen Brennan, Anna Mikkola, Kylie Peppler, Elisabeth Soep</i>	

Papers

Leveling Transparency via Situated Intermediary Learning Objectives (SILOs)	23
<i>Dor Abrahamson, Kiera Chase, Vishesh Kumar, Rishika Jain</i>	
Supporting Middle Schoolers' Use of Inquiry Strategies For Discovering Multivariate Relations In Interactive Physics Simulations	31
<i>Luke D. Conlin, Nicole R. Hallinen, Daniel L. Schwartz</i>	

Klauer's Inductive Reasoning Training as a Cognitive Apprenticeship Approach for Special-Needs Students.....	38
<i>Antonia E. E. Baumeister, Heiner Rindermann</i>	
Development of an Empirically-Based Learning Performances Framework for 3rd-Grade Students' Model-Based Explanations about Hydrologic Cycling	46
<i>Cory T. Forbes, Christina V. Schwarz, Laura Zangori</i>	
Problematizing as Scaffold for Engaging in Scientific Argumentation.....	54
<i>Mon-Lin Monica Ko</i>	
Three Diagnoses of Why Transfer Across Disciplines Can Fail and Their Implications for Interdisciplinary Education.....	62
<i>Eric Kuo, Danielle Champney</i>	
Diving into Practice with Children and Undergraduates: A Cultural Historical Approach to Instatiating Making and Tinkering Activity in a Designed Learning Ecology.....	70
<i>Lisa Hope Schwartz, Daniela K. DiGiacomo, Kris D. Gutiérrez</i>	
Facilitating Design Research by Mapping Design Research Trajectories.....	78
<i>Guanzhong Ma, Jan van Aalst</i>	
Connected Gaming: Towards Integrating Instructionist and Constructionist Approaches in K-12 Gaming.....	86
<i>Yasmin B. Kafai, Quinn Burke</i>	
Learning With Multiple Visualizations in the Science Museum	94
<i>Joyce Wang, Susan Yoon</i>	
The Role of Inconsistencies in Collaborative Knowledge Construction	102
<i>Martina Bientzle, Ulrike Cress, Joachim Kimmerle</i>	
More Than Just Plain Old Technology Adoption: Understanding Variations in Teachers' Use of an Online Planning Tool.....	110
<i>Heather Leary, Victor R. Lee, Mimi Recker</i>	
Using an Adaptive Expertise Lens to Understand the Quality of Teachers' Classroom Implementation of Computer-Supported Reform Curricula in High School Science	118
<i>Susan Yoon, Jessica Koehler, Joyce Wang, Emma Anderson, Eric Klopfer</i>	
Promoting 5th Graders' Views of Science and Scientific Inquiry in an Epistemic-Enriched Knowledge-Building Environment.....	126
<i>Feng Lin, Carol K. K. Chan, Jan van Aalst</i>	
Becoming an Activist-Mathematician in an Age of Austerity.....	134
<i>Indigo Esmonde, Joe Curnow, Dominique Riviere</i>	
Measuring Affective Experience in the Midst of STEM Learning.....	142
<i>Jayson Nissen, Jonathan Shemwell</i>	
Epistemic Networks for Epistemic Commitments.....	150
<i>Simon Knight, Golnaz Arastoopour, David Williamson Shaffer, Simon Buckingham Shum, Karen Littleton</i>	
Framing Reflections on Instruction: A Precursor to Noticing	158
<i>Vicky Pilitsis, Ravit Golan Duncan</i>	
"Teach me how to Facebook!" Design-Based Research about Risk Prevention on Social Network Sites	166
<i>Ellen Vanderhoven, Tammy Schellens, Martin Valcke</i>	

The Role of Identity Development within Tensions in Ownership of Science Learning	174
<i>Jason Yip, Tamara Clegg, June Ahn, Elizabeth Bonsignore, Michael Gubbels, Emily Rhodes, Becky Lewittes</i>	
Developing Mechanistic Model-Based Explanations of Phenomena: Case Studies of Two Fifth Grade Students' Epistemologies in Practice Over Time	182
<i>Christina V. Schwarz, Li Ke, May Lee, Joshua Rosenberg</i>	
MOOCs: A Perspective from the Learning Sciences	190
<i>Michael Eisenberg, Gerhard Fischer</i>	
How Interpreters Make Use of Technological Supports in an Interactive Zoo Exhibit	198
<i>Brian Slattery, Leilah Lyons, Priscilla Jimenez Pazmino, Brenda A. López Silva, Tom Moher</i>	
Becoming an Activist: Learning the Politics and Performances of Youth Activism Through Legitimate Peripheral Participation.....	206
<i>Joe Curnow</i>	
"We should all help each other": Latina undergraduates' practices and identities in the figured world of computing	214
<i>Heather Thiry, Sarah Hug</i>	
Identifying Transfer of Inquiry Skills across Physical Science Simulations using Educational Data Mining.....	222
<i>Michael Sao Pedro, Yang Jiang, Luc Paquette, Ryan S. Baker, Janice Gobert</i>	
Reactivation of Multimodal Representations and Perceptual Simulations for Meaningful Learning: A Comparison of Direct Embodiment, Surrogate Embodiment, and Imagined Embodiment	230
<i>Saadia A. Khan, John B. Black</i>	
Creativity as Practice(d) in a Design Studio	238
<i>Christoph Richter, Julia Lembke, Elisa Ruhl, Heidrun Allert</i>	
Fostering Scientific Reasoning. A Meta-analysis on Intervention Studies.....	246
<i>Katharina Engelmann, Frank Fischer</i>	
Knowledge Organization with Multiple External Representations for Socioscientific Argumentation: A Case on Nuclear Energy.....	254
<i>Bahadır Namdar, Ji Shen</i>	
Students' Resources for the Construction of Scales for Graphing	262
<i>Cesar Delgado, Margaret Lucero</i>	
Scientific Practices Through Students' Eyes: How Sixth Grade Students Enact and Describe Purposes for Scientific Modeling Activities Over Time	270
<i>Christina Krist, Brian J. Reiser</i>	
The Impact of a Social Robot's Attributions for Success and Failure in a Teachable Agent Framework	278
<i>Kasia Muldner, Victor Girotto, Cecil Lozano, Winslow Burlison, Erin A. Walker</i>	
Characterizing a New Dimension of Change in Attending and Responding to the Substance of Student Thinking.....	286
<i>Jennifer Richards, Andrew Elby, Ayush Gupta</i>	
Spatial Practices in CSCL Discussions.....	294
<i>Benzi Slakmon, Baruch B. Schwarz</i>	
Collective Immersive Simulations: A New Approach to Learning and Instruction of Complex Biology Topics.....	301
<i>Michelle Lui, James D. Slotta</i>	

A Study of Subjective Emotions, Self-Regulatory Processes, and Learning Gains: Are Pedagogical Agents Effective in Fostering Learning?	309
<i>Nicholas Mudrick, Roger Azevedo, Michelle Taub, Reza Feyzi Begnagh, François Bouchet</i>	
Design-Based Research Process: Problems, Phases, and Applications	317
<i>Matthew W. Easterday, Daniel Rees Lewis, Elizabeth M. Gerber</i>	
Modeling the Construction of Evidence through Prior Knowledge and Observations from the Real World	325
<i>Lauren Barth-Cohen, Daniel Capps, Jonathan Shemwell</i>	
Automatic Coding of Questioning Patterns in Knowledge Building Discourse	333
<i>Jin Mu, Jan van Aalst, Carol K. K. Chan, Ella Fu</i>	
Enacted Misconceptions: Using Embodied Interactive Simulations to Examine Emerging Understandings of Science Concepts	341
<i>Robb Lindgren, Michael Tscholl</i>	
Friendship, Participation, and Site Design in Interest-Driven Learning among Adolescents	348
<i>Ashley Cartun, Ben Kirshner, Emily Price, Adam York</i>	
Using Models for Reasoning and Content Learning: Patterns of Bootstrapping Towards Earth Science Understandings	354
<i>Ann E. Rivet, Cheryl A. Lyons, Alison R. Miller</i>	
Towards a Complex Systems Meta-Theory of Learning as an Emergent Phenomenon: Beyond the Cognitive Versus Situative Debate	362
<i>Michael J. Jacobson, Manu Kapur, Peter Reimann</i>	
Adventures in Argument: Training in Argumentation Influences Student Resource Use in Collaborative Meaning Making	370
<i>Julia Gressick, Sharon J. Derry</i>	
Collective Engagement in a Technologically Mediated Science Learning Experience: A Case study in a botanical garden	378
<i>Fariha H. Salman, Heather Toomey Zimmerman, Susan M. Land</i>	
Promoting Student Learning through Automated Formative Guidance on Chemistry Drawings	386
<i>Anna N. Rafferty, Libby Gerard, Kevin McElhaney, Marcia C. Linn</i>	
Computer-Enhanced Dialogic Reflective Discourse	394
<i>Shiri Mor-Hagani, Dani Ben-Zvi</i>	
A Case Study Examining the Microdynamics of Social Positioning within the Context of Collaborative Group Work	402
<i>Lesley Dookie</i>	
Showing What They Know: Multimedia Artifacts to Assess learner understanding	410
<i>Cindy E. Hmelo-Silver, Carolyn A. Maher, Marjory F. Palius, Robert Sigley, Alice Alston</i>	
Varied Appropriation of Tools from Professional Development: Moving Beyond Levels	418
<i>Huy Q. Chung, Elizabeth A. van Es</i>	
Learning in Low-Performing School Districts: Conceptual and Methodological Challenges Resulting from Network Churn	426
<i>Kara S. Finnigan, Alan J. Daly</i>	
Shifts in Identification in a Hybrid Space	434
<i>Kok-Sing Tang</i>	

Shared Epistemic Agency and Agency of Individuals, Collaborative Groups, and Research Communities	440
<i>Crina Damşa</i>	
Interrogating the Divide: A Case Study of Student Technology Use in a One-to-one Laptop School	448
<i>Nicholas Wilson</i>	
Capturing Personal and Social Science: Technology for Integrating the Building Blocks of Disposition	455
<i>Tamara Clegg, Elizabeth Bonsignore, June Ahn, Jason Yip, Daniel Pauw, Michael Gubbels, Becky Lewittes, Emily Rhodes</i>	
Moving Beyond Case Studies: Using Social Network Analysis to Study Learning as Participation in Communities of Practice	463
<i>Julia Eberle, Karsten Stegmann, Frank Fischer</i>	
Designing Critique to Improve Conceptual Understanding	471
<i>Elissa Sato, Marcia C. Linn</i>	
Being Mathematical Relations: Dynamic Gestures Support Mathematical Reasoning	479
<i>Candace Walkington, Rebecca Boncoddio, Caroline Williams, Mitchell J. Nathan, Martha W. Alibali, Erica Simon, Elizabeth Pier</i>	
"Learning to Live": Expansive Learning and Mo(ve)ments Beyond 'Gang Exit'	487
<i>Line Lerche Mørck</i>	
Analyzing Equity in Collaborative Learning Situations: A Comparative Case Study in Elementary Computer Science	495
<i>Niral Shah, Colleen Lewis, Roxane Caires</i>	
Hands-on Small Group versus Whole Class Use of an Interactive Simulation: Qualitative Comparisons	503
<i>A. Lynn Stephens, John J. Clement</i>	
The Contours and Possibilities of Desire in Sociocultural Research on Learning and Becoming	511
<i>Ian Parker Renga</i>	
How Do Learners Process Information in Lectures? The Role of Projected Slides and Type of Note-taking	519
<i>Christof Wecker</i>	
Using Analytics for Improving Implementation Fidelity in an Large Scale Efficacy Trial	527
<i>Mingyu Feng, Jeremy Roschelle, Robert Murphy, Neil T. Heffernan</i>	
Where Are We Now? Research Trends in the Learning Sciences	535
<i>Elizabeth Koh, Young Hoan Cho, Imelda Caleon, Yu Wei</i>	
Defining Success in an Alternative High School: Resources for the Reframing of Education	543
<i>Gavin Tierney</i>	
"Case n' Point": Discovering Learning in the Nonce	551
<i>Timothy Koschmann, Alan Zemel, Michael Neumeister</i>	
Time Needed: Growth of Preservice Science Teachers' Knowledge of Inquiry and Practice of Lesson Design	559
<i>Augusto Z. Macalalag Jr.</i>	
Becoming Agents of Change through Participation in a Teacher-Driven Professional Research Community	567
<i>Michael Ross, Ben Van Dusen, Valerie Otero</i>	
'Mangling' Science Instruction: Creating Resistances to Support the Development of Practices and Content Knowledge	575
<i>Eve Manz</i>	

An Analytic Tool for Supporting Teachers' Reflection on Classroom Talk	583
<i>Gaowei Chen, Sherice N. Clarke, Lauren B. Resnick</i>	
The Discourse of Creative Problem Solving in Childhood Engineering Education.....	591
<i>Elise Deitrick, Brian O'Connell, R. Benjamin Shapiro</i>	
Supporting Pre-Service Science Teachers' Planning of Task-Based Classroom Discussions	599
<i>Danielle Ross, Aaron M. Kessler, Jennifer Cartier</i>	
Understanding the Relationships Within and Between Constructs of a Learning Progression: Combining Multidimensional Item Response Modeling and Latent Class Analysis	607
<i>Jinnie Choi, Ravit Golan Duncan</i>	
Students' Use of Evidence and Epistemic Criteria in Model Generation and Model Evaluation	615
<i>Ravit Golan Duncan, Carol Tate, Clark A. Chinn</i>	

Volume 2

Papers (continued)

The Role of Scientific and Social Academic Norms in Student Negotiations while Building Astronomy Models.....	625
<i>Melissa Sunshine Cook, Noel Enyedy</i>	
The Beginnings of Engineering Design in an Integrated Engineering and Literacy Task	633
<i>Mary McCormick, David Hammer</i>	
The Roles of Teacher Questioning in Argument-based Inquiry (ABI): Approaches that Promote Cognitive Thinking and Dialogical Interaction	641
<i>Ying-Chih Chen, Brian Hand</i>	
Hear What They Say and Watch What They Do: Predicting Valid Mathematical Proofs Using Speech and Gesture	649
<i>Elizabeth Pier, Candace Walkington, Caroline Williams, Rebecca Boncoddo, Jessica Waala, Martha W. Alibali, Mitchell J. Nathan</i>	
I Want to be a Game Designer or Scientist: Connected Learning and Developing Identities with Urban, African-American Youth	657
<i>June Ahn, Mega Subramaniam, Elizabeth Bonsignore, Anthony Pellicone, Amanda Waugh, Jason Yip</i>	
Reflecting on Educational Game Design Principles via Empirical Methods	665
<i>Osvaldo Jimenez</i>	
Exploring Group-Level Epistemic Cognitions within a Knowledge Community and Inquiry Curriculum for Secondary Science.....	673
<i>Alisa Acosta, Michelle Lui, James D. Slotta</i>	
High School Students' Parameter Space Navigation and Reasoning during Simulation-Based Experimentation.....	681
<i>Hee-Sun Lee, Amy Pallant, Robert Tinker, Paul Horwitz</i>	
Learner Alignment with Expansive Framing as a Driver of Transfer	689
<i>Diane P. Lam, Adam Mendelson, Xenia S. Meyer, Lloyd Goldwasser</i>	
Game-Enabled Agency: Outcomes that Matter	697
<i>Sasha Barab, Anna Arici</i>	
Model-Based Reasoning: A Framework for Coordinating Authentic Scientific Practice with Science Learning	705
<i>Julia S. Gouvea, Arash Jamshidi, Cynthia Passmore</i>	

The Impact of Text Genre on Science Interest in an Authentic Science Learning Environment.....	713
<i>Steven McGee, Amanda M. Durik, Dena Ann Pastor</i>	
When Experts Disagree: Sourcing Practices While Reading Conflicting Online Information Sources.....	721
<i>Sarit Barzilai, Eynav Tzadok, Yoram Eshet-Alaklai</i>	
Learning and Becoming in an After School Program: The Relationship as a Tool for Equity within the Practices of Making and Tinkering.....	729
<i>Daniela K. DiGiacomo, Kris D. Gutiérrez</i>	
“It’s Intentional”: Co-Construction of Transformational Processes and Pathways within and across Hubs of Interdependence in an Urban Community	737
<i>Joanne Larson, Courtney Hanny, Joyce Duckles, Joel Gallegos Greenwich, Eric Meyer, Robert Moses, George Moses, Kimberly Jones, Jeremy Smith</i>	
Middle School Learners' Ontological 'Trying-on' of Dimensions: A Phenomenological Investigation.....	745
<i>Keri Duncan Valentine, Theodore J. Kopcha</i>	
Towards the Facilitation of an Online Community of Learners: Assessing the Quality of Interactions in Yammer.....	753
<i>Marcela Borge, Sean Goggins</i>	
Communities of Learning Practice: Balancing Emergence and Design in Educational Settings.....	761
<i>Filitsa Dingyloudi, Jan-Willem Strijbos</i>	
Dynamic Visualization of Motion for Student-Generated Graphs	769
<i>Jonathan M. Vitale, Kevin Lai, Marcia C. Linn</i>	
Supporting Teacher Learning for Pedagogical Innovation Through Collaborative Co-Design: Issues and Challenges.....	777
<i>Nancy Law, Johnny Yuen, Yeung Lee</i>	
Inexplicable Silence: An Uncomfortable Analysis of the Social Silences	785
<i>Daniel Steinbock</i>	
Designing for Democracy in Education: Participatory Design and the Learning Sciences	793
<i>Betsy DiSalvo, Carl DiSalvo</i>	
Leveraging the Cultural Practices of Science for Making Classroom Discourse Accessible to Emerging Bilingual Students.....	800
<i>Enrique Suárez, Valerie Otero</i>	
Using Contrasting Video Cases of the Enactment of Cognitively Demanding Science Tasks in Professional Development	808
<i>Miray Tekkumru Kisa, Mary Kay Stein</i>	
Insights Into Teacher Reflective Practice During Planning.....	816
<i>Michael T. Dianovsky, Donald J. Wink</i>	
Analyzing Online Knowledge-Building Discourse Using Probabilistic Topic Models	823
<i>Weiye Sun, Jianwei Zhang, Hui Jin, Siwei Lyu</i>	
Developing an Orchestration Framework for Collective Inquiry in Smart Classrooms: SAIL Smart Space (S3).....	831
<i>Mike Tissenbaum, James D. Slotta</i>	
Bidirectional Analysis of Creative Processes: A Tool for Researchers.....	839
<i>Alecia Marie Magnifico, Erica Rosenfeld Halverson, Christopher T. Cutler, TJ Kalaitzidis</i>	
Learning from Self-Explaining Emergent Phenomena.....	847
<i>Kasia Muldner, Winslow Burleson, Michelene T. H. Chi</i>	

The Programmers' Collective: Connecting Collaboration and Computation in a High School Scratch Mashup Coding Workshop	855
<i>Deborah A. Fields, Veena Vasudevan, Yasmin B. Kafai</i>	
A Tale of Two Worlds: Using Bifocal Modeling to Find and Resolve "Discrepant Events" Between Physical Experiments and Virtual Models in Biology	863
<i>Tamar Fuhrmann, Shima Salehi, Paulo Blikstein</i>	
"Deep Hanging": Mentors Learning and Teaching in Practice	871
<i>Déana Scipio</i>	
Not a Magic Bullet: The Effect of Scaffolding on Knowledge and Attitudes in Online Simulations	879
<i>Ido Roll, Adriana Briseño, Nikki Yee, Ashley Welsh</i>	
Explanations that Make Sense: Accounting for Students' Internal Evaluations of Explanations	887
<i>Shulamit Kapon, Orit Parnafes</i>	
Exploring A Digital Tool for Exchanging Ideas During Science Inquiry	895
<i>Camillia Matuk, Marcia C. Linn</i>	
Framing Sociocultural Interactions to Design Equitable Learning Environments	903
<i>Bryant Jensen</i>	
Learning to Practice Data-Driven Instructional Leadership: Confronting Cultural and Historical Contradictions	911
<i>Raymond Kang</i>	
Tensions and Possibilities for Political Work in the Learning Sciences	919
<i>Angela N. Booker, Shirin Vossoughi, Paula K. Hooper</i>	
"So, I think I'm a Programmer Now": Developing Connected Learning for Adults in a University Craft Technologies Course	927
<i>Deborah A. Fields, Whitney L. King</i>	
Report and Reflection Papers	
Metacognitive Planning and Monitoring: 9th Graders Performing a Long-Term Self-Regulated Scientific Inquiry in A Complex System	937
<i>Billie Eilam</i>	
Investigating the Effect of Curricular Scaffolds on 3rd-Grade Students' Model-Based Explanations for Hydrologic Cycling	942
<i>Laura Zangori, Cory T. Forbes, Christina V. Schwarz</i>	
Exploring How Mobile Technology Provides Inquiry Supports for Middle School Students in Conducting Scientific Practices in a Ubiquitous Learning Context	947
<i>Wan-Tzu Lo, Alex Kuhn, Chris Quintana, Ibrahim Delen, Steven McGee, Jennifer Duck</i>	
Empowering Under-Represented Middle School Youth in Engineering Knowledge and Productive Identity Work	952
<i>Angela Calabrese Barton, Daniel Birmingham, Edna Tan, Takumi Sato</i>	
Becoming a Youth Worker in a Classroom Community of Practice	957
<i>Laurie Ross</i>	
Designing for Engagement in Environmental Science: Becoming "Environmental Citizens"	962
<i>Susan Bobbitt Nolen, Gavin Tierney, Alexandra Goodell, Nathanie Lee, Robert D. Abbott</i>	
Student Regulation of Collaborative Learning in Multiple Document Integration	967
<i>Jun Oshima, Ritsuko Oshima, Keita Fujii</i>	
Evaluating Lesson Design and Implementation within the ICAP Framework	972
<i>Rod D. Roscoe, Pedro J. Gutierrez, Ruth Wylie, Michelene T. H. Chi</i>	

Sequencing Sense-Making and Fluency-Building Support for Connection Making between Multiple Graphical Representations	977
<i>Martina A. Rau, Vincent Aleven, Nikol Rummel</i>	
Taking DALITE to the Next Level: What Have We Learned from a Web-Based Peer Instruction Application?	982
<i>Elizabeth S. Charles, Chris Whittaker, Nathaniel Lasry, Michael Dugdale, Kevin Lenton, Sameer Bhatnagar, Jonathan Guillemette</i>	
Representational Competence and Spatial Thinking in STEM	987
<i>Mike Stieff, Matthew Lira, Dane DeSutter</i>	
Expansive Framing and Preparation for Future Learning in Middle-School Computer Science	992
<i>Shuchi Grover, Roy D. Pea, Stephen Cooper</i>	
Conceptualizing Teacher's Practices in Supporting Students' Mathematical Learning in Computer-Directed Learning Environments	997
<i>Aaron M. Kessler, Melissa D. Boston, Mary Kay Stein</i>	
Teaching Struggling Middle School Readers to Comprehend Informational Text	1002
<i>Donna Caccamise, Angela Friend, Christine Groneman, Megan Littrell-Baez, Eileen Kintsch</i>	
Becoming a Computer Scientist: Early Results of a Near-Peer and Social Justice Program with Latino/a Children	1007
<i>Jill Denner, Jacob Martinez, Heather Thiry</i>	
Science, Technology, Body and Personhood: The Concept of Health Emerging in High-Tech Modern Medicine Practice	1012
<i>Federica Raia, Martin Cadeiras, Ali Nsair, Daniel Cruz, Grecia Ramos, Claire Alvarenga, Valeria M. Rivera, Kristina Barrientos, Mario C. Deng</i>	
"With-Me-Ness": A Gaze-Measure for Students' Attention in MOOCs	1017
<i>Kshitij Sharma, Patrick Jermann, Pierre Dillenbourg</i>	
A Design Inquiry: Bridging Assessment and Curriculum Frameworks to Engage Students in Science Practices	1022
<i>Angela Haydel DeBarger, Erika Tate, Yves Beauvineau, Mingyu Feng, Patricia Schank, Tamara Heck, Michelle Williams</i>	
Design Principles for Motivating Learning with Digital Badges: Consideration of Contextual Factors of Recognition and Assessment	1027
<i>Cathy Tran, Katerina Schenke, Daniel T. Hickey</i>	
What's Happening in the "Quantified Self" Movement?	1032
<i>Victor R. Lee</i>	
The Nature of Student Thinking and Its Implications for the Use of Learning Progressions to Inform Classroom Instruction	1037
<i>Alicia Alonzo, Andrew Elby</i>	
Teaching about Confidence Intervals: How Instructors Connect Ideas Using Speech and Gesture	1042
<i>Elise Lockwood, Amelia Yeo, Noelle Crooks, Mitchell J. Nathan, Martha W. Alibali</i>	
Studying Students' Early-Stage Software Design Practices	1047
<i>Brian A. Danielak, William E. J. Doane</i>	
Can Scaffolds from Pedagogical Agents Influence Effective Completion of Sub-Goals during Learning with a Multi-Agent Hypermedia-Learning Environment?	1052
<i>Michelle Taub, Roger Azevedo, Nicholas Mudrick, Erika Clodfelter, François Bouchet</i>	
Civilian Analogs of Army Tasks: Supporting Pedagogical Storytelling Across Domains	1057
<i>Andrew Gordon, Mark Core, Sin-Hwa Kang, Catherine Wang, Christopher Wienberg</i>	

From Playing a Game to Solving an Equation.....	1062
<i>Shulamit Kapon</i>	
Supporting Conceptual Understandings Outdoors: Findings from the Tree Investigators Mobile Project	1067
<i>Heather Toomey Zimmerman, Susan M. Land, Brian J. Seely, Michael R. Mohny, Gi Woong Choi, Lucy Richardson McClain</i>	
Analyzing University Students' Participation in the Co-Design of Learning Scenarios.....	1072
<i>Iolanda Garcia</i>	
Authenticity Matters: Youth and Science Participation in Design-Based Learning Environments	1077
<i>Gavin Tierney, Déana Scipio</i>	
Personas as a Powerful Methodology to Design Targeted Professional Development Resources.....	1082
<i>Adrian Madsen, Sarah B. McKagan, Eleanor C. Sayre, Mathew "Sandy" Martinuk, Alexander Bell</i>	
Engaging in Educational Design Processes for Sustainable Learning: Learning and Becoming in Practice.....	1087
<i>Mona Holmqvist Olander, Clas Olander</i>	
Supporting Computational Algorithmic Thinking (SCAT): Exploring the Development of Computational Algorithmic Thinking Capabilities in African-American Middle-School Girls	1092
<i>Jakita O. Thomas</i>	
Investigating Student Generated Computational Models of Science	1097
<i>Satabdi Basu, Anton Dukeman, John S. Kinnebrew, Gautam Biswas, Pratim Sengupta</i>	
Perspectival Computational Thinking for Learning Physics: A Case Study of Collaborative Agent-Based Modeling	1102
<i>Amy Voss Farris, Pratim Sengupta</i>	
Children's Use of Inscriptions in Written Arguments About Socioscientific Issues	1107
<i>Sihan Xiao, William A. Sandoval</i>	
An Interactional Analysis of Gaze Coordination during Online Collaborative Problem Solving Activities.....	1112
<i>Murat Perit Cakir, Selin Deniz Uzunosmanoğlu</i>	
Multimedia Educative Curriculum Materials: Designing Digital Supports for Learning to Teach Scientific Argumentation	1117
<i>Suzanna Loper, Katherine L. McNeill, Raphaela Peck, Jeremy Price, Jacqueline Barber</i>	
How Good Is This Evidence? Students' Epistemic Competence in Evidence Evaluation.....	1122
<i>Clark A. Chinn, Leah C.-C. Hung, Randi M. Zimmerman, Ravit Golan Duncan</i>	
Mathematical Tasks as Boundary Objects in Design-Based Implementation Research	1127
<i>Raymond Johnson, Samuel Severance, Heather Leary, Susan Miller</i>	
Evolution of Communities of Learning Practice in Higher Education: Collective Units of Analysis	1132
<i>Filitsa Dingyloudi, Jan-Willem Strijbos</i>	
Recasting the Textbook: Student Creation of Interactive Digital History Texts with Primary Source Documents	1137
<i>Molly B. Zielezinski, Paul Franz</i>	
Sources of Affect around Interdisciplinary Sense Making	1142
<i>Benjamin D. Geller, Julia S. Gouvea, Vashti Sawtelle, Chandra Turpen</i>	
The Impact of Principle-Based Reasoning on Hands-on, Project-Based Learning	1147
<i>Marcelo Worsley, Paulo Blikstein</i>	

Emotional Engagement in Agentive Science Learning Environments	1152
<i>Andrew Morozov, Leslie Rupert Herrenkohl, Kari Shutt, Phonraphee Thummaphan, Nancy Vye, Robert D. Abbott, Giovanna Scalone</i>	
Filling in the Gaps: Capturing Social Regulation in an Interactive Tabletop Learning Environment	1157
<i>Abigail Evans, Jacob O. Wobbrock</i>	
Reverberating Words and GED 2014 Academic Writing Instruction: Reflecting on a Functional Linguistics-Based Approach to Grammar Foregrounding the Social Concept of Identity	1162
<i>Sasha Lotas</i>	
Influence of Public Design Critiques on Fifth Graders Collaborative Engineering Design Work	1166
<i>Michelle E. Jordan</i>	
Tensions in a Multi-Tiered Research-Practice Partnership	1171
<i>Samuel Severance, Heather Leary, Raymond Johnson</i>	

Volume 3

Symposia

Combining Generation and Expository Instruction to Prepare Students to Transfer Big Ideas Across School Topics	1179
<i>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</i>	
The Interplay of Domain-Specific and Domain-General Factors in Scientific Reasoning and Argumentation	1189
<i>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</i>	
Science Sims and Games: Best Design Practices and Fave Flops	1199
<i>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</i>	
Teachers' Learning about Equitable Practice through Talk with Colleagues	1209
<i>Ilana Seidel Horn, Irene Yoon, Britnie Delinger Kane, Nicole Bannister, Elizabeth A. van Es, Victoria Hand, Ilana Seidel Horn</i>	
Motivating and Broadening Participation: Competitions, Contests, Challenges, and Circles for Supporting STEM Learning	1219
<i>Yasmin B. Kafai, Natalie Rusk, Quinn Burke, Chad Mote, Kylie Peppler, Deborah A. Fields, Ricarose Roque, Orkan Telhan, Karen Elinich, Alecia Marie Magnifico</i>	
Research and Design of Learning Experiences for Families	1228
<i>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</i>	
Re-Placing the Body in Children's Learning	1237
<i>Katie Headrick Taylor, Rogers Hall, Jasmine Y. Ma, Ananda Marin, Nathan Phillips</i>	
Connecting Learning and Becoming: Studying Epistemologies and Identities as Interconnected, Dynamic Systems	1247
<i>Eli Gottlieb, Leslie Rupert Herrenkohl, Stanton Wortham, Catherine Rhodes, Martin J. Packer, Martha Rocío Gonzalez, Anna Sfard</i>	

Toward an Argumentative Grammar for Socio-Cultural/Cultural-Historical Activity Approaches to Design Research.....	1254
<i>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</i>	
Mapping the Distribution of Children's Digital Media Practices: Methodological Innovations and Challenges.....	1264
<i>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</i>	
Becoming Reflective: Designing for Reflection on Physical Performances	1273
<i>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</i>	
Enrollment of Higher Education Students in Professional Knowledge and Practices.....	1283
<i>Crina Damşa, Hanni Muukkonen, Sten Ludvigsen, Monika Nerland, Minna Lakkala, Auli Toom, Kari Kosonen, Liisa Ilomäki, Lina Markauskaite, Peter Goodyear, Agnieszka Bachfischer</i>	
Becoming More Mathematical: New Directions for Describing and Designing for Positive Dispositions Toward Mathematics	1293
<i>Melissa Kumar, Tesha Sengupta-Irving, Noel Enyedy, Melissa Gresalfi, Jennifer Langer-Osuna, Anna Sfard, Kris D. Gutiérrez</i>	
Theorizing Learning in the Context of Social Movements.....	1302
<i>A. Susan Jurow, Ben Kirshner, José Antonio Torralba, Sherine El Taraboulsi, Leah Teeters, Barbara Guidalli, Nosakhare Griffin-EL, Samuel Severance, Molly Shea, Erik Dutilly, Rogers Hall</i>	
Concept Formation in Activity	1308
<i>Yrjö Engeström, Annalisa Sannino, Yuri Lapshin, Maria Safronova, Jaakko Virkkunen, Irene Vänninen, Marco Pereira Querol, Rogers Hall</i>	
Cyberinfrastructure for Design-Based Research: Toward a Community of Practice for Learning Scientists	1315
<i>Sharon J. Derry, Alan J. Hackbarth, Sadhana Puntambekar, William A. Sandoval, Carlos Gonzalez, Katherine Bielaczyc, Richard Lehrer, Allan Collins</i>	
Is the Sum Greater than Its Parts? Reflections on the Agenda of Integrating Analyses of Cognition and Learning	1323
<i>Mariana Levin, Orit Parnafes, Timothy Koschmann, Andrea A. diSessa, Reed Stevens, Rogers Hall, Joshua A. Danish, Noel Enyedy</i>	
Reimagining Cultural Forms, Ethnographic Methods and Researcher Responsibilities in Studying Engineering and Science Learning: Honoring and Building on the Work of Margaret Eisenhart	1332
<i>Cory Buxton, Karen L. Tonso, Heidi Carlone, Angela Johnson, Jrene Rahm</i>	
When Friends Argue: Investigating Argumentative Learning Processes in Facebook	1342
<i>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</i>	
Enhancing Self-Regulated Learning through Metacognitively-Aware Intelligent Tutoring Systems.....	1352
<i>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</i>	
History Learning and Teaching Today: Learning What? Becoming What? By What Practice?	1362
<i>Yifat Ben-David Kolikant, Sarah Pollack, Bradley Fogo, Tsafirir Goldberg, D. Kevin O'Neill, Laura D'Amico, Sheryl Guloy, Avishag (Abby) Reisman</i>	

Learning and Thinking in Practice: Complex Systems Thinking "In the Wild"	1372
<i>Izabel Duarte Olson, Ananda Marin, Douglas L. Medin, Alon Hirsh, Sharona T. Levy, Megan Bang, Priya Pugh, Megan McGinty, Uri Wilensky</i>	
Learning as Multi-Dimensional Psychological and Cultural Ecological Spaces.....	1382
<i>Carol D. Lee, Shirin Vossoughi, Kris D. Gutiérrez, Na'ilah Suad Nasir, Maxine McKinney de Royston, Barbara Rogoff</i>	
Combining Video Games and Constructionist Design to Support Deep Learning in Play	1388
<i>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</i>	
Disrupting Learning: Changing Local Practice for Good.....	1396
<i>Jasmine Y. Ma, Charles Munter, Einat Heyd-Metzuyanim, James G. Greeno, Molly Kelton, Rogers Hall, Melissa Gresalfi</i>	
Differing Notions of Responsive Teaching across Mathematics and Science: Does the Discipline Matter?	1406
<i>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</i>	
Making the Most Out of It: Maximizing Learners' Benefits from Expert, Peer and Automated Feedback across Domains.....	1416
<i>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</i>	
Leveraging Educative Approaches to STEM Disciplinary and Instructional Practices	1426
<i>Philip Bell, Jeanne Chowning, Elaine Klein, Veronica McGowan, Tana Peterman, Kerri Wingert, Anna Maria Arias, Elizabeth A. Davis, Annemarie S. Palincsar, April Luehmann</i>	
Learning and Becoming Through Art-Making: Relationships among tools, phenomena, people, and communities in shaping youth identity development	1436
<i>Noel Enyedy, Joseph L. Polman, Cynthia Graville-Smith, Megan Bang, Beth Warren, Ann Rosebery, Jeff Burke, Fabian Wagnister, Amy Bolling, Taylor Fitz-Gibbon, Erica Rosenfeld Halverson, Na'ilah Suad Nasir</i>	
Teacher Facilitation of Whole-Class Discussion in Secondary History Classrooms	1446
<i>Avishag (Abby) Reisman, Chava Shane-Sagiv, Lisa M. Barker, Bradley Fogo, Joseph L. Polman</i>	
Synergistic Scaffolding of Technologically-Enhanced STEM Learning in Informal Institutions	1456
<i>Leilah Lyons, Emma Anderson, Michael Carney, Karen Elinich, Robb Lindgren, Michael Tscholl, Chris Quintana, Jessica Roberts, Joyce Wang, Susan Yoon, Iris Tabak</i>	
Education for Sustainability and Resilience in a Changing Climate	1466
<i>Sameer Honwad, Ofelia Mangen Sypher, Christopher Hoadley, Armanda Lewis, Kenneth Tamminga, Rose Honey, Daniel C. Edelson</i>	
Learning Across Settings: Towards Transformative Trajectories of Practice	1474
<i>Shirin Vossoughi, Meg Escudé, Fan Kong, Elizabeth Mendoza, Molly Shea</i>	
Posters	
What Do They Do?: Tracing Students' Patterns of Interactions within a Game-Based Intelligent Tutoring System.....	1481
<i>Erica L. Snow, G. Tanner Jackson, Danielle S. McNamara</i>	
Implementation Model for Developing Training Measures to Foster Values in an Organization	1483
<i>Sandra Niedermeier, Heinz Mandl</i>	
"It's Not as Bad as Using the Toaster All of the Time": Trade-offs in a Scratch Game about Energy Use.....	1485
<i>Gillian Puttick, Amanda Strawhacker, Debra Bernstein, Elisabeth Sylvan</i>	

Controlling for Statistical Dependencies in CSCL Using General Estimating Equations.....	1487
<i>E. Michael Nussbaum, Gwen C. Marchand</i>	
Between the Lines: The Role of Curriculum Materials and Teacher Language in Communicating Ideas about Scientific Modeling.....	1489
<i>Carrie-Anne Sherwood, Carrie Allen Bemis, Savitha Moorthy, Cynthia D'Angelo, Tina Stanford, Christopher Harris</i>	
Exploring the Role of Theory of Mind and Executive Functions in Preschool Children's Hypothesis Testing and Revision.....	1491
<i>Jamie Liberti, Susan Golbeck</i>	
Learning to Survive “Home-Free”: Compulsory Learning and the Politics of Freight-Hopping Mobility.....	1493
<i>F. Alvin Pearman II</i>	
Sink or Swim: Understanding the Evolution of User Behaviors in an Online Educational Community	1495
<i>Min Yuan, Lei Ye, Mimi Recker</i>	
Using Deficient Models as Scaffolds for Learning Engineering Concepts of Tradeoffs and Optimization	1497
<i>Chandan Dasgupta, Tom Moher</i>	
Intersections of Science Learning and Language Development within Scientific Argumentation: Implications for English Language Learners.....	1499
<i>Maria González-Howard, Katherine L. McNeill</i>	
Exploring the Use of Elaborative Interrogation in an Introductory Physics Course	1501
<i>Robert Zisk, Elana Resnick, Eugenia Etkina</i>	
Fusing a Crosscutting Concept, Science Practice, and a Disciplinary Core Idea in Single Learning Progression.....	1503
<i>Hayat Hokayem, Amelia Wenk Gotwals</i>	
Redefining Engagement and Participation: The Co-Construction of Student Learning Practices	1505
<i>Simona Goldin, Michaela Krug O'Neill, Shweta Naik</i>	
Reflective Decision Making within the Discourse of Urban Elementary Engineering Classrooms	1507
<i>Christopher G. Wright, Kristen B. Wendell, Patricia Paugh</i>	
Adventure Learning @ the Learning Sciences	1509
<i>Brant G. Miller, R. Justin Hougham, Christopher Cox, Von Walden, Karla Bradley Eitel</i>	
Visualizing Three-Dimensional Spatial Relationships in Virtual and Physical Astronomy Environments.....	1511
<i>Patricia Udomprasert, Alyssa A. Goodman, Philip Sadler, Erin Lotridge, Jonathan Jackson, Ana- Maria Constantin, Zhihui Helen Zhang, Susan Sunbury, Mary Dussault</i>	
Upper-Level Physics Students' Perceptions of Physicists.....	1513
<i>Paul W. Irving, Eleanor C. Sayre</i>	
Designing and Validating a Story-Based Socio-Emotional Learning Assessment Instrument	1515
<i>Mitra Fatolapour, Soyeon Hwang, Mario Piergallini, Julie Shim, Steven Dow, Carolyn Penstein Rosé</i>	
Detecting Iterative Cycles of Engineering Design from Student Digital Footprints in Computer-Aided Design Software.....	1517
<i>Zhihui Helen Zhang, Charles Xie, Saeid Nourian</i>	
Enculturation: Contemporary Use in the Learning Sciences from a Historical Perspective	1519
<i>Ornit Sagy, Yotam Hod</i>	
How do Children Draw, Describe, and Gesture about Motion?	1521
<i>Donna Kotsopoulos, Michelle Cordy, Melanie Langemeyer, Laaraib Khattak</i>	

Design Principles for Science Laboratory Instruction Using Augmented Virtuality Technologies.....	1523
<i>Crystal J. DeJaegher, Jennifer L. Chiu, Jie Chao</i>	
The Role of Feedback in Interest Development in an Out-of-School Engineering Setting.....	1525
<i>Joseph E. Michaelis, Mitchell J. Nathan</i>	
Comprehension SEEDING: Providing real-time formative assessment to enhance classroom discussion	1527
<i>Ruth Wylie, Michelene T. H. Chi, Robert Talbot, Erik Dutilly, Susan Trickett, Brandon Holding, Rodney D. Nielsen</i>	
Conditions for Successful Learning-by-Teaching: Lessons Learned in Training Prospective Science Teachers	1529
<i>Moseli A. Mafa</i>	
Learners' Intuitions about Geology	1531
<i>Lauren Barth-Cohen, Jonathan Shemwell, Daniel Capps</i>	
Development of Integrated Physics Identity through Physics Learning Assistant Experience	1533
<i>Eleanor W. Close, Jessica Conn, Hunter G. Close</i>	
Assessment Analytics in CSCL: Activity Theory Based Method	1535
<i>Wanli Xing, Robert R. Wadholm, Sean Goggins</i>	
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
<i>Kathryn Lanouette, Eric Berson, Kathleen E. Metz</i>	
Cognitive Ethnographies of Heterogeneous Engineering Design	1539
<i>Carlye Lauff, Joanna Weidler-Lewis, Kevin O'Connor, Daria A. Kotys-Schwartz, Mark E. Rentschler</i>	
The Role of Stated Relationships in Detecting Contradictions Between Multiple Representations in Science	1541
<i>Candice Burkett, Susan R. Goldman, M. Anne Britt</i>	
Improving Online Collaboration by Fostering Norm-Oriented Content Based Knowledge Awareness.....	1543
<i>Michail D. Kozlov, Tanja Engelmann, Richard Kolodziej, Roy B. Clariana</i>	
Fostering Awareness Content Creation by Self-Determined Regulation	1545
<i>Tanja Engelmann, Katrin König, Michail D. Kozlov</i>	
Learning Integrated STEM Using Tangible Agent-Based Modeling	1547
<i>Gokul Krishnan, Pratim Sengupta</i>	
Linked Reading and Writing Using Wikilinking: Promoting Knowledge Building within Technology-Enhanced Classroom Learning Communities.....	1549
<i>Tamar Novik, Dani Ben-Zvi, Yotam Hod</i>	
Building A Framework For the Process of Crafting and Using Definitions.....	1551
<i>Angela Little</i>	
Re-grow Your City: A NetLogo Curriculum Unit on Regional Development.....	1553
<i>Arthur Hjorth, Uri Wilensky</i>	
Colors of Nature: Connecting science and arts education to promote STEM-related identity work in middle school girls.....	1555
<i>Carrie Tzou, Laura Conner, Stephen Pompea, Mareca Guthrie</i>	
Reel Science: Identity Development through Filmmaking.....	1557
<i>Rachel Chaffee</i>	
Identifying and Assessing Computational Thinking Practices	1559
<i>Wendy Martin, Karen Brennan, William Tally, Francisco Cervantes</i>	

Undergraduate Attitudes Towards Help-Seeking	1561
<i>Iris Howley, Carolyn Penstein Rosé</i>	
Becoming a Professional through Virtual Practice	1563
<i>Tabitha McKay, Andrea Cantarero, Vanessa Svihla, Elizabeth Yakes Jimenez, Tim Castillo</i>	
Community-Based Engineering and Novice Elementary Teachers' Knowledge of Engineering Practices	1565
<i>Kristen B. Wendell, Tejaswini Dalvi</i>	
Using Classroom-Based Authentic Research Experiences to Foster Scientific Thinking and Representational Competence	1567
<i>Kristy L. Halverson, Kari L. Clase</i>	
Scaffolding Argumentation Competence: The Shift from First to Second Order Skill Acquisition	1569
<i>Omid Noroozi, Paul A. Kirschner, Harm J. A. Biemans, Martin Mulder</i>	
Multiple-Text Processing in Text-Based Scientific Inquiry	1571
<i>Katherine James, Susan R. Goldman, Mon-Lin Monica Ko, Cynthia Greenleaf, Willard Brown</i>	
Building a Learning Progression for Chromosome Segregation Using Phenomenographic Variation Theory	1573
<i>Stanley M. Lo, Stephanie Kim, Su Swarat, Gregory J. Light</i>	
Tools for Sustained Student Engagement in InterLACE (Interactive Learning and Collaboration Environment)	1575
<i>Chris Teplovs, Leslie Schneider</i>	
“Are You ‘In’ or Are You ‘Out’?” Investigating the Factors Affecting Immersion in a Location-Based AR Game for IBSE	1577
<i>Yiannis Georgiou, Eleni A. Kyza</i>	
Collaborative Hypothesis-Building Using Immersive Virtual Environments for Ecosystems Science	1579
<i>Shari Metcalf, Amy Kamarainen, Tina Grotzer, Chris Dede</i>	
Understanding Data Variability in Ecosystems: Blending MUVE and Mobile Technologies to Support Reasoning with Real World Data	1581
<i>Amy Kamarainen, Shari Metcalf, Tina Grotzer, Chris Dede</i>	
Identifying Affordances of 3D Printed Tangible Models for Understanding Core Biological Concepts	1583
<i>Jodi L. Davenport, Matt Silberglitt, Jonathan Boxerman, Arthur J. Olson</i>	
Comparison of Specific and Knowledge Integration Automated Guidance for Concept Diagrams in Inquiry Instruction	1585
<i>Kihyun Ryoo, Marcia C. Linn</i>	
Elementary Students Becoming Engineers through Practice	1587
<i>Cathy P. Lachapelle, Christine Gentry, Jonathan D. Hertel, Christopher San Antonio-Tunis, Christine M. Cunningham</i>	
Local Ground: A Toolkit Supporting Metarepresentational Competence in Data Science	1589
<i>Sarah Van Wart, Tapan Parikh</i>	
Make to Relate: Narratives Of, and As, Community Practice	1591
<i>Colin Dixon, Lee Martin</i>	
Gesture Enhancement of a Virtual Tutor via Investigating Human Tutor Discursive Strategies: Forms and Functions for Proportions	1593
<i>Virginia J. Flood, Alyse Schneider, Dor Abrahamson</i>	
Beyond ‘Solve for x’: Integrating Equations with Conceptual Understanding	1595
<i>Matthew Lira</i>	

Purposeful Learning across Collaborative Educational Spaces.....	1597
<i>Teresa Cerratto Pargman, Nuno Otero, Marcelo Milrad, Daniel Spikol, Ola Kutsson, Robert Ramberg</i>	
Taking a New Perspective on Spatial Representations in STEM	1599
<i>Dane DeSutter, Mike Stieff</i>	
The Power of Networks as an Engineering Sophomore	1601
<i>Janet Y. Tsai, Daria A. Kotys-Schwartz, Daniel W. Knight</i>	
Getting your Drift: Activity Designs for Grappling with Evolution.....	1603
<i>Corey Brady, Michael Horn, Uri Wilensky, Aditi Wagh, Arthur Hjorth, Amartya Banerjee</i>	
Modeling the Dynamics of Ontological Reasoning in Physics	1605
<i>Lele Mathis, Ayush Gupta</i>	
Finding Productivity in Design Task Tinkering	1607
<i>Gina Quan, Ayush Gupta</i>	
Making Mathematical Meaning through Robot Enactment of Mathematical Constructs	1609
<i>Scot Sutherland, Tobin White, Jason Huang, Harry Cheng</i>	
Predicting Performance Behaviors during Question Generation in a Game-Like Intelligent Tutoring System.....	1611
<i>Carol Forsyth, Arthur C. Graesser, Borhan Samei, Brea Walker, Philip Pavlik, Jr.</i>	
Promoting Science Identification and Learning through Contemporary Scientific Investigations Using Practice-Focused Instruction.....	1613
<i>Katie Van Horne</i>	
Creating Material Representations of Practice at the Boundary of Professional Development and Classroom Practice.....	1615
<i>Scott McDonald, Jessica Thompson</i>	
Examining How Students Make Sense of Slow-Motion Video.....	1617
<i>Min Yuan, Nam Ju Kim, Joel Drake, Scott Smith, Victor R. Lee</i>	
Identity, Digital Learning Environments and Academic Success	1619
<i>Mirlanda E. Prudent</i>	
Science Teacher Pedagogical Design Capacity with Technology in an Integrated Teacher Preparation Program.....	1621
<i>Aaron M. Kessler, Jennifer Cartier</i>	
Family Creative Learning: Engaging parents and children as learning partners in creative technology workshops	1623
<i>Ricarose Roque, Natalie Rusk, Luisa Beck, Xiaodi Chen</i>	
Teachers Becoming (Temporary) Engineers to Become Better Teachers	1625
<i>Ayesha Livingston, Jamie Collins, Ara Kooser, Vanessa Svihla</i>	
Contradictions on the Process of Becoming a Physics Teacher	1627
<i>Glauco Silva, Alberto Villani</i>	
Women Becoming Engineers	1629
<i>Joanna Weidler-Lewis</i>	
Characterizing Teachers' Analysis of Student Work	1631
<i>Jason Silverman, Valerie Klein, Wesley Shumar, Cheryl Fricchione</i>	
Beyond Databases: Librarians in a Project-Based Language Arts Curriculum	1633
<i>Sarah Evans</i>	

Characterizing Teachers' Support of Modeling Practices in Science Classrooms	1635
<i>Deborah Peek-Brown, Shawn Stevens, LeeAnn Sutherland, Sung-Youn Choi, Namsoo Shin, Joseph Krajcik</i>	
Advancing Epistemological Frame Analysis to Refine Our Understanding of Inquiry Frames in Early Elementary Interviews	1637
<i>Alejandro Andrade-Lotero, Joshua A. Danish</i>	
Capturing Qualities of Mathematical Talk via 'Coding And Counting'	1639
<i>Einat Heyd-Metzuyanim, Michal Tabach, Talli Nachlieli, Carolyn Penstein Rosé</i>	
Kinecting in Physics: Student Conceptualization of Motion Through Visualization	1641
<i>Janice Anderson, Steven Wall</i>	
"These are Facts": Opportunities for and Barriers to Policy Changes that Support Learning	1643
<i>Abigail Stiles, Julie Bryant, Kersti Tyson, Vanessa Svihla</i>	
Tug of War: What is it Good For? Measuring Student Inquiry Choices in an Online Science Game	1645
<i>Nicole R. Hallinen, Julius Cheng, Min Chi, Daniel L. Schwartz</i>	
Educational Games in the Classroom: Design-Based Research and Methods for Classroom Mediation	1647
<i>Rachel S. Phillips, Theresa Horstman, Carmen Petrick Smith, Christy Ballweber, Noelle Conforti Preszler, Nancy Vye, John Bransford</i>	
"What in the world?" Animated Worlds in Multivariable Modeling with Motion Chart Graph Arguments.....	1649
<i>Jennifer Kahn</i>	
Productive Disciplinary Engagement: Examining Negotiation of Group Activity with Multiple Frameworks.....	1651
<i>Debra Gilbuena, Marja-Liisa Makela, Tuuke Iiskala, Simone Volet, Susan Bobbitt Nolen, Milo Koretsky, Marja Vauras</i>	
Everyday Life Science and Engineering: Bridging the Gap Between Formal and Informal Learning among Native American Students in the Northwestern United States	1653
<i>Marcie Galbreath, Rose Honey, Sameer Honwad, Anne Kern, Chris Meyer, Laura Laumatia</i>	
Distributed Cognition and Gesture: Propagating a Functional System Through Impromptu Teaching.....	1655
<i>Robert F. Williams, Simon Harrison</i>	
Two Systems, Two Stances: A Novel Theoretical Framework for Game-Based Learning	1657
<i>Mario M. Martinez-Garza, Douglas B. Clark</i>	
Keeping Up: Shifting Access to Gateway Resources in a Cycling Community of Practice	1659
<i>Joel Drake, Victor R. Lee</i>	
Examining Teacher Assignments and Student Work at the Intersection of Content and Practice in Middle School Science	1661
<i>Britte Haugan Cheng, Jeremy Fritts, Tiffany Leones, Bowye Gong</i>	
Examining the Use of Technology: Affordances and Constraints in a Blended Learning Environment	1663
<i>Déana Scipio, Annie Camey Kuo</i>	
Demystifying Success in a Summer Bridge Program: Investigating Students' Intrinsic Motivation and Mastery Goals in the Context of a Learning Analytics Intervention	1665
<i>Steven Lonn, Stephen Aguilar, Stephanie D. Teasley</i>	
Designing Collaborative Learning Activities for Two Outcomes: Deep Structural Knowledge and Idea Generation.....	1667
<i>Rachel J. Lam</i>	

Promoting Diversity within the Maker Movement in Schools: New Assessments and Preliminary Results.....	1669
<i>Paulo Blikstein, Vivian Chen, Andrew Martin</i>	
Talk as a Window into Collaborative Lesson Design: Designing a Common Rubric in an Elementary School Work Circle.....	1671
<i>Kimberley Gomez, Nicole Mancevice, Ung-Sang Lee, Jahneille Cunningham</i>	
What Is and Who Can Do Science? Supporting Practice-Linked Identities in Science for Racially/Ethnically Underrepresented Youth	1673
<i>Tammie Visintainer</i>	
Workshops	
Current Research and Practice on Learning Communities: What We Know, What are the Issues, and Where Are We Going?.....	1677
<i>Katerine Bielaczyc, Dani Ben-Zvi, Yotam Hod</i>	
Analytics for Learning and Becoming in Practice	1680
<i>Golnaz Arastoopour, Simon Buckingham Shum, Wesley Collier, Paul A. Kirschner, Simon Knight, David Williamson Shaffer, Alyssa Friend Wise</i>	
Social, Motivational and Affective Dimensions of Learning through Social Interaction	1684
<i>Christa S. C. Asterhan, Sherice N. Clarke</i>	
Exposing and Assessing Learners' Epistemic Thinking	1686
<i>Maggie Renken, Clark A. Chinn, Penelope Vargas, William A. Sandoval</i>	
MOOCShop 2014	1691
<i>Steven Lonn, Christopher Brooks, Zachary A. Pardos, Barry Peddycord III, Emily Schneider, Ido Roll, Ashley Shaw</i>	
Mediated Action and Mediated Discourse Analysis: Studyig Learning and Becoming at the Nexus of Practice.....	1692
<i>Ingrid de Saint-Georges, Kevin O'Connor</i>	
Interaction Analysis of Student Teams Enacting the Practices of Collaborative Dynamic Geometry.....	1697
<i>Gerry Stahl</i>	
Writing Competitive Proposals for Programs in NSF's Division of Research in Learning in Formal and Informal Settings.....	1698
<i>Ellen McCallie, Christopher Hoadley, Michael J. Ford</i>	
Tightening Research-Practice Connections: Taking ISLS Findings to Public Debate.....	1700
<i>Susan McKenney, Kimberley Gomez, Brian J. Reiser</i>	
Constructing Assessment Items that Blend Core Ideas and Science Practices.....	1703
<i>Angela Haydel DeBarger, Joseph Krajcik, Christopher Harris</i>	
NAPLeS: Networked Learning in the Learning Sciences	1704
<i>Freydis Vogel, Frank Fischer, Daniel Sommerhoff</i>	
Designing for Student Agency and Authority around Issues of Climate Change	1709
<i>Victoria Hand, Leilah Lyons, Chrystalla Mouza, Elizabeth Walsh</i>	
The Learning Theater: Designing a Flexible Reconfigurable Space for Ambitious Learning and Teaching on Campus.....	1709
<i>Gary Natriello, Hui Soo Chae</i>	
Research-Practice Partnerships Workshop	1710
<i>William R. Penuel, Philip Bell</i>	

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

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
<i>Haiwen Chu</i>	

Designing Culturally Relevant Technology Based Learning Environments Across Formal and Informal Spaces	1718
<i>Sameer Honwad</i>	

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

Exploring Learning in Out-of-School Settings	1720
<i>Nan Renner</i>	

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

Designing for Persistence after Failure.....	1722
<i>Cathy Tran</i>	

Toward a Multimodal Semiotics of Mathematics Teaching and Learning.....	1723
<i>Daniel Ginsberg</i>	

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

Leadership Development for Collaborative Educational Data Use	1725
<i>Emily S. Lin</i>	

Making Learning: Leverage Makerspaces as Learning Environments.....	1726
<i>Breanne K. Litts</i>	

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

Early Career Workshop

Digital Media for Learning and Assessment	1731
<i>Jody Clarke-Midura</i>	

Research Summary: Simulations and Games for Science Learning.....	1732
<i>Cynthia D'Angelo</i>	

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

Social Aspects of Legitimate Peripheral Participation	1734
<i>Julia Eberle</i>	

Scaffolding Undergraduates' Argumentative Reasoning	1735
<i>Julia Gressick</i>	

Transforming Technological and Scientific Practice: A Research Summary	1736
<i>Katie Headrick Taylor</i>	

Searching for the 'Dialogic' within the Dialogue: Temporal Analysis of Teacher Orchestration in Technology Enhanced Elementary School Classrooms	1737
<i>Andrew Joyce-Gibbons</i>	
Undergraduate Students' STEM-Based Computing Practices	1738
<i>Alejandra J. Magana</i>	
Integrating the Epistemic, Conceptual, and Social Aspects of Science in Elementary School Instruction.....	1739
<i>Eve Manz</i>	
Adding a Human Touch to Personalized Digital Learning.....	1740
<i>Amy Ogan</i>	
Learning and Becoming in Practice	1741
<i>Jimmy Scherrer</i>	
Climate Science in Three Part Harmony: Conceptual, Epistemological and Social Learning Gains.....	1742
<i>Asli Sezen-Barrie</i>	
Epistemic Thinking and Evidence-based Reasoning with Multiple Information Sources in Health Decision-Making.....	1743
<i>Shira Soffer-Vital</i>	
Listening and Learning in Education	1744
<i>Kersti Tyson</i>	
Supporting Equitable Participation in Socially Relevant Science: Scientific Practice, Partnerships and Technological Innovation	1745
<i>Elizabeth Walsh</i>	
Expressive Technologies for Middle School Math and Science	1746
<i>Michelle Hoda Wilkerson-Jerde</i>	
Doctoral Consortium	
Learning to Teach Elementary Students to Construct Evidence-Based Claims about Natural Phenomena.....	1749
<i>Anna Maria Arias</i>	
Attributions and Epistemology in Conversation: How Math Tutors and Students Co-Construct Accounts of Failure and Knowledge.....	1750
<i>David DeLiema</i>	
Identity and Collaborative Learning: Examining the Microdynamics of Social Positioning in Mathematical Group Work	1751
<i>Lesley Dookie</i>	
Towards a Theoretical Model of Immersion in Inquiry-Based Science Learning.....	1752
<i>Yiannis Georgiou</i>	
Co-Constructing Opportunities for Inquiry: A Cross-Context and Development Ethnography of Young Children's Inquiry.....	1753
<i>Danielle Keifert</i>	
Learning to Notice: Supporting Students' Meaningful Engagement in Scientific Practices.....	1754
<i>Abraham S. Lo</i>	
Collaborative Learning: An Effective Component of Productive Failure?	1755
<i>Claudia Mazziotti</i>	
Drawing and Dissection: Improving Understanding in Anatomy	1756
<i>Dimitrios Panagiotopoulos</i>	

Have I Become What I Once Beheld? Identity Construction of the Other in Virtual Worlds	1757
<i>Natasha Anne Rappa</i>	
The Role of Evidence and Models in Science Education	1758
<i>Ronald W. Rinhart</i>	
Exploring Culturally Responsive Computing Education: Learning with Electronic Textiles in an American Indian Community School	1759
<i>Kristin Anne Searle</i>	
Play Your Way Into Math: Supporting Prevocational Students in a Computer Game-based Learning Environment.....	1760
<i>Judith ter Vrugte</i>	
Outside the Margins: Examining the Identities, Communities, and Contexts of Youth Who Attend Alternative Schools.....	1761
<i>Gavin Tierney</i>	
Design-Based Learning in a Community-Based Youth Program: Affordances for Youth Learning and Development.....	1762
<i>Steven Worker</i>	
Children's Use of Inscriptions in Argumentation About Socioscientific Issues	1763
<i>Sihan Xiao</i>	

Papers

(Continued from Volume 1)

The Role of Scientific and Social Academic Norms in Student Negotiations while Building Astronomy Models

Melissa Sunshine Cook and Noel Enyedy, University of California, Los Angeles, 2323 Moore Hall
Los Angeles, CA 90095-1521, enyedy@gseis.ucla.edu, cook.ucla@gmail.com

Abstract: This paper contrasts two groups of students as they build models to explain the phases of the moon. The group with a more successful model and more coherent final explanation judged their own work against scientific norms explicitly throughout their building process, while the less successful group remained focused on social and academic norms as they evaluated their work. When students explicitly invoked particular norms, subsequent arguments tended to coalesce around the explicitly identified norm. Particular strategies for dispute resolution aligned students with either scientific or social academic norms. We suggest the need to make norms and the negotiations around norms visible for students and to encourage science students to use scientific norms when evaluating their work.

Introduction

In this paper we examine peer interaction in two groups of children attempting to build models of the moon-sun-earth system that can explain the phases of the moon. One group's model was quite successful; in a series of presentations multiple group members were able to use the model to accurately explain a new moon, full moon, and the waxing and waning quarter moons. In addition to the model's suitability for conveying key ideas about the earth-moon-sun system, the social process of building the model was an effective platform for collaborative learning: every group member was able to produce an accurate written explanation of the phases of the moon by the time they completed their models. The other group, in contrast, built a static model that only worked to explain a quarter moon. Their verbal explanation of their model was also plagued by confusion between the conditions for a new moon and the conditions for a lunar eclipse. At the end of the project, only half the group members were able to explain the phases of the moon on paper.

In this paper we argue that the trouble experienced by the less successful group comes from a tension between the social/academic norms of their school at large and the specifically scientific norms that are also invoked in the double-layered social space of the science classroom. We will show students in the less successful team "talked past each other" by speaking from different sets of norms whereas students in the more successful team managed to engage more productively in debates that drew on social/academic norms at some times and in science-based norms at other times. We will point out particular moves made by students in the successful group to increase intersubjectivity, make the grounds of disputes explicit, and enable them to reach flexible resolutions more quickly. Our primary aim here is to use examples of small group discourse to illustrate how these two sets of norms operated and sometimes conflicted, the active role children took in negotiating which norms were relevant to their work, and the consequences of particular orientations towards these sets of norms for children's learning.

Background

Academic researchers typically conceive of children negotiating different sets of social norms and values as they move between different social spaces through the course of a typical day, such as the transition from home to school and back (e.g. Heath, 1983). But even within one context, multiple sets of norms and values may be at play. For instance, within a science classroom students may experience tension as they try to make moves that are valuable within peer culture on the one hand or within scientific culture on the other (Brown, 2004, Enyedy, Danish, & Fields, 2011). We adapt a distinction made by Cobb, Stephan, McClain, and Gravemeijer (2001) to identify social academic norms and science/engineering specific norms. Social academic norms delineate expectations for students and the teacher irrespective of the subject at hand, whereas science-specific norms derive from scientific discourse and are specific to work in science class.

Additionally, we recognize that students' modeling activities are heavily mediated by and negotiated with their peers. Peers offer recommendations, critiques, and demand that others be accountable to norms that directly or indirectly renegotiate the various norms of behavior that have been established within the classroom. Through the experience of building a representation, students make use of a variety of resources including norms about representation and particular skills and practices related to modeling. Danish and Enyedy (2007) showed that students appear to respond to local contingencies to decide which norms are relevant to the decisions they are faced with as they create representations. We build on this notion in the current paper by looking closely at the ways in which interactional moves are aligned with or undercut the norms and engineering practices of two groups of students. Our goal is to examine how each group's alignment with particular norms affects their cooperation, their learning, and the model they create.

Method

Setting and Curricular Unit

Data for this study come from a classroom of fifth grade students in a progressive laboratory school in an urban center of Southern California with an ethnically and socio-economically diverse student body that roughly mirrored the population of the State of California. The school uses the Reggio Emilia approach to education, stressing the importance of making and evaluating representations in a broad range of educational activities. On a day-to-day basis at REA, the central way students were evaluated by teachers was through their ability to contribute in valued ways to whole-class discussion. Instead of grades, report cards at REA include detailed assessments of how students are performing in specific areas of academic and personal development.

Data come from a project in which students were challenged to build a model that could explain the change in appearance of the moon from our vantage point on Earth. For two weeks, students marked the position and appearance of the moon in the sky at night as homework. In groups of 4, students compared their observations to discover that the moon looks very different in the sky from night to night. Each group discussed their own theories as to why that might be and devised a brief skit to demonstrate their own theory of why the moon looks different. Initial explanations demonstrated a variety of conceptions about the apparent changes in the moon. For instance, one group expressed the idea that the moon is just covered up with clouds on some nights while another suggested that the moon folds in upon itself to become a thin sliver one night and re-expands to a large ball another night. Some groups included a sun in their model, some included an earth, and some included an observer. All groups had one student playing the moon, but other “characters” also appeared, such as clouds, stars, an observer on the earth, and a narrator. After expressing initial ideas in these models, each group was asked to write a list of questions they would need to answer in order to best explain why the moon looks different on different days or nights. Finally, each group was given the moon model challenge in which they were presented with one of their own observation drawings that they were challenged to explain with their model. They were given and a planning sheet to help them organize their ideas for model building, a variety of craft supplies, and a list of web resources that might be helpful for research.

Participants

For the present analysis we selected 2 student groups: Team Orbiting Moon, who constructed an especially successful model and in which all group members displayed an excellent understanding of the phases of the moon by the end of the project; and team Fixed Moon, whose model was less successful and in which half of the group was still unable to explain the phases of the moon accurately after the unit. The teacher assembled groups to be gender balanced (2 boys, 2 girls) and heterogeneous with regard to students’ prior science performance. Both groups of students began their task with comparable background knowledge: each group had two members who responded to the pre-interview question on why the moon looks different with some version of “I don’t know” while 2 other group members had partial understandings of the mechanisms behind the phases of the moon. Both groups spent a similar proportion of their time “on task” (working on some aspect of planning or building their model). Their interactions with the teacher were similar in nature: she visited each group about half a dozen times during the build and her visits typically refocused them from off-task conversations.

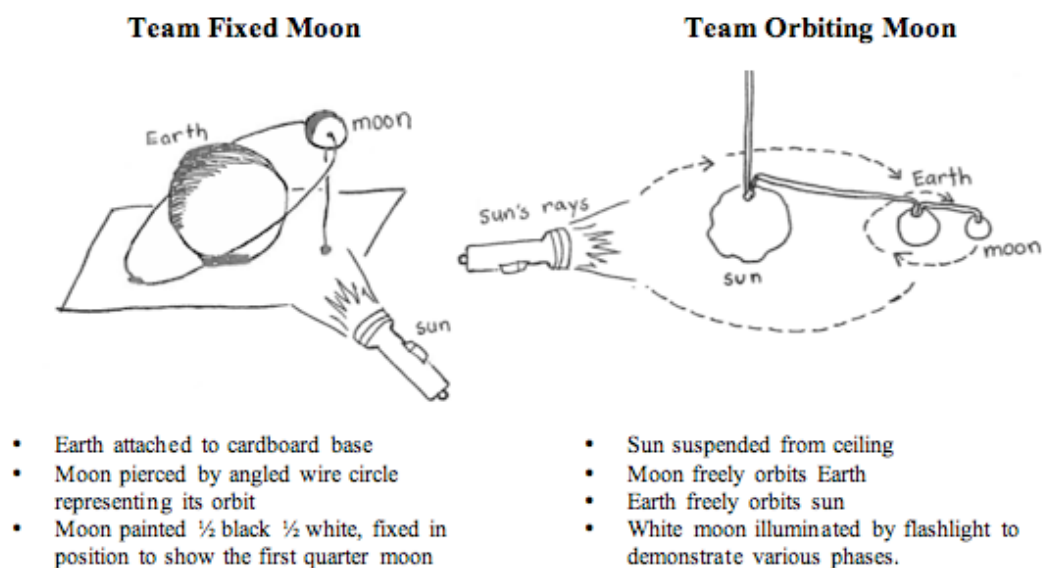


Figure 1. A comparison of Team Fixed Moon and Team Orbiting Moon’s models

Team Fixed Moon has made a few representational commitments that limit their ability to successfully depict all the phases of the moon (see Figure 1). They attached their moon to the cardboard base and to the Earth so that if they want to represent a different phase of the moon, they have to either move the sun (giving the unfortunate illusion that the sun rotates around the earth) or break the model and rotate the moon, in which case the side they painted dark becomes incorrectly positioned relative to the sun. Additionally, while presenting the model, more than one group member makes a critical mistake which belies basic confusion over how the phases of the moon work. Two of the students spoke of the shadow of the earth blocking the moon from the sun's light (a lunar eclipse) as the cause of new moon. The alignment of the planetary bodies they used to illustrate a new moon would actually produce a full moon. This confusion is common because the motions that produce eclipses are complex and the variable angles of the moon's orbit relative to the earth's orbit is rarely well-depicted.

Team Orbiting Moon's model, in contrast, was better able to represent multiple phases of the moon because the moon could be rotated around the earth freely and because the moon was not painted, allowing for a light and shadow to change as the moon orbited the earth. More members of Team Orbiting Moon skillfully explained this model during their final presentation without any major mistakes. Finally, all of Team Orbiting Moon's participants could successfully explain the cause of a new moon and a quarter moon in their individual final interviews after the unit was complete.

Data and Analysis

There were three sources of data for this project. Classroom video of students presenting their final models was coded to indicate the sophistication and scientific accuracy of the model and the explanation students gave of their model. Individually administered written and oral pre-post questions prompted students to explain the phases of the moon and were coded for accuracy so that we could find out which group members could successfully explain the phases of the moon on their own. Finally, classroom videos from the 3 key lessons of this unit (an exploring day, a planning day, and a building day) were recorded. We analyzed these recordings for student reinforcement of norms.

During the flow of work, norms are not always apparent to an observer. Participants make norms visible when they evoke them in evaluating themselves and their work. Therefore, we began by identifying all the times norms were invoked in each group across the classroom footage to identify any/all the times students made judgments or evaluations of their work. Norms about how students should contribute to class surfaced in students' corrections, compliments, and complaints to each other as well as in more formal evaluations such as their final presentations. Students may explicitly flag the set of values by which they are judging contributions to the model or may make comments that implicitly suggest the student is speaking from a particular set of values as they make their judgment.

These evaluations often sparked discussions and arguments about the work as students disagreed either with each other's evaluations of the model or with proposals to improve it. We termed these conversations in which the students explicitly debated judgments of their own work (or the social process by which that work was being accomplished) "evaluative exchanges." Evaluative exchanges differed from other on-topic working conversations because they involved critique: someone's contribution was being judged against some set of norms. We identified the 31 evaluative exchanges that took place within the three selected lessons. Within each exchange, we identified all of the norms being invoked by students as they made their judgments. The specific norms identified were then grouped into two categories: social academic norms or Science-specific norms. This parallels a distinction commonly made in mathematics education research between social norms and sociomathematical norms (Cobb, Stephan, McClain, and Gravemeijer, 2001). Social academic norms help participants coordinate school-based activities by delineating expectations for students and the teacher. These norms would not be expected to differ substantively if the teacher and students were working on an art project or history lesson rather than constructing scientific models. Science-specific social norms, on the other hand, outline a set of shared expectations for scientific work in the classroom. In a prior analysis of a classroom at this school, Cook (2011) found that some of students' science-specific classroom norms echoed the commitments of formal scientific discourse in the adult world while others stemmed from popular images of science and scientists. Therefore, to determine if a norm was science-specific, we considered first whether it would be relevant in a non-science lesson and secondly whether it corresponded to the core commitments of scientific discourse and/or popular notions of what being scientific means. Occasionally group members made it explicit which set of norms guided their evaluation (see science-specific norm 1 below for an example). In these cases, we were able to bolster our analytical judgments with participants' expressed perspectives.

In some evaluative exchanges, more than one set of norms were referenced during the course of the discussion as the students evaluated their work in terms of both social and scientific expectations. In most of those cases, exchanges drew more heavily on one set of norms than on others and were classified in terms of the set of norms underpinning the majority of evaluations within the exchange. One exchange (discussed in the subsequent section) was identified in which students drew equally on science and school norms. Two researchers reviewed all of the exchanges and agreed on their classification.

In this paper, we selected one evaluative exchange from each team for a closer qualitative analysis. To make this selection, we looked for the moment at which each group made the key representational decisions that led to the differences between their final models. For the selected conversations, we created vignettes transcribing and describing each argument. We then analyzed these vignettes them in terms of the way students alluded to the norms that influenced them, the rhetorical moves they employed, and the implicit rules for resolution the group seemed to be operating under as they worked together. After comparing the selected evaluative exchanges on these terms, we looked back across all the exchanges to see if similar patterns were evident.

Results

Types of Norms Referenced by Students

This list of key Science-specific and social academic norms is not exhaustive but represents the most frequently referenced norms in the group work we reviewed.

Table 1. Most frequently referenced norms during the moon model challenge.

Science-specific Norms
1) <i>Judge ideas by how well they correspond with your scientific prior knowledge.</i> Example of norm in use: Atticus made a perfectly reasonable suggestion, from an engineering perspective, when he suggested that the group could spin the moon, let it twist, and then let it automatically untwist itself when it naturally began to spin in the opposite direction. His idea was rejected by two group members, Iman who said “I am pretty sure the moon doesn’t go like this then go like this [pantomiming clockwise and then counterclockwise rotation] and Marie, who explicitly framed her objection to this idea as a science-based argument by sarcastically saying “Ooooooh so scientific”
2) <i>Prioritize scientific accuracy over other engineering concerns.</i> Example of norm in use: Team Orbiting Moon works hard to engineer a structure that will let them accurately represent their understanding of the motion of the planetary bodies and reject construction options that would have been physically robust but would not have moved in the ways they believed the planets to move.
3) <i>Omit details that don’t affect the phenomenon you are modeling.</i> Example of norm in use: Team Fixed Moon cut clouds out of their design fairly quickly. Like the fists, clouds appeared in many students’ initial drawings, but Team Fixed Moon decided midway through Design Day to omit them despite the fact that Sammy had the materials selected. Marie judged that “we don’t need clouds” explaining that clouds weren’t part of what they needed to show.
Social Academic Norms
1) <i>Everyone should contribute ideas to the project and share responsibility for work.</i> Example of norm in use: Students routinely called each other out for not contributing enough to the work and frequently paused to review their work and assign authors to all the ideas currently represented in the representation for the purpose of assessing the division of labor.
2) <i>Each group member should understand / be able to explain the project.</i> Example of norm in use: Students held the expectation that any of them might be called upon to explain any part of the project at any time and felt that part of building a good model was that each team member be able to use the model to explain the phenomena of interest.
3) <i>Pick the easiest way to do your project.</i> Example of norm in use: In both groups students offered evaluations based on the ease of a proposed idea: “[choosing to answer] that question will be too hard” or “It’s gonna be way easier to act this out than to build it.”
4) <i>Judge ideas based on how good of a grade you think you might get for using them (or the chance of provoking other teacher reactions such as ‘getting in trouble’ for a particular choice)</i> When reasoning from school-based norms, both groups balanced priorities between choosing an easy approach and trying to get a better grade, as when Team Orbiting Moon rejected the idea of acting out their entire model because they believed “we’ll get a better grade [if we build a model] than if we act it though”

Socio-academic norms at this school draw on reform schooling discourses and have been carefully established and cultivated through the sustained effort of teachers, staff, and students. Years of being pushed to take shared intellectual responsibility for group projects has led to a stronger commitment to collective ownership of ideas and democratic processes for creating and editing in small groups here than one might expect to see in a more conventional setting (Ryu & Sandoval, 2012). In addition to norms based on the practices of progressive education, participants also invoked more generic academic norms derived from conventional

schooling practices and are related to the procedural aspects of moving through work efficiently and in a way that will be well-rewarded.

Differing Strategies for Dispute Resolution

When multiple sets of norms are at play and students have prioritized them differently, disagreements are inevitable. In the evaluative exchanges we examined, students use a variety of strategies for dispute resolution.

Appeals to authority versus focus on convincing each other: During their debate over whether or not their model should have fists in it, Team Fixed Moon appealed to the following sources of authority: the text of the assignment itself, whether or not another group was using the fists, the teacher's judgment, the opinion of a neighbor and a lunar simulation website the students consulted. During the set of debates Team Orbiting Moon had about how to make their moon orbit their earth, no outside authority was consulted.

Convincing each other was more productive in terms of student learning for a few reasons: Primarily it was because the intersubjectivity built through the kind of argumentation Team Orbiting Moon engaged in helped all students to better understand the science behind their model. Secondly: in these two cases convincing each other was actually also a faster way to resolve disputes which allowed Team Orbiting Moon to move on to the subsequent decision and get further with building and practicing their model.

Trial and error versus planning first, execution second: Team Orbiting Moon used "let's just try it" to resolve a few of the disputes considered in this analysis, Team Fixed Moon got hung up on unanimously agreeing on a plan before they started building. Being willing to operate with trial and error gave Team Orbiting Moon more time to work with their model in various forms and gave students concrete objects to reason with as they argued for or against particular representational or construction decisions.

Majority rules / trial and error versus unanimous decisions only: Both groups occasionally invoked majority rules and each group had at least one instance of a group member holding up the flow of work until a unanimous decision was reached. Unlike the other two dispute resolution strategies, it is not clear from the evaluative exchanges in these two groups that one style of dispute resolution is more or less productive than another: The dark side of majority rules decisions was that they could involve some self-suppression or oppression as students who weren't part of the majority "put up with" the majority view. On the other hand, holding out for a unanimous decision resulted in less work time for Team Fixed Moon.

Which Type of Norms Decided Arguments?

Team Orbiting Moon had a greater number of evaluative exchanges over the course of their project than Team Fixed Moon, with the balance shifting away from socio-academic disputes in the towards arguments about science and engineering as the lessons progressed. Across all three lessons, about 2/3 of the evaluative exchanges in Team Orbiting Moon were centered on scientific or engineering concerns, compared to about 1/3 of the evaluative exchanges in Team Fixed Moon. One exchange in Team Fixed Moon was double coded because it relied equally on social academic and Science-specific norms; it will be discussed in greater detail in the next section.)

Table 2: Types of evaluative exchanges for each group in each lesson

	Team Orbiting Moon (more successful model)		Team Fixed Moon (less successful model)	
Lesson:	Social/Academic	Science/Engineering	Social/Academic	Science/Engineering
Planning Day	2	0	0	0
Design Day	6	6	5	1
Building Day	1	7	2*	2*
total	9	13	7	3

* One evaluative exchange from Team Fixed Moon was double coded because it hinged equally on social academic and Science-specific norms.

Key Evaluative Exchanges from Each Group

Team Fixed Moon

The critical representational and engineering decisions which sent Team Fixed Moon into an unproductive cycle of planning and re-planning their model stemmed from a central argument over whether or not the group should build a model that accounted for the way the moon's distance from the horizon changes over time, which the group had measured by charting how many of their own fists they could fit between the moon and the horizon when they found the moon in the sky. The argument began when Zeke proposed building a model of the earth, moon, and sun as seen from outer space very similar to what he saw on one of the reference websites students accessed. Jean Carlos quickly agreed, but Megan (who was holding the planning sheet the group was supposed

to use to draft a design of their model) didn't write down their design.

"If you have the general idea of your model in mind, what is preventing you from putting anything on paper?" the TA asked. "We need to understand it," said Megan. [*Social Academic Norm 2*] "Who doesn't understand it?" asked the TA. "I don't understand it" Megan said. The TA asks someone who DID understand the model to explain it to Megan. Jean Carlos began first with an air of exaggerated patience "OK. MEGAN you know how it's a half moon right there? [pointing at the Lunar Simulator on his computer screen and continuing to speak without pausing for Megan's response] We're gonna build that. Exactly like that."

Zeke points to the materials list where Megan has listed 8 small foam balls (she planned to stack them on the earth to represent the 8 fists in Gianna's observation.) He asks Megan "do we have to have the fists? Why do we need the fists?" [Zeke's object to including the fists seems aligned with *Science Norm 3* about omitting unnecessary detail]

Megan answers that "that's what it says" [on the assignment sheet] and that they "have to represent what is in their drawing" so they do need them. (Megan's concern with the fidelity to the assignment references *Social Academic Norm 4* about evaluating ideas based on how they will be graded or judged by the teacher.) Zeke countered "[the other group] don't have the fists in their model." Megan asks Jordan from Group D if their drawing had fists in it. She said no, so Megan argued, "see? Not all of them have fists in them!"

Zeke was the first to call on the other group, using their model's lack of fists as evidence that the fists weren't needed. He used the appeal to authority (in this case, the authority of peers) to begin to resolve the dispute between his science-based objection to including the fists from Gianna's drawing and Megan's Social Academic argument that they should be included. Megan argued that Team Fixed Moon was given a different observation of the moon in the prompt for their model and therefore had a different task, reinforcing social academic norms. The group went on to repeat the argument several times, appealing to the TA, to the teacher, and to a researcher. Megan continued to voice objections to omitting the fists, emphasizing that they all had to understand and be able to explain the model and that she didn't understand how Zeke's model could show the fists (*Social Academic Norm 2*) without saying much about what she didn't understand. Eventually, with very little time left to build, the team settled on a plan that was an amalgam of Zeke's proposal and Megan's proposal. In their hasty compromise, they focused more on incorporating ideas from each student "fairly" (as social academic norm) than the science norm of correspondence to their understanding of reality.

Team Orbiting Moon

The evaluative exchange from Team Orbiting Moon presented here was one in which a key representational decision coalesced: the idea to have a freely orbiting moon attached to the earth by a rotating wire. The team had previously toyed with many potential strategies for creating a model that hung in midair and was able to demonstrate an orbiting moon. After debating whether or not to use string or wire, Marie began this evaluation by critiquing Atticus' latest suggestion that all the planetary bodies should be hung by string from a beam in the ceiling of their classroom. This key conversation came after brief exchange in which students debated using wire versus using string that we will summarize as well. Taken together, the exchanges show while a variety of types of norms were invoked in these exchanges, science norms began to take on more and more explicit priority over other norms.

The Wire Versus String Evaluative Exchange began when Atticus asked "should we use string or wire" and all group members began to argue for one material or the other. Students made clear references to engineering norms as they argued by describing wire as sturdier or saying that with string it will be easier to spin the moon (*Science Norms 3 and 4*). Then the terms of the debate shifted when Sammy pointed out that the group had already requested wire on their planning sheet. This meant that, if the group was going to respect the order of planning and building their teacher had laid out in the assignment, they would have to settle for using wire because it was the only material they had asked for and the time to request materials was over. (Her respect for the procedural directions in the assignment and concerns about taking steps out of order reflect *Social Academic Norm 4*) Iman's response to this limitation was to propose cheating. He suggested that Sammy sneak the word "string" on to the materials list the group had written the day before using the same pencil she had used yesterday so that the group could go get some string and try both string and wire in their project. Sammy declined and reminded him that the group was being filmed. (Again, the concern for adhering to the teacher's directed procedure reflects *Social Academic Norm 4*) Then Sammy said "wait though, do you even think string will be better? Guys I think wire will

be better. Wire will be better to hold it up.” (Sammy shifted to an argument that drew upon Science Norm 3 at this point, emphasizing which design would be most robust). At this point Atticus agreed, “Ok thin wire then.”

A new evaluative exchange (the Orbit Exchange) began a few moments after this decision was settled. Atticus said that the group should hang the earth and moon next to one another on two wires and use their hands to make each one spin. “No” Marie said, “It can’t all be hanging because it has to go around” she said while circling her left hand around her clinched right fist in the air in front of Atticus. (*Science-specific Norm 2*) Marie continued, “If the wire’s supposed to be hanging [as she said this she traced two invisible wires down from the ceiling with the tips of each of her index fingers, then made fists at the bottom of the wires where the moon and earth would hang in Atticus’ proposed solution] how do you make one go around the other [at this point she moved her right fist in a circle around her stationary left fist] without the wire getting twisted up?” (*Science-specific Norm 2*)

“I thought you meant it was just going to be spinning in it’s own place” said Iman as he twisted his finger in front of him without moving his hand around the table (a gesture that evoked rotation as opposed to orbit) Marie banged the table and began to yell “Moons and earths don’t/!” when Iman interrupted “yeah I know, I’m kidding I’m kidding.” (Marie’s objection, though unfinished, seems to have been an appeal to *Science-specific Norm 1* as she compares the proposal with her understanding of what moons and earths do.) After a pause Atticus said, “we could let it get twisted up I guess. And then untwist it.” Marie shrugged forward and said “soooooo very scientific” as she shot a sarcastic half-grinning look at Atticus. “Yeah!” Atticus insisted, “It will go forward around, then it will go backward around.” As he said this he spun his hands around each other clockwise and then counterclockwise to demonstrate a pair of objects tangling and untangling while hanging on two separate lines of string or wire. (*Science-specific Norm 2*) Marie said nothing, just let her half-grinning sarcastic look transform into a glare while Iman said, “I’m pretty sure the moon doesn’t go doop doh doop” as he mimicked Atticus’ hand gesture. (*Science-specific Norm 1*) “Wait, I have an idea!” said Marie. It was at this point that she sketched the initial draft of what eventually became team Orbiting Moon’s final model. “We hang the sun from here, we hang the earth from the sun, and we hang the moon from the earth. So they can all go around each other,” she explained, holding up the plan drawing. At first Atticus and Iman both objected that “this isn’t gonna work” based on their understanding of her drawing. “These two would be going around the sun the whole entire time” complained Atticus, pointing at the moon and the earth in Marie’s drawing. (*Science-specific Norm 1*) “no no no” Marie explained, “We can make it so that the moon is attached to the earth not the sun. so the moon can go around the earth.” After leaning in over her drawing Atticus said “OH YEAH!” and team began to plan their model.

In the Wire Versus String Evaluative Exchange, Iman’s suggestion of cheating suggests an explicit prioritizing of engineering and science norms over social academic norms because it demonstrates that he is willing to violate social academic norms in order to get the materials he believes will result in a sturdier build. While she declines to write the new addition to their materials list, Sammy doesn’t object to Iman based on Social Academic norms, she offers her own engineering norm-based argument in favor of wire, which ultimately settles the exchange. Therefore the debate, while touching on multiple sets of norms, is settled in terms of science-specific norms. In the Orbit Evaluative Exchange, both Atticus and Marie objected to each others engineering plans by comparing the motions they would create to their understanding of the actual motion of the planetary bodies. What is more, Marie explicitly named the body of norms against which she was judging Atticus’ suggestion when she sarcastically mocked his idea as “so scientific.” Atticus later critiqued her proposal in a very similar fashion, arguing against her idea because he believed she was proposing to have the moon orbit the sun, which he called out as inaccurate.

Discussion

This analysis demonstrated that children in this science class had to negotiate multiple sets of norms in the dual-layered social space of their science classroom. The norms were not always harmonious, and there was not general agreement about whether social academic norms or science norms should take priority when they pointed students in differing directions. Instead, children used a variety of strategies for resolving conflicts that implicitly indexed different sets of norms. These resolution strategies were not neutral with respect to norms. In fact, depending on the strategy they employed, children were aligned towards either social academic norms or towards science norms. For instance, the strategy of temporarily agreeing to a solution and subjecting it to trial and error aligned Team Orbiting Moon with several science norms and led to deep conversations about the

science behind the phases of the moon. Agreeing to trial-and-error allowed the group to achieve some degree of consensus around their design while maintaining flexibility. The dispute resolution strategy of trying to convince each other aligns children with the science norm that ideas should be judged based on their correspondence to reality, whereas making appeals to authority aligns students with the social academic value that ideas should be judged based on how well they will be received by the teacher or other academic stakeholders. The dispute resolution strategy of unanimous decision-making aligns students with the social academic norm that all group members should understand the group's project.

Which norms apply to which decisions was socially negotiated by children in the moment. These momentary negotiations built upon each other: one appeal to authority by Team Fixed Moon was countered by another appeal to a different authority and so on until the team had consulted nearly everyone in the room about their standoff over "showing the fists." This is an extension of what Anderson et al (1001) called snowballing. Anderson found that after the first appearance of a particular rhetorical stratagem, the probability that it would appear again rose. This analysis suggests that it is not only rhetorical moves that can snowball in group activities, but that orientations towards one set of norms or another and strategies for dispute resolution may snowball as well. Team Orbiting Moon provides an example of a productive snowball effect: their evaluative exchanges were more and more likely to hinge on scientific norms as time went by, particularly the norm of matching reality. This led to a deeper engagement with the science behind the phenomena they were trying to represent. For team Fixed Moon, it was not a particular norm that snowballed, but the strategy of appealing to authority.

In most of their evaluative exchanges, children negotiated which norms should apply to particular decisions about their projects implicitly. However, Team Orbiting Moon benefited from particular conversational moves made that made science norms the explicit criteria of the groups' self-evaluation, such as when Marie teased Atticus that his idea to let the moon rotate the earth in one direction and then in another was unscientific. This comment and other explicit references to science served to reorient the group to science-related norms and set the stage for their productive engagement with the model-building challenge. After Marie's teasing, the group noticeably shifts towards arguing about the science of the moon-sun-earth system instead of their prior focus on the parameters of the assignment and the engineering benefits of using wire versus string.

To become more skilled at scientific representation, students need to practice evaluating their models with a critical eye, not just building them. This paper demonstrates that while building models, students used a variety of kinds of norms as reference for their judgments and that these norms were rarely explicitly invoked. When norms were explicitly invoked (such as when Marie teased Atticus for his 'unscientific' idea) students tended to coalesce around the explicitly named norm as a reference point for their subsequent judgments. A framework for explicitly describing the nature of their own evaluations and objections could help students to have better arguments and build better models by pushing them to articulate their arguments in terms of the norms upon which they are drawing. Future work should explore: 1) ways of making norms and the negotiations around norms visible for students and 2) ways of encouraging students to draw upon science-specific norms when they evaluate work in science class rather than relying only/primarily upon social academic norms.

References

- Anderson, R. C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S., Reznitskaya, A., et al. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition & Instruction*, 19, 1–46.
- Brown, B. (2004) Discursive Identity: Assimilation into the culture of science and its implications for minority students. *Journal of Research in Science Teaching*, Vol. 41, No.8, pp.810-834.
- Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *The Journal of the Learning Sciences*, 10, 113–163
- Cook, M. (2011). The identity-related consequences of a cultural model of scientific scholarship and individuals' appropriations of it. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Danish, J. A., & Enyedy, N. (2007). Negotiated Representational Mediators: How Young Children Decide What to Include in Their Science Representations. *Science Education*, 91(1), 1-35.
- Enyedy, N., Danish, J. A., Fields, D. (2011). Negotiating the Relevant in Culturally Relevant Mathematics. *Canadian Journal for Science, Mathematics, and Technology Education* 11(3).
- Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge University Press.
- Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children's epistemic understanding from sustained argumentation. *Science Education*, 96 (3), 488-526.

The Beginnings of Engineering Design in an Integrated Engineering and Literacy Task

Mary McCormick and David Hammer, Tufts University, Medford, MA
Email: mary.mccormick@tufts.edu, david.hammer@tufts.edu

Abstract: In *Integrating Engineering and Literacy* (IEL), students identify and engineer solutions to problems that arise for fictional characters in stories they read for class. There are advantages to this integration, for both engineering and literacy goals of instruction: The stories provide “clients,” to support students’ engagement in engineering, and understanding clients’ needs involves close attention and interpretation of text. Results are encouraging, but mixed, in part for variation in how students frame the task. For instance, while students often pay close attention to the stories, interpreting and anticipating their fictional clients’ needs, they sometimes focus more on the teacher and what they think she would like to see. This variation occurs both within and across groups of students, and it motivates studying the dynamics of student framing. Here we examine two students’ stability in framing their work as nascent engineers.

Introduction

The following excerpt takes place in a fourth grade classroom in a rural Massachusetts town. The students are involved in an Integrated Engineering and Literacy (IEL) design task based on the book *Shiloh*, by Phyllis Reynolds Naylor. They are to design and construct a small scale model of a dog pen for Marty, a young boy trying to protect Shiloh, a small beagle, from an abusive owner. Ms. C, the students’ teacher, has explained they are to use craft materials (e.g., cardboard, tape, glue, felt) to create a model that will “fit on their desk.” She noted that their designs will be “tested,” but she does not say how. Stella and Alexi are working on their sketch (Figure 1) when Ms. C asks them about their design.

- Ms. C: What kind of entrance is it?
Stella: It’s...
Alexi: Just a little door, like the walls are two feet.
Ms. C: Two feet thick? Or two feet high?
Alexi: No, two feet high.
Ms. C: Two feet high, okay.
Alexi: I wrote it on this side somewhere (flipping over the paper), here it is. Oh yeah, oh, it’s three feet.
Ms. C: Three feet! Wow. Why three feet, what made you decide three feet?
Alexi: Um... so it wouldn’t be too short, like when Marty wants to go in, since it has, like, that glass that doesn’t break on the top, he doesn’t have to scrunch down (positioning her body to make scrunching gesture).
Ms. C: Ah, so Marty could go in as well?
Alexi: Yeah.
Ms. C: Alright, very cool. Did the graph paper help you guys draw your diagram?
Alexi: Yeah.
Stella: Like, she did six squares that way, and then eight squares that way (pointing to lines sketch).
Ms. C: Oh, so you’re using *measurement* as well, excellent work!
Alexi: (to Stella) And now, you can draw the bolts (pointing to outer edge of hexagon)...
Stella: Okay. Right, so the door is, like, right here. Okay?

Stella and Alexi continue discussing the advantages and disadvantages of entrance locations, adjusting their sketch.

Our first purpose in this paper is to argue that Stella’s and Alexi’s considerations and reasoning evidence nascent abilities for engineering design. For example, Stella and Alexi consider their clients and what might help them, and they use graph paper to generate an appropriate scale. Explaining their estimate of “three feet,” Alexi describes the possibility of Marty needing to enter the tunnel and a height that would allow him to fit comfortably. Like engineers, they negotiate criteria (e.g. access, height, location) and constraints (e.g., abutting glass elements), and they “prioritize the needs of their clients” (Ropohl, 1997, p. 70).

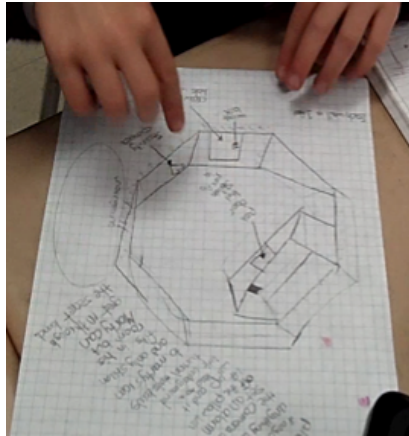


Figure 1. Design sketch

Our second purpose is to consider how Stella's and Alexi's pursuit of a design solution reflects their understanding of what it is they are doing. It is one thing to have abilities, for example, to scale a diagram; it is another thing to recognize a need to make use of those abilities. For the girls, scaling their diagram to "real life" was not part of their assignment; they recognized a need to scale in order to support their thinking about their design. That is, part of the dynamics of their activity in this moment involves their understandings of "what is it that's going on" (Goffman, 1974, p. 8), for their client in the story as well as for them in the classroom, in other words their *framing* (Goffman, 1974; Tannen, 1993).

In the following sections, we discuss the construct of framing, in particular of epistemic activity (Redish, 2004; Hammer, Elby, Scherr & Redish, 2005), in a classroom setting and as applied to research on engineering (Dym, Agogino, Eris, Frey, & Leifer, 2005; Schön, 1983; Vincenti, 1990). We then turn to Stella and Alexi to interpret how they frame what they are doing, in particular to their stability in framing: The girls show resilience in their understanding the activity as involving designing and presenting their ideas for the fictional clients. In this we propose that part of their stability was their involvement in the story itself, including their empathy for the characters. We close the paper with a discussion of further questions for research and possible implications for instruction.

Framing

In a given situation, whether it involves playing soccer, learning science, or designing a bridge, people form a sense of what is taking place, what researchers have called a "frame" (Goffman, 1974; Tannen, 1993). Forming that sense, or "framing," reflects structures of expectations formed from previous experiences (Tannen, 1993). In these accounts, frames are knowledge structures that both shape and are shaped by experience, and framing is a dynamic interaction between expectations and perceptions. Frames are not static, rigid structures, but are active and responsive, perpetually evolving as they are informed, shaped, and tuned with new experiences; in this sense, they are "schemas" (Bartlett, 1932) of activity. "One's structures of expectation make interpretation possible, but in the process, they also reflect back on the perception of the world to justify that interpretation" (Tannen, 1993, p. 20-21).

For Stella and Alexi, part of the challenge was to form a sense of their task, engineering for Marty and Shiloh, and that would involve their tapping into patterns of their previous experiences of telling stories, doing projects in school, making things, and so on. Part of the challenge, too, was in understanding the situation in the story. Their experiences similarly shaped their comprehension of the novel, in structures of expectations about caring for dogs, ownership and protection, and so on. At the same time, their experiences in this task contribute to those patterns, perhaps helping them understand future experiences. Reading the story, for example, may be their first encounter with the idea of an abusive owner; designing the protective pen may be one of their first experiences of engineering.

Epistemological Framing in a Classroom Setting

There are many aspects to framing, at multiple scales and with complex, nested relationships. Someone baking has an overall sense of what baking involves, but may cue finer-grained framings within subtasks of measuring, mixing, frosting, etc. In Stella's and Alexi's case, their framing of being students in a classroom may be constituted by expectations for sitting at their desks, listening to their teacher or an adult in charge, and enacting certain actions for specific time blocks. Within that, they may activate frames for "learning science" that involve experimenting and making sense of phenomena, and other frames for "learning spelling" that involve memorizing sequences of letters. Thus, across and within different activities or classroom contexts, students activate and

tune their expectations, including with respect to knowledge and learning, that is their “epistemological framing,” (Redish, 2004).

Research in science education has paid significant attention to students’ expectations with respect to knowledge. A variety of studies have documented students experiencing science class as focusing on the authority of the teacher or textbook (Lemke, 1990; Jimenez-Aleixandre, Rodriguez & Duschl, 2000; Hammer, 1994; Redish, Steinberg & Saul, 1998), rather than on making tangible sense of natural phenomena. In these cases, students frame what they are doing as memorizing, storing, and reproducing known information, rather than, for example, producing and assessing knowledge. Recent accounts have built on this work by attending to the local dynamics of students’ framing (Hammer, 2004; Louca, Elby, Hammer & Kagey, 2004; Rosenberg, Hammer & Phelan, 2006), evidencing the sensitivity to features of context and social interactions. Researchers’ findings indicate that for students to be actively learning science, they must not only frame what they are doing as sense-making about natural phenomena, they must do so with stability, e.g. for resilience against the familiar “classroom game” (Lemke, 1990) that focuses less on the natural world than on the authority of the teacher or text (Hutchinson & Hammer, 2010).

In this work we study how students frame their work in engineering. We are interested to understand aspects of framing that are productive for engineering as well as in the local dynamics, stabilities, and variations. As in science, students may frame what they are doing in ways that are counterproductive for engineering, including following a sequence of steps (e.g., Massachusetts Curriculum Frameworks, 2010), or assuming there is a single “right answer” (Johnsey, 1995, 1997; Hennessy & McCormick, 1994; Welch, 1995). Accordingly, research in engineering education often focuses on students’ abilities follow these steps, such as planning in the beginning of a design endeavor. In such cases, when students do not follow the prescribed sequence (e.g., planning while constructing), they may be diagnosed as lacking in engineering ability. A framing perspective, however, offers an alternative possibility: Students’ understanding of what is taking place have them invoke abilities they have, e.g. for planning (Portsmore, 2010). This motivates attention in engineering education beyond abilities, both in interpreting students’ work and in planning objectives for lessons, in particular to cultivate productive framings for engineering.

Productive Framing in Engineering

A view of framing sees engineers’ understanding of design as involving patterns of familiar experiences, tuned to the particulars of situation. This is the heart of schema theory; a schema is “an active organization...of past experiences” (Bartlett, p. 201), active to include local tuning. As Schön (1983) describes, engineers “are not confronted with problems that are independent to each other, but with dynamic situations that consist of complex systems of changing problems,” (p. 16). In “making sense of a situation” (Schön, 1983, p. 40), an engineer maintains a heightened awareness of the overarching design task, while attending to the multiplicity of interacting subtasks (Dym et al., 2005). Accordingly, design tasks generally involve subtasks, and this is part of engineers’ framings. For example, an engineer’s framing of a bridge design project may involve optimally meeting the client’s needs while adhering to situational constraints. Within this overarching framing, the engineer is simultaneously recognizing subtasks, such as researching the environment, developing and analyzing computer models, and negotiating with contractors and community members. At each decision juncture, the engineer must reflect on the big picture, recognizing clients’ needs and design constraints, and respond with appropriate modes of reasoning and action, such as analyzing, evaluating, constructing, etc (Trevelyan, 2010).

Analogously, students’ framing of an IEL task may involve reflecting on the story and responding to characters’ needs. Our early findings suggest that a story setting provides a sufficiently “messy” (Schön, 1983, p. 33) design context, in which story characters become clients with wants, needs, and potential dilemmas, and there are implicit physical, social, and economic constraints (McCormick & Hynes, 2012). Thus, in framing a complex design task as beginning engineers, students may recognize a need to reason, make decisions, and act as engineers: to develop an optimal solution for their client. We argue that engineering abilities, or “technical know how” (Ropohl, 1997), should not be our sole end goal in engineering education. Fostering productive framing should be a central target for research and practice, such that students recognize a need to use their engineering abilities.

Integrating Engineering and Literacy

This study is part of an NSF funded project at University *Integrating Engineering and Literacy*. The primary goal of IEL is to support elementary school teachers’ incorporating engineering into work with children’s literature. Participating teachers develop and implement IEL units using stories that are already part of their curricula. In preparation, teachers attend approximately forty hours of professional development per school year at Tufts University to work with researchers in developing lessons and implementation strategies.

Research Aim

Over the last three years, our research team has collaborated with teachers to explore IEL activities in elementary classrooms, with a variety of book genres, materials, and lesson structure. Much of our interest is to understand what comes or may come of these choices, in particular with respect to the students' learning of engineering design and development of literacy skills. We have noticed a wide spectrum of ways in which students engage in the assignment. For example, while some stay anchored in the story, others focus on what they think their teacher wants to see; while some stay focus on tangible mechanisms, other incorporate fantastical elements in their design. We see this variation as reflecting dynamics of students' framing, including where students direct their attention, what ideas they consider, and how they evaluate those ideas.

In this paper, we focus on a pair of students who, the evidence suggests, are stable in their framing. While many of their classmates respond to prompts in the classroom, apparently shifting from focusing on the situation for their fictional clients to focusing on their sense of the teacher's requirements, Stella and Alexi persist in designing a solution that fits within the story context. With this analysis, we aim to understand how they are framing the task, why they do so with stability, and how their framing supports their engagement in engineering design.

Research Setting

This case study comes from a in a fourth grade classroom in a rural town in Massachusetts, about forty miles from Boston. The teacher, Ms. C, had attended approximately thirty-five hours of professional development as part of the IEL project and was excited to try an IEL activity using the book *Shiloh*. She was devoted two hours per day for three days: Day One involved class read-aloud, discussing the major problems in the book, and starting individual plans; Day Two working with a partner on design plans and building, and Day Three finishing designs followed by group tests and presentations.

Data Collection and Analysis

Members of the researchers team are often in classrooms during IEL activities, providing materials, supporting teachers and students during building, as well as taking field notes and video recording. In this case, two researchers were present, including the first author.

The following draws from multiple viewings of the video in the research group, including consideration from multiple theoretical lenses to understand student engagement (Jordan & Henderson, 1995). Here we focus on evidence of framing (Tannen, 1993), in students' discourse, gestures, as well as writings and sketches.

Phases of the Task

In the following, we show three excerpts of Stella's and Alexi's work in chronological order. We highlight these moments to show evidence of how they frame the task, and the stability with which they do so. In the first excerpt, Stella and Alexi explain their design decisions, focusing on their fictional clients' needs. In the second and third, they show resilience in this framing against competing expectations regarding testing and evaluating criteria.

Design Considerations (Day One)

During the initial phase of their design, all of the students in class are working in pairs or groups of three to co-construct a sketch of a dog pen for Shiloh. When the materials for building (e.g., cardboard, paper, glue, etc.) become available, many students rush to grab them. Others, including Stella and Alexi, continue to work on the details of their design sketches. In the following, the first author asks the girls about their work.

- Mary: That's a cool design. What is, so what do you have?
- Alexi: It's like, in this [unclear], and there's a little lock, so Marty can just turn the lock, and there's a little door that Shiloh just fits in. And if the camera sees something that it doesn't recognize, like, if it's not Marty's family or something, or if it's something else, it'll, like, this door will go automatically open, and the pillow will come out, and there's underground tunnels, and there's, like, a little, um, there's kinda, like, a little box in here — I kinda drew dotted lines.
- Mary: That's really cool!
- Alexi: -and then there's tunnels leading to Marty's room, and an alarm will go off in Marty's room, so he can just crawl through the tunnels and get to Shiloh.
- Mary: Oh, that's really cool! So you're thinking about how Marty can — is Marty the owner of the dog?
- Alexi: Yeah.
- Stella: Well, not necessarily the owner...
- Alexi: But he wants to be the owner!
- Mary: (laughing) He wants to be!

Alexi: So that's why he's trying to keep it very secret.

The girls focus on keeping Shiloh safe, comfortable, and accessible to Marty. They describe the functional issues of the tunnel connection, with attention to details from the story: the tunnel is accessed only through the pillow door, and provides a direct route for Shiloh to Marty's room. As they imagine Shiloh's escape route, they consider multiple perspectives: Alexi describes the path Shiloh will take to get to Marty's room, as well as a way for Marty to be alerted so he can quickly rescue Shiloh in the case of danger. They develop contingency plans to account for implicit "what if" circumstances, such as sizing the tunnel door so that "just Shiloh fits in," in case the bigger dog gets past the first barrier, while maintaining "must haves" (i.e., an door to the pen) (Schön, 1983, p. 101). Although keeping Shiloh hidden or "secret" was never discussed as a classroom requirement, Stella and Alexi make it a top priority, realizing that if he is caught, he will likely be abused again.

Thus Stella and Alexi coordinate their overarching design goals of keeping Shiloh safe and secret with subtasks of developing and evaluating components. Their decisions are not driven primarily by the classroom requirements of size and testability but by the girls' sense of the physical and social setting of the story. In these ways, the evidence suggests they frame the engineering design task as situated in the story.

How Do You Test It? (Day Two)

At the start of Day Two, with the class assembled as a whole, Ms. C. calls on Stella to summarize the requirements.

Ms. C.: Stella, can you give a quick summary of what our requirements would be?

Stella: Oh, okay. Um, must fit on top of our desk and the test must be able to fit inside (referring to "inside" the dog pen).

Ms. C.: Whatever we choose, however we choose to test, it (referring to testing object) must be able to fit inside (the dog pens) so we can see if Shiloh would be able to get out and if something would be able to get in. And there was one more on the bottom, it has to be some sort of...

Stella: Pen.

Ms. C: Pen, right? Some sort of enclosure.

The students' design task, as Stella remembers, is to construct a model of Shiloh's dog pen that is scaled to "fit on top of our desk," and the scaled model must be functional. Ms. C. confirms and elaborates several criteria for the test, referring to a wind-up toy they will use in class to represent Shiloh: (1) the object must fit inside the model; (2) the model must have boundaries that will prevent the object from leaving ("we can see if Shiloh would be able to get out"), and (3) the model must protect keep outside objects from getting in.

As the students in the class construct their pens, they all evaluate their projects, but by different criteria. For instance, while some evaluate based on how well it will work for Marty and Shiloh, others prioritize "classroom" expectations, anticipating how their projects will be assessed in comparison to their classmates'. In the following, Stella and Alexi are working on their project when another student, Owen, who has finished his dog pen, comes to look at their work.

Owen: Did you guys see ours?

Alexi: Yeah, yours is awesome. Did yours make it through the tests?

Owen: Not yet.

Mary: How are you guys testing it?

Stella: Um, over there, I don't know what she's doing (pointing towards Ms. C).

Mary: How do you think you'd want to test it?

Alexi: I think she's gonna take, like, a little wind-up toy, and it's just gonna walk around and it can't, your thing can't fall over.

Stella: Well this is felt, so I don't even know if it would be able to walk. But the felt is good, cause then it's soft.

The interaction between Stella and Owen evidences competing expectations for the design task. The "classroom" expectations involve passing the test with the wind-up toy; the client-focused expectations involve optimizing a design for Marty and Shiloh that makes sense in the story context. When Alexi asks Owen if his dog pen "made it through the tests," she shows an awareness that their projects will be tested when they are done, that Ms. C is "doing" the test, and that her design may be compared to the other students' designs based on their relative success on the test. When pressed on what the test involves, Stella reacts dismissively: She gestures to the other side of the classroom, but quickly resumes her focus on constructing, biting her lip as she figures out how to attach the roof. She is clearly uncertain about the parameters of Ms. C.'s test, but does not seem phased by this. Alexi then elaborates that the test involves an action that "she" (her teacher) will perform using a "wind-

up toy” to make sure the “thing can’t fall over.” And, when Stella sees a feature of their design that might perform badly in the test—the felt might keep the toy from being able to move—she keeps it anyway, thinking of her clients.

In a classroom framing, the test is a familiar event that makes sense; it adheres to classroom expectations. However, for the girls, “test” cues up a pro-forma event that is disconnected from the story context and their goals. In this event, their teacher performs an action, Shiloh is a “wind up toy,” and their dog pen is a “thing.” Although they recognize that other students may be prioritizing the test, Stella and Alexi remain rooted in the story context, as evidenced in Stella’s comment in the last line. Her explicit prioritization of Shiloh’s comfort over classroom testing criteria suggests that she is aware of the competing sets of expectations but committed to her own.

Evaluating for the Client (Day Three)

On Day Three, all of the students take turns presenting and testing their designs. Ms. C announces that the dog pen test is two parts: (1) a “small dog” test, which involves letting a small wind-up toy scurry about inside the pen for thirty seconds without escaping, and (2) a “big dog” test, which involved winding up two bigger toy cars (to represent big dogs) and letting them crash into the sides of the pen. During Stella and Alexi’s presentation, they highlight meaningful features of their design, elaborating on how the tunnel will function as an escape route in case the antagonists of the story come after Shiloh. When they are ready to test, Ms. C suggests that the first test should be for the small dog to slide down the secret tunnel part of the design. The students are gathered around Stella and Alexi’s design to observe, hoping to see the small dog emerge from the bottom of the tunnel.

- Students: He’s at the bottom! (“He” refers to Shiloh and/or the toy).
- Ms. C: Oh, he came out! All right, so the small dog was able to go through the tube (referring to the tunnel). Why might it be tricky to test going *up* the tube?
- Alexi: (without pausing) He (referring to Shiloh) doesn’t go up the tube because Marty lives on the bottom of the hill and Shiloh’s pen is on a hill. So he would just like, Marty would walk him up the hill.
- Ms. C: Okay, so he’s not expected to go back up the tube. He’s expected to start at the top and go all the way down.

During the test, Ms. C raises the question of whether using the toy would be appropriate to find out whether the dog could go up the tunnel. For Alexi, though, that question is moot. She responds by describing how her design works in the story setting, insinuating that there is no reason to test the small dog going up the tube because that is not how the tunnel is designed to work. Rather than adapting their framing to incorporate the classroom expectations, Stella and Alexi persistently remain focused on designing for their clients.

Stable Beginnings in Engineering

Our initial motivation to study this case was to examine Stella’s and Alexi’s abilities to reason and act as engineers. In early analyses, we examined how they spontaneously planned by considering multiple aspects of the design context and their clients’ needs, and generated appropriate scales to ensure accuracy in a “real life” context. In accounting for social and physical dimensions, they seemed to tacitly recognize that “design does not take place for its own sake or in isolation, but rather is directed at a practical set of goals intended to serve human beings” (Vincenti, 1990, p. 6). Much like engineers, the girls demonstrated “design thinking,” making informed assumptions, reasoning to narrow uncertainty, and considering outcomes of hypothetical situations (Dym et al., 2005).

As we continued to study Stella and Alexi, we became more interested in their framing of what they were doing, itself an aspect of their nascent engineering. Like engineers, Stella and Alexi were continually reflecting on and responding to their clients’ needs within the context of the story, in contrast to some other groups that evidently framed what they were doing more directly in terms of their own needs within the context of the classroom.

We were struck as well by Stella’s and Alexi’s stability in their framing, in contrast, for example, to the ways another group shifted their framing in response to classroom cues. That group, of three boys, discussed their initial design decisions based on “keeping Shiloh safe,” and ensuring access to sunlight so Shiloh “doesn’t feel trapped,” evidence of framing comparable to Stella’s and Alexi’s. Later, when presenting their design to their teacher, the boys made a point of using terms from geometry, including “rhombus,” “square,” and “hexagon,” which they rightfully expected Ms. C would appreciate. In another instance, a pair of girls incorporated LEGO figurines as “body guards” and pipe cleaners as “laser beams” to protect Shiloh. Because their initial design sketch did not include these imagined features, we suspect the girls’ interest in craft materials triggered a shift in their framing away from the situation of the story. That is, they adjusted their framing of what they were doing, essentially shifting the genre of the story as written, to include elements of fantasy or science fiction (1).

To summarize, many of the students' framing of the IEL task dynamically evolved as they responded to classroom prompts, interactions with other students, or materials in the classroom. Stella and Alexi, however, remained stable in their focus on designing for their fictional clients, within the context of the story, even in potentially pivotal moments.

Their stability has piqued our interest and spawned a larger research question: What was it about their framing that enabled them to be stable? Based on this analysis, our conjecture is that Stella's and Alexi's stability in this task came in part from their investment in the story, including their caring for the characters and their problem. We see evidence of the story holding their attention in their responses to questions about their activity, with references to details about the situation, such as Marty wanting to be Shiloh's owner, as well as signs of their imagining aspects of the situation not explicitly in the story, such as how their system might need to let Marty's family in, that Marty might need to get in the tunnel himself, or that the real need for the tunnel would be to escape from the pen to Marty's room if endangered. In this they demonstrate *design empathy* (Kouprie & Visser, 2009), an understanding of and concern for their clients, ensuring that Marty and Shiloh will have access to each other and that Shiloh's pen will provide safety, comfort, and security. By imaginatively projecting themselves into Marty's and Shiloh's situations (Koskinen & Battarbee, 2003), Stella and Alexi are able to deeply discern their clients' circumstances and perspectives (Battarbee, 2004), and to design a solution to best meet their needs.

While the importance of empathy in design is well recognized (Battarbee & Koskinen, 2005), many researchers have noted that it is often lacking in the design process (Fulton Suri, 2003; Mattelmaki & Battarbee, 2002), and have developed a number of tools and techniques to enhance designers' empathy (Kouprie et al., 2009). In this case of children engaging engineering design, however, we see the opposite: Stella and Alexi's ability to empathize not only informs their design decisions, it supports and sustains their framing of the task as engineers.

Conclusion

In this study, we showed elementary students' abilities to reason and act as engineers for fictional clients in children's literature. In developing an optimal solution for Marty and Shiloh, Stella and Alexi considered design criteria and constraints that were implicit to the story; their need to understand the characters was synergistic with their need to comprehend the text. As they interacted with a complex design situation, the girls were not following a procedural set of engineering design steps, but instead were invoking their engineering abilities purposefully and with agency; within their framing of task, their actions made sense. In contrast to many of their classmates, Stella and Alexi remained stable in their framing, even when their design criteria were in tension with the classroom expectations for testing.

This case study is part of a larger project to understand students' framing in engineering. From this and other observations and analyses, we suggest that student framing should be a central target in engineering education research and practice. By attending to student framing in research, we may illuminate not only students' engineering abilities, but also their reasons for enacting those abilities. Moreover, we may be better equipped to foster and cultivate productive framing during engineering activities in practice, providing students with opportunities to design for clients and to interact with multidimensional problem situations. Our hope is that as students gain experience in framing as engineers, they may strengthen their abilities to navigate complex design situations, such that their engineering "ways of knowing" become "tacit, spontaneous, and automatic" (Schön, 1983; p. 60).

Endnotes

- (1) Of course, Stella's and Alexi's design was also unrealistic—it would be quite difficult to dig that tunnel! Our claim is that imagining a tunnel is much closer to the story context than imagining bodyguards and lasers.

References

- Bartlett, F.C (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.
- Bateson, G. (1972). A theory of play and fantasy (Originally published 1954). In G. Bateson (Ed.), *Steps to an ecology of mind; collected essays in anthropology, psychiatry, evolution, and epistemology* (pp. 177–193). San Francisco, CA: Chandler Publishing Company.
- Battarbee, K. and Koskinen, I., 2005. Co-experience: user experience as interaction. *CoDesign*, 1 (1), 5–18.
- Berland, L. K., & Hammer, D. (2012). Framing for scientific argumentation. *Journal of Research in Science Teaching*, 48 (1), 68–94.
- Dym, Clive. (1994). *Engineering Design: A Synthesis of Views*. Cambridge, UK: Cambridge University Press.
- Dym, C, Agogino, A., Eris, O., Frey, D., & Leifer, L. (2005). Engineering Design Thinking, Teaching, and Learning. *Journal of Engineering Education*. pp. 103– 120.
- Fulton Suri (2003). Empathic design: informed and inspired by other people's experience. In: I.Koskinen, K.

- Battarbee, and T. Mattelmäki, eds. *Empathic design, user experience in product design*. Helsinki: IT Press.
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Cambridge, MA: Harvard University Press.
- Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction*, 12 (2), 151-183.
- 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.
- Hennessy, S., & McCormick, R. (1994). The general problem-solving process in teaching education: Myth or reality? In F. Banks (Ed.), *Teaching technology*. London: Routledge/ Open University Press.
- Hutchison, P., & Hammer, D. (2010). Attending to Student Epistemological Framing in a Science Classroom. *Science Education*, 94(3), 506-524.
- Jimenez-Aleixandre, M., Rodriguez, A., & Duschl, R., (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education*, 84, 757 – 792.
- Johnsey, R. (1995). The place of the process skill making in design and technology: Lessons from research into the way primary children design and make. Paper presented at the IDATER95: International Conference on Design and Technology Educational Research and Curriculum Development, Loughborough, UK: Loughborough University of Technology.
- Koskinen, I. & Battarbee, K. (2003). Introduction to user experience and empathic design. In: I. Koskinen, K. Battarbee, and T. Mattelmäki, eds. *Empathic design, user experience in product design*. Helsinki: IT Press/
- Kouprie, M. & Visser, F. (2009). A framework for empathy in design: stepping into and out of the user's life. *Journal of Engineering Design*, 20 (5), 437-448.
- Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex Publishing.
- Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist*, 39 (1), 57-68.
- Massachusetts Department of Elementary and Secondary Education. (2010). *Technology/Engineering Concept and Skill Progressions*. Massachusetts: Massachusetts DEoE. Document Number)
- Mattelmäki, T. and Battarbee, K., 2002. Empathy probes. In: T. Binder, J. Gregory, and I. Wagner, eds. *Proceedings of the participatory design conference 2002*. Palo Alto CA: CPSR, 266–271.
- Minsky, M. (1975). A framework for representing knowledge. In P. Winston (Ed.), *The psychology of computer vision* (pp. 211–277). New York: McGraw-Hill. Moje, E. B., Ciechanowski, K. M., Kramer, K., Ellis.
- Portsmore, M. (2010). *Exploring How Experience with Planning Impacts First Grade Students' Planning and Solutions to Engineering Design Problems*. Education. Medford, MA, Tufts University, PhD.
- Redish, E. F. (2004). A theoretical framework for physics education research: Modeling student thinking. In E. F. Redish & M. Vicentini, (Eds.), *Proceedings of the Enrico Fermi Summer School Course, CLVI* (pp. 1–63). Bologna, Italy: Italian Physical Society.
- Ropohl, G. (1997). Knowledge types in technology. *International Journal of Technology and Design Education*, 7.
- Rosenberg, S.A., Hammer, D., & Phelan (2006) Multiple epistemological coherences in an eighth-grade discussion of the rock cycle. *Journal of the Learning Sciences* 15(2), 261-292.
- Scherr, R., & Hammer, D. (2009). Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics. *Cognition and Instruction*, 27(2), 147–174.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books. (Reprinted in 1995).
- Vicenti, W. (1990). *What Engineers Know and How They Know It*. Baltimore and London: The Johns Hopkins University Press.
- Welch, M. (1995). Analyzing the Tacit Strategies of Novice Designers. *Research in Science & Technology Education*, 17(1).

The Roles of Teacher Questioning in Argument-Based Inquiry (ABI): Approaches that Promote Cognitive Thinking and Dialogical Interaction

Ying-Chih Chen, Arizona State University, Mary Lou Fulton Teachers College, chenpeter78@gmail.com
Brian Hand, University of Iowa, Teaching and Learning, brian-hand@uiowa.edu

Abstract: The purpose of this study was to investigate the various questioning roles elementary teachers adopt to scaffold dialogical interaction, students' cognitive responses, and the use of evidence for constructing and critiquing ideas in argumentation over time. This study was designed as a follow-up study after a four-year professional development program that emphasized an argument-based inquiry approach. Data sources included 30 science lessons focusing on whole-class discussion from three early elementary teachers' classes. Data were analyzed through constant comparative method and enumerative approach. The findings indicated: (1) teachers used multiple roles in establishing argumentative discourse as they persistently implemented an argument-based inquiry approach, (2) as teachers used multiple roles in establishing patterns of questioning, framing classroom interaction, students' higher levels of cognitive thinking was promoted, and (3) as teachers' patterns of questioning changed, the frequency of students' talk increased and the dialogical interaction between students and teachers became more evidence-based and connected.

Introduction

The emphasis on argumentation in science education shifts the focus of science classrooms from memorizing facts to engaging students in an authentic scientific practice in which they search for data patterns to shape evidence for the support of scientific claims and debate those claims publicly to identify the weaknesses of their arguments (Cavagnetto, 2010; Berland & Reiser, 2009). In this form, learning science is not simply about how to define vocabulary to explain content, but is rather about the ways in which questioning and evolving questioning techniques can be used to extend one's conceptual understanding of the subject matter. The importance of argumentation has been explicitly endorsed by two recent U.S. reform documents, *Taking Science to School* (National Research Council (NRC), 2007) and *A Framework for K-12 Science Education* (NRC, 2012), as a critical approach utilized in science classrooms for promoting student conceptual understanding and cognitive thinking.

However, *A Framework for K-12 Science Education* (NRC, 2012) raised the critical point that argumentation "has too often been underemphasized in the context of science education" (p. 44, NRC, 2012). The problem has not been so much the content of the lessons (BaniLower et al., 2013) or the pedagogical knowledge of the teachers (Ball, 2000), but rather the way teachers ask questions to engage students in the practice of actively constructing and critiquing scientific knowledge (Hmelo-Silver & Barrows, 2008). A national observation survey in the United States conducted by BaniLower et al. (2013) found that more than 90% of questioning patterns in science lessons consist of low-level "fill-in-the-blank" questions, asked in rapid-fire fashion to obtain the correct answer and move on, in effect short-circuiting student cognitive thinking. Teachers struggle with developing appropriate and diverse roles for asking a series of high quality questions to establish argumentative discourse and foster students' conceptual understanding and cognitive thinking (Oliveira, 2010). For Chin (2007), teachers' questions "play an important role in determining the nature of discourse" (p. 815) for argumentation, including eliciting, comparing, challenging, and synthesizing students' arguments in different conditions.

Developing appropriate and multiple roles of questioning in the establishment of argumentative discourse is challenging and takes time (Martin & Hand, 2009; McNeill, & Pimentel, 2010). Addressing this challenge requires examining teacher roles of questioning in the establishment of argumentative discourse as well as examining the dynamic discourse patterns among teachers and students that can influence students' cognitive thinking (Oliveira, 2010). Many studies of this subject have dichotomized the types of teacher questions into open- and closed-ended questions, which tend to oversimplify the role of teacher questioning in the complexity of argumentative dialogue. This dichotomy also prevents teachers from developing multiple roles for argumentation. There is still much to be learned about supporting the development of teacher roles of questioning in promoting student ownership of learning, as well as about the relationship between teacher roles of questioning and the way students use evidence to actively construct and critique knowledge.

In addition, although U.S. national reform documents (NRC, 2007, 2012) have called for argumentation as a core practice of science classrooms across the entire K-12 grade span, the opportunities for early elementary students to participate in scientific argumentation are practically "epistemologically impoverished" (p. 51, Metz, 2011). Deficiencies in the cognitive development of early elementary students have

been misinterpreted as reasons why they are not able to participate successfully in scientific argumentation. Most previous studies have focused on upper elementary (e.g., Martin & Hand, 2009), middle (e.g., Nam, Choi, & Hand, 2011; Chin, 2007), and high school levels (e.g., McNeill, & Pimentel, 2010; Scott, Mortimer, & Aguiar, 2006). Little information currently exists on early elementary classrooms, even though research suggests that early elementary science practices are fundamental for developing cognitive thinking and scaffolding readiness capacities in multiple domains (Lee & Kinzie, 2012).

This multiple-case study aimed to explore and conceptualize the roles of questioning that developed as early elementary teachers attempted to implement an argument-based inquiry approach over four successive years and how those changes in teacher roles impacted on student engagement in argumentative processes. The study was framed by the following research questions: (1) What roles do early elementary teachers adopt in argumentative discourse, especially in whole class discussion, when they use questioning as a tool to engage students in cognitive response over the course of four years? (2) How do the roles teachers adopt impact on students' cognitive response? (3) How do the roles teachers adopt impact on the way students use evidence for dialogical interactions?

Theoretical Framework

Research has suggested that teacher questioning is a major contributing factor to shape the role of teachers for promoting dialogical interaction and students' ownership of ideas (Scott, Mortimer, & Aguiar, 2006). However, Roth (1996) argued that teacher questioning is activity dependent. That is, the function of teacher questions is not just about discussing ideas with students, but the teachers also need to have awareness of the ownership of activity situated in the moment of that context. For example, some classroom activities are still controlled by the teachers even if the ownership of discussion ideas belongs to students (Chin, 2007). In contrast, there are cases where activities are maintained by students but teachers may serve in the role of a coach to nudge students towards an understanding of canonical science knowledge. Thus, the role of teachers' questions should be conceptualized based on ownership of ideas and activities. The following four categories of dialogical approach emerge.

- a. *Teacher's ownership of ideas/ Teachers' ownership of activities* involves the teacher guiding students and directing them to develop ideas and strategies for argument. Teachers control the ideas and activities during dialogue.
- b. *Students' ownership of ideas/ Teacher's ownership of activities* involves encouraging students to develop their own ideas through teacher-led dialogue. The teacher only intervenes in recognizing, comparing, and integrating students' ideas to reach consensus. Teachers control the activity, but the direction of dialogue follows students' ideas.
- c. *Teacher's ownership of ideas/ Students' ownership of activities* involves the teacher allowing students to conduct activities but challenging students' ideas and resolving their difficulties by asking questions. The teacher guides students' ideas during the students' activities.
- d. *Students' ownership of ideas/ Students' ownership of activities* involves the teacher and the students exchanging ideas and developing activities collaboratively through student-led negotiation. Students control their ideas and activities and the teacher is open to learning new concepts.

By analyzing various questioning approaches and the associated teacher–student discourse, we used these four aspects to explicate the critical roles of the teacher in argumentation and to reveal some of the limitations and functions of teacher roles of questioning for establishing dialogical patterns and students' ownership of learning.

Methods

This study took place in three early elementary science classrooms (Brielle and Susan were third grade teachers; Lynette was a second grade teacher) utilizing an argument-based inquiry (ABI) approach (Chen, & Steenhoek, 2014; Martin & Hand, 2009) over a span of four consecutive years. This multi-case study employing a mix methods approach (Creswell, 2003) was designed as a follow-up study after a four-year professional development project that emphasized learning science as a negotiation process by embedding arguments in scientific inquiry activities using the ABI approach. This project attempted to aid elementary science teachers in designing instruction around unit big ideas and provided opportunities for teachers to tie learning theory to pedagogical practice. Three teachers were purposefully selected for this study from the 31 participating teachers. The criteria of selection was that (1) they taught early elementary grade levels, (2) they had no experience with implementing argument-based inquiry in their classrooms before they participated in the project, and (3) they had completed data sources that enabled the researchers to trace their changes over time.

Data Collection

The major data sources were 30 science lessons taught by the three teachers over four consecutive years. Those lessons were recorded by the three teachers when they engaged their students in public negotiation about claims and evidence. Each lesson recorded was 25–45 minutes in length and focused on whole-class discussions after a small group of students presented their written claim and evidence about a concept in science. The total time of Brielle's video-taped classroom observation was 364 minutes; Lynette's was 280 minutes; and Susan's was 337 minutes. The units taught in the three teachers' classrooms were covered by National Standards (NRC, 1996) and State Standards, and included units such as plants, force, the three phases of water, and sound, etc.

Data Analysis

To triangulate the dynamics of and changes in the role of teachers' questioning and the resulting impact on students' engagement in argumentation, data were analyzed through two approaches: (1) the constant comparative method (Strauss & Corbin, 1990) and (2) the enumerative approach (LeCompte & Preissle, 1993). What follows is a description of each approach.

Constant Comparative Method

All 30 science lessons focusing on whole class discussion were first transcribed and each transcript was broken into individual utterances. An utterance was defined as a unique idea contributing to the discussion. The total number of utterances in Brielle, Lynette, and Susan's classrooms were 1991, 1451, and 2407, respectively. Four different coding frameworks were developed to code the utterances to address three research questions respectively: teacher roles for questioning to address Research Question 1; student cognitive responses to address Research Question 2; evidence of quality and dialogical interaction to address Research Question 3. The coding system was established by five graduate students in the field of science education through an interactive process of reviewing the transcripts. Any disagreements were discussed weekly until a consensus was reached.

Teacher roles for questioning: The coding system was established through an interactive process of reviewing the transcripts without utilizing a pre-existing coding system. As a result, the coding system included: *lecture, guide, recognize, compare, integrate, challenge, elicit, exchange, and encourage*. In order to simplify and capture the representation of teacher questioning roles, we attempted to group these nine codes into more comprehensive categories. The nine codes were categorized based upon these four roles. As a result, *lecture* and *guide* were categorized as dispenser; *recognize, compare, and integrate* were categorized as moderator; *challenge and elicit* were categorized as coach; and *exchange and encourage* were categorized as participant.

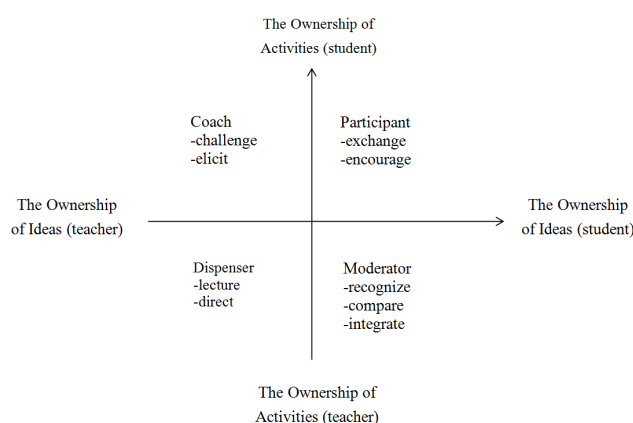


Figure 1. Framework for teacher roles of questioning from two combinations of ownership of ideas and ownership of activity

Student cognitive responses: Codes developed through the analysis included: *retrieve, express, elaborate, reframe, defend, synthesize, challenge, and justify*. In order to categorize the codes into a hierarchical typology, the codes for cognitive processes were clustered into three levels using Bloom's taxonomy. As a result, *retrieve* and *express* were clustered in the low level (knowledge/comprehension); *elaborate* and *reframe* were clustered in the medium level (application/analysis); and *defend, synthesize, challenge, and justify* were clustered in the high level (evaluation).

Evidence of quality: Evidence is an explanation consisting of data and reasoning to show how or why a claim is supported (Chen, Hand, & McDowell, 2013). Two coding schemes were developed to classify the quality of student evidence: *without reasoning* and *with reasoning*. If students simply expressed experience, reported data, or quoted information from books or the Internet as their evidence, it was coded as without

reasoning. If the evidence included a justification for why it supported or rejected a student's claim or evidence, it was coded as with reasoning.

Dialogical interaction: The scheme included three codes: independence, trivialization, and connection. Utterances coded as *independence* represent any reaction that was not connected to evidence or ideas that had been previously raised in classroom discussion. Utterances coded as *trivialization* represent any reaction that was shut down or that ignored previous evidence or ideas. Utterances coded as *connection* represent any reaction that challenged, rejected, supported, or asked for elaboration about evidence or ideas that had been previously proposed during discussion.

Enumerative Approach

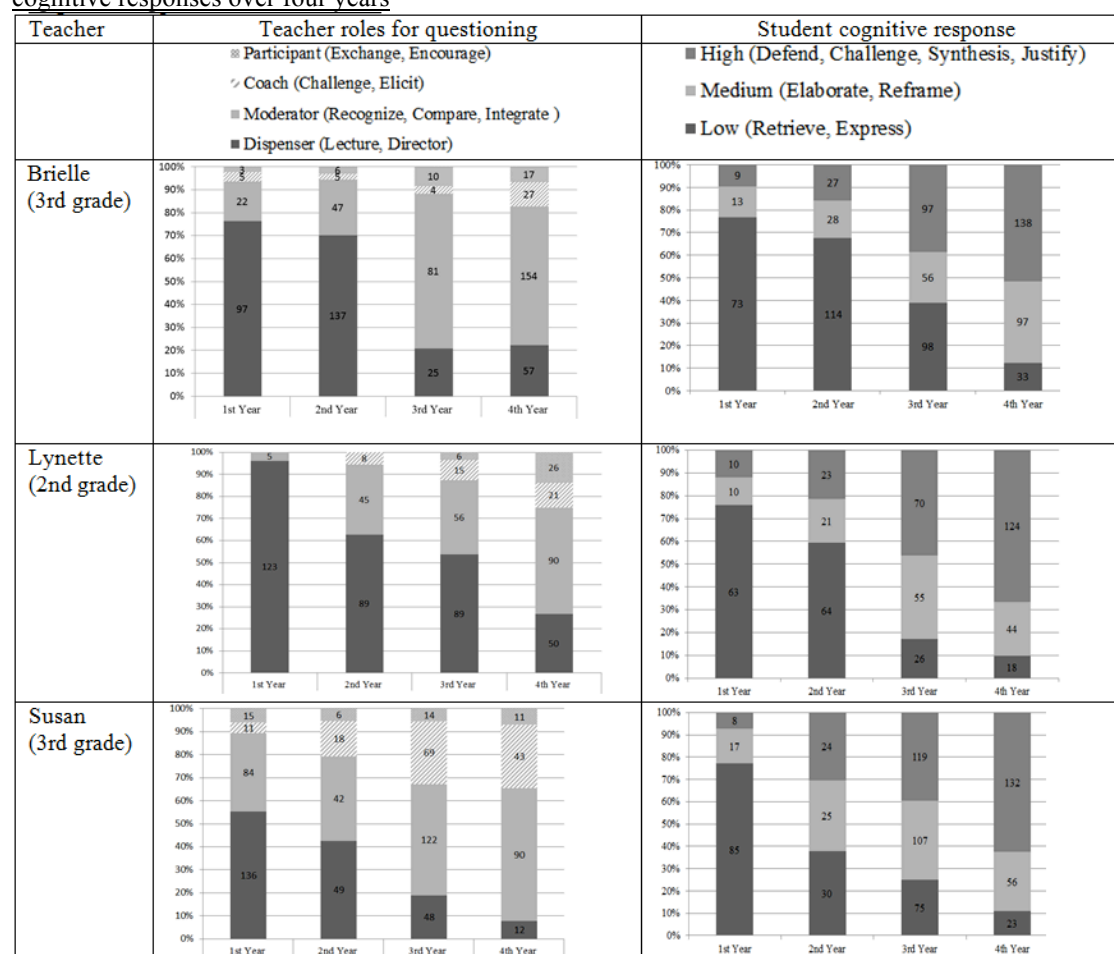
In order to clearly and explicitly portray the changes over a four-year period in teacher use of questioning to engage students in argumentation, an enumerative approach was employed to quantify verbal data (LeCompte & Preissle, 1993).

Findings

Teachers Increasingly Used Multiple Roles in Establishing Argumentative Discourse as They Persistently Implemented an Argument-Based Inquiry Approach

Four different roles in establishing patterns of questioning were increasingly adopted by the three teachers over the four-year period. These included: dispenser, moderator, coach, and participant. In the first year, the most typical role adopted by teachers in classroom discussion was the dispenser, focused on getting a specific response. As Table 1 shows, 76% (193/254), 96% (123/128), and 56% (136/246), respectively, of Brielle, Lynette, and Susan's questioning strategies fell in the role of dispenser in the first year.

Table 1. Number and percentage of teacher roles for questioning used by the three teachers and students' cognitive responses over four years



However, as teachers consistently implemented an argument-based inquiry approach, they developed different roles for questioning in order to establish students' ownership of ideas and activities. The moderator

role was employed to synthesize students' ideas as a consensus, leading the students to the development of canonical science concepts that were consistent with the teachers' lesson goals. This type of questioning is similar to Chin's (2007) "verbal jigsaw" and "semantic tapestry" approach in that the teachers still maintained their position of ownership for the discussion activity but students played active roles in verbalizing their ideas. Taking Brielle as an example, the percentage of utterances coded as moderator increased from 17% (44/254) in the first year to 60% (154/255) in the fourth year.

Another role developed by the three teachers was that of the coach. For example, the coach role in Susan's class increased from 4% (44/246) in the first year to 28% (43/156) in the fourth year. The purpose of this role was to elicit and challenge students to resolve discrepant views to build and strengthen the connections among their preconceptions and a new concept, thereby broadening and deepening their conceptual network.

The participant role was another salient strategy that the three teachers developed for argumentative environments. For example, the coach role in Lynette's class increased from 0% (0/128) in the first year to 14% (26/187) in the fourth year. This role provides students ownership of ideas and activities in which they actively ask questions, pose problems, and seek an explanation as a class.

Taken together, the findings show a series of shifts in teacher roles for questioning from one single role focusing on controlling the ownership of ideas and activities toward multiple roles that provide for students' ownership of ideas and activities.

As Teachers Used Multiple Roles in Establishing Patterns of Questioning, Framing Classroom Interaction, Students' Higher Levels of Cognitive Thinking Was Promoted

The findings show that as these teachers simply adopted the dispenser role for interacting with students, the students mainly engaged in lower-level cognitive activities, such as retrieving scientific vocabularies and providing short answers (See Table 1). While the dispenser role is usually pitched at short answers and lower-level cognition, the moderator, coach, and participant roles usually required much longer student responses to defend, challenge, synthesize, and articulate claims and evidence as well as further prompted higher-level cognition.

Taking Brielle as an example, as her questioning patterns shifted away from a single, more teacher-centered role to multiple, student-centered roles, students' higher-level cognitive thinking was encouraged. Evidence of the shift in students' cognitive response is supported by the increase of their utterances categorized as medium and high, which grew from 14% (13/95) and 9% (9/95), respectively, in the first year to 36% (97/268) and 51% (138/268) in the fourth year.

In the following discussion, several examples will be provided to demonstrate the differences and trends in the four roles of teacher questioning over the four years as well as to illustrate how those questioning patterns impacted students' cognitive levels of response.

As Teachers' Patterns of Questioning Changed, the Frequency of Students' Talk Increased and the Dialogical Interaction Between Students and Teachers Became More Evidence-Based and Connected

As the teachers' roles in establishing patterns of questioning were examined further, the data showed a parallel shift between teachers' questioning patterns and students' verbal participation. Figure 2 shows that students' argumentative utterances in an hour increased from the first to the fourth year (Brielle's class from 140.1 to 301.1; Lynette's class from 177.6 to 295; Susan's class from 199.8 to 321.1). Figure 3 portrays the trend of the changes in students' verbal participation in an hour. These changes in the relative frequency of students' argumentative utterances suggest that the learning environment became more student-centered and unthreatening, and that thus students were more willing to articulate their arguments.

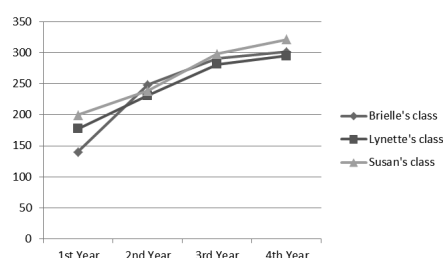


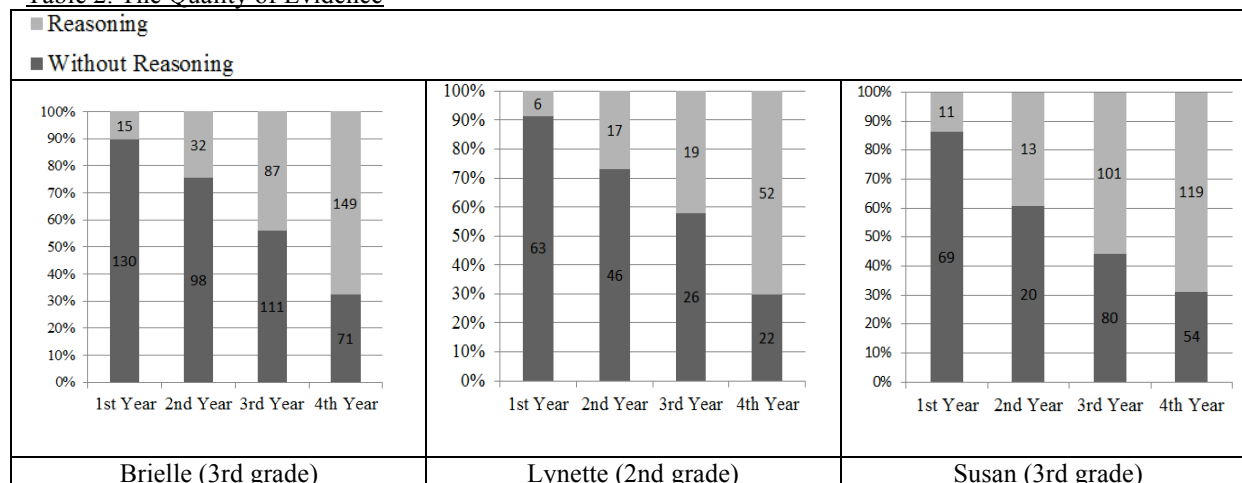
Figure 2. The Trend of Students' Utterances over Four Years

As students' oral participation increased over the years, two distinct shifts regarding the quality of discussion were found: (1) students came to discuss their ideas based on evidence, and (2) students began to connect pieces of evidence to other evidence they had found.

Evidence-Based Discussion

Table 2 shows that the proportion of utterances where students used evidence to discuss their ideas increased from 10% (15/145) in Brielle's classroom, 9% (6/69) in Lynette's classroom, and 14 % (11/80) in Susan's classroom in the first year to 68 % (149/220), 70 % (52/74), and 69% (119/173), respectively, in the fourth year.

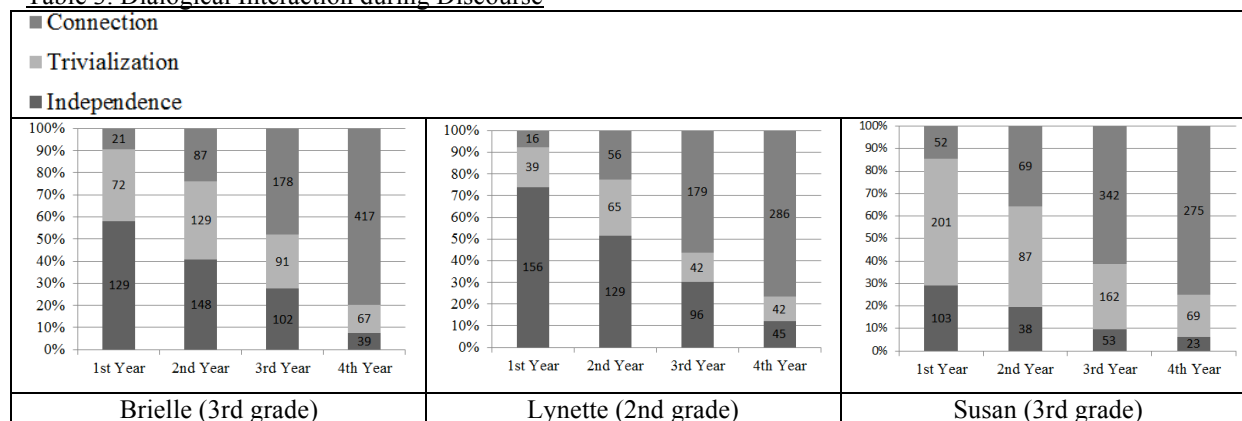
Table 2. The Quality of Evidence



Dialogical Interaction

Figure 5 presents our analysis of the interactions during classroom discussion. In the first year, only 10% (21/222) of the utterances during discussion in Brielle's class, 8% (16/211) in Lynette's class, and 15% (52/356) in Susan's class were linked to previous ideas in that the ideas reworded, justified, clarified or posed an elaborating question. However, in the fourth year, the utterances connected to previous ideas increased to 80% (417/523) in Brielle's classroom, 77% (286/373) in Lynette's classroom, and 75% (275/ 367) in Susan's classroom.

Table 3. Dialogical Interaction during Discourse



Summary of Findings

Table 4 summarizes the findings across the four major features that emerged along the time span involved. The first feature was that during the four years of the study the teachers increasingly used four roles to establish argumentative discourse; this is referred to as an increasing richness of the teacher role, rather than relying only on the dispenser role as they had in the first year. As the shift in the use of multiple roles occurred, students' cognitive responses were promoted to higher and more complex levels. Students were also observed constructing and critiquing ideas in a more evidence-based form. The final feature was that students also became more likely to link their ideas back to previous contributions made by the teacher and their peers through defending, challenging, synthesizing, and justifying. The findings imply that to promote student engagement teachers should go beyond one single role for questioning and should play multiple roles to tackle different situations by considering student ownership of ideas and activities (Crawford, 2000; Walshaw & Anthony, 2008). With the support of teacher questioning, this study suggests that even early elementary students

can successfully engage in productively argumentative practices.

Table 4. Trend of features of teacher questions and classroom discourse over four years

	Year 1	Years 2 & 3	Year 4
Teacher Role of Questioning	Single role: Teachers focused on dispenser role to lecture and direct classroom discourse. The ownership of ideas and activities controlled by the teachers.	Transition: Teachers gradually developed moderator, coach, and participant roles in questioning to promote students' diverse dialogue. The ownership of ideas and activities shifted from the teacher to the students.	Richness: Teachers used multiple roles of questioning to support students' thinking and conversation. Students had many opportunities to take ownership of ideas and activities.
Students' Cognitive Response	Simplicity: Students' discourse was restricted to low-level cognition, such as retrieving and expressing ideas.	Transition: Students' discourse shifted to medium- and high-level, such as elaborating, reframing, defending challenging, synthesizing, and justifying ideas.	Complexity: Students' discourse was in high-level cognition, such as defending, challenging, synthesizing, and justifying ideas.
Evidence-based Discussion	Simple: High proportion of utterances focused on simple response.	Transition: Students came to use evidence to discuss their ideas.	Evidence-based: High proportion of utterances focused on evidence.
Dialogue Interaction	Independence: Students' utterances were independent. Classroom discussion was dominated by teacher talk.	Transition: Students came to link back to what other students had said. Student voices increased, and teacher's voice decreased.	Connection: Students actively connected their utterances to previous ideas. Classroom discussion was dominated by student talk.

Discussion

While many studies on teacher role of questioning have simply dichotomized the type of questions into open/closed end and disconnected the two types of questions, this current study conceptualizes four roles of teacher questioning based on the tension between ownership of both ideas and activities to better represent the complexity of argumentative environment. Significant is that the ABI approach requires students to engage in knowledge construction and critique; therefore, the function of the shifting roles of teacher questioning is to fully implement an approach that pushes students to be involved in and take ownership of construction and critique in generating arguments. The function of teacher questions is not just about discussing ideas with students, but the teachers also need to have awareness of the ownership of activity situated in the moment of that context. As the examples demonstrate in the four role, the purpose and function of teacher questions can be varied depend on the tension between ownership of ideas and activities.

Although the moderator, coach, and participant roles promoted higher-level cognition in student responses, the three teachers in this study also continued to adopt the dispenser role in the fourth year (Brielle: 22%; Lynette: 31%; Susan: 8 %). This finding raises essential questions about whether teachers should eliminate the dispenser role to successfully engage students in argumentative practices. Oliveira (2010) suggested that teachers should aim to better understand the important cognitive functions of different questioning strategies and the situations in which they can apply those strategies to help students effectively and productively develop conceptual understandings. Thus, instead of simply abandoning the adoption of the dispenser role, educators should realize how the different roles can elicit different cognitive results for different purposes and contexts. More research is needed to determine both the best times and contexts to use the different roles in argumentative environments and the most appropriate sequence for using different roles to advance students' conceptual understanding.

The development of different roles to support student argumentation requires time. In this study, it took the three teachers years to shift their classroom environments to a more reform-based model. An increasing body of empirical evidence suggests that it takes more than 18 months before significant shifts in teachers' questioning pedagogy are observed (Martin & Hand, 2009). Therefore, this study supports Luft's (2001) suggestion that teacher professional development should be designed systematically and should consist of ongoing training. The various roles of teacher questioning should be viewed as an essential component of

professional development in order to expedite the shift in teachers' practices from a teacher-centered approach to a more student-centered orientation.

References

- Ball, D. L. (2000). Bridging practices: Intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education*, 5(3), 241-247.
- Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 National Survey of Science and Mathematics Education*. Horizon Research, Chapel Hill, NC.
- Berland, L. K., & Reiser, B. J. (2011). Classroom communities' adaptations of the practice of scientific argumentation. *Science Education*, 95(2), 191-216.
- 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, Y.-C., Hand, B., & McDowell, L. (2013). The effects of writing-to-learn activities on elementary students' conceptual understanding: Learning about force and motion through writing to older peers. *Science Education*, 97(5), 745-771.
- Chen, Y.-C., & Steenhoek, J. (2014). Arguing like a scientist: Engaging students in core scientific practices. *The American Biology Teacher*, 76(4), 231-237.
- Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of Research in Science Teaching*, 44(6), 815-843.
- Crawford, B. A. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science Teaching*, 37(9), 916-937.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publications.
- Hmelo-Silver, C. E., & Barrows, H. S. (2008). Facilitating collaborative knowledge building. *Cognition and Instruction*, 26(1), 48-94.
- LeCompte, M. D., & Preissle, J. (Eds.). (1993). *Ethnography and qualitative design in educational research* (2nd ed.). San Diego, CA: Academic Press.
- Lee, Y., & Kinzie, M. (2012). Teacher question and student response with regard to cognition and language use. *Instructional Science*, 40(6), 857-874.
- Luft, J. A. (2001). Changing inquiry practices and beliefs: The impact of an inquiry-based professional development programme on beginning and experienced secondary science teachers. *International Journal of Science Education*, 23(5), 517-534.
- Martin, A. M., & Hand, B. (2009). Factors affecting the implementation of argument in the elementary science classroom: A longitudinal case study. *Research in Science Education*, 39(1), 17-38.
- McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education*, 94(2), 203-229.
- Metz, K. E. (2011). Disentangling robust developmental constraints from the instructionally mutable: Young children's epistemic reasoning about a study of their own design. *Journal of the Learning Sciences*, 20(1), 50-110.
- Nam, J., Choi, A., & Hand, B. (2011). Implementation of the science writing heuristic (SWH) approach in 8th grade science classrooms. *International Journal of Science and Mathematics Education*, 9(5), 1111-1133.
- National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academy Press.
- National Research Council. (2012). *A Framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Oliveira, A. W. (2010). Improving teacher questioning in science inquiry discussions through professional development. *Journal for Research in Science Teaching*, 47(4), 422-453.
- Scott, P. H., Mortimer, E. F., & Aguiar, O. G. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Science Education*, 90(4), 605-631.
- Strauss, A., & Corbin, J. (1990). Open coding. In A. Strauss & J. Corbin (Eds.), *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed., pp. 101-121). Thousand Oaks, CA: Sage.
- Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, 78(3), 516-551.

Hear What They Say and Watch What They Do: Predicting Valid Mathematical Proofs Using Speech and Gesture

Elizabeth Pier, University of Wisconsin-Madison, epier@wisc.edu
Candace Walkington, Southern Methodist University, cwalkington@smu.edu
Caroline Williams, University of Wisconsin-Madison, ccwilliams3@wisc.edu
Rebecca Boncoddio, Central Connecticut State University, boncoddio@ccsu.edu
Jessica Waala, Martha Wagner Alibali and Mitchell J. Nathan, University of Wisconsin-Madison,
waala@wisc.edu, mwalibali@wisc.edu, mnathan@wisc.edu

Abstract: In mathematics, practices of proof are notoriously difficult for learners to adopt. In prior work, we found that when providing verbal justifications, learners' speech patterns predict whether their justifications are mathematically sound. However, current views on the embodied nature of cognition suggest that actions and speech may co-constitute reasoning processes. The current study investigated whether the *gestures* learners use while formulating proofs also predict proof validity. 120 undergraduates provided verbal justifications for two mathematical tasks. We analyzed speech patterns in participants' justifications using text analysis software, and we coded participants' gestures as *dynamic* or *static*. Results showed that dynamic gestures were correlated with mathematically valid proofs and with specific patterns of speech. A stepwise logistic regression model found that speech and gesture separately account for unique variance in a model predicting proof validity, indicating that gesture contributes to mathematical reasoning in an abstract task domain.

Introduction

Recent research has highlighted ways that mathematical cognition is *embodied*, or formulated through perception and action, and has specifically identified gesture as an important cognitive resource for mathematics learning. In mathematics classrooms, teachers and students spontaneously use gesture to connect disparate mathematical representations and ideas (Alibali et al., 2014; Alibali & Nathan, 2012; Nathan et al., 2013). Teachers' strategic use of gestures can support appropriate mathematical generalizations (Radford, 2003), and attending to students' gestures is important when evaluating their mathematical reasoning (Boncoddio et al., 2013; Nemirovsky & Ferrera, 2009; Williams et al., 2012). The present study is part of an emerging research program investigating how gesture and action can influence students' mathematical cognition. We identify the different types of speech and gestures that students use as they generate and communicate mathematical proofs. We examine how these speech and gesture categories relate to mathematically valid proofs in order to better understand how learners can be supported in adopting proof practices.

The ICLS theme, "Learning and Becoming in Practice," is particularly relevant for this line of investigation. Recent *in situ* classroom research has shown that teachers and students in mathematics and engineering classrooms spontaneously use gesture to enact connections across representations (Alibali et al., 2014; Nathan et al., 2013). In some cases, by using gesture and body-based action, they can effectively *become* the mathematical objects and relations they intend to convey. For example, students and teachers often use their bodies to dynamically simulate mathematical objects and relations, such as representing changes in linear slopes with an arm (Alibali & Nathan, 2012). When the body is used as a tool for depicting and understanding a fluid, moving transformation of a mathematical entity, we refer to this as a *dynamic depictive gesture* (Walkington et al., this volume).

In this work, we examine how the production both of dynamic depictive gestures and of particular speech patterns (e.g., "if... then" statements) is associated with valid proofs. Our goal is to understand the complex interplay between learners' use of gesture and speech as they produce mathematical justifications. While many studies have examined discourse patterns or gesture patterns in isolation, research on how these two modalities interact and work together to form complex patterns of interpersonal, embodied communication is more rare (though, see Radford, 2003). Here, we take a novel approach in which we use quantitative measures of speech and gesture to predict accuracy of justifications for mathematical tasks. The emergence of recent software tools for automated text analysis, coupled with the identification of dynamic gestures during mathematical reasoning from video-based classroom studies, gives us leverage in understanding how speech and gesture contribute to proof practices.

Theoretical Framework

The Importance of Proof Practices in Mathematics

Both the Common Core State Standards (2010) and the National Council of Teachers of Mathematics (NCTM) Standards (2000) identify constructing mathematical arguments and engaging in proof practices as crucial components of mathematical understanding. In addition, constructing valid proofs is central to both the professional domain of mathematics and mathematics education (Schoenfield, 1994), as it is the means by which mathematicians communicate key ideas and novel understandings to each other. However, student difficulties with proof are well documented (e.g., Harel & Sowder; Dreyfus, 1999; Knuth, Choppin, & Bieda, 2009); therefore, there is a need to explore how to better identify and support valid proof practices. We follow Harel and Sowder's (2005) definition of proof as the process of reasoning about a conjecture and communicating the believed truth or falseness of the conjecture, and consequently we use the terms *proof* and *justification* interchangeably. According to Harel and Sowder (2005), mathematically valid proofs must be (1) general in nature, such that they hold for all cases, (2) based on logical inferences in which conclusions are drawn from valid premises, and (3) based on operational thought, in which there is progression through goals and subgoals. Harel and Sowder's (2005) taxonomy focuses on traditional methods of communicating proof with written and spoken language. Although mathematical proofs often are verbalized instead of formally written in K-12 classrooms, the content and structure of students' speech alone may not provide a complete picture of students' mathematical knowledge. In particular, studies have shown that people often express information in gestures that they do not express in speech (Alibali & Goldin-Meadow, 1993; Alibali & Kita, 2010; Nathan & Johnson, 2010; Williams et al., 2012). We therefore argue that the body is an important modality in which complementary or even novel mathematical information can be communicated. In this work, we extend Harel and Sowder's (2005) framework by invoking views of embodied cognition to include gestures as an additional means of supporting both mathematical cognition and mathematical communication.

Gestures and Embodiment in Proof Practices

Theories of embodied cognition reject traditional views of cognition as based on manipulating amodal symbol systems in the brain based on purely syntactic rules (Shapiro, 2011). In contrast, within an embodied framework, cognition comprises a mutual feedback loop between the brain and the body. For example, gestures may reveal cognitive processes, and the process of gesturing in turn affects those cognitive processes (Alibali & Nathan, 2012; Goldin-Meadow, Cook, & Mitchell, 2009; Hostetter & Alibali, 2008; Shapiro, 2011). Several embodied cognition scholars (Barsalou, 2008; Hostetter & Alibali, 2008) acknowledge the role of gesture as simulated action that reenacts perceptual, motor, and mental states that arise from prior experience. In this work, we investigate whether certain types of gesture are associated with valid proof construction. We draw upon work by Göksun, Goldin-Meadow, Newcombe, and Shipley (2013), who categorized gestures used in a mental rotation task as "static" if they referred to individual objects, and "dynamic" if they indicated movement such as rotation or direction. Newcombe and Shipley (2012) also differentiate static gestures, those that *indicate* spatial features or locations, versus dynamic gestures that *transform* spatial features or locations. We build upon these definitions by operationalizing simulated actions as *static depictive gestures* if the hands display a fixed, unchanging mathematical object and *dynamic depictive gestures* if the hands are used to enact transformations on mathematical entities and relations. We refer to these in the rest of the paper simply as "dynamic" or "static" gestures, but we are referring to the dynamic and static nature of the *objects*, and not whether or not the gesture itself contains movement. We make this distinction because we believe that dynamic gestures are closely related to learners' use of simulated action to represent multiple cases and fluid transformations between depicted entities, which can promote sophisticated mathematical reasoning. Furthermore, our prior work has found a significant correlation between using dynamic gestures and generating valid mathematical justifications across six different mathematical tasks (Walkington et al., this volume).

Use of Speech in Proof Practices

Although theories of embodied cognition underscore the importance of the body when engaging in cognitive tasks such as proof production, proofs themselves are linguistic in nature. Mathematical proofs can be thought of as a specific kind of disciplinary discourse practice (e.g., Gee, 2007; Gresalfi & Cobb, 2006), and in K-12 classrooms, proofs often take spoken—rather than formal, written—forms (Healy & Hoyles, 2000). Here we focus on mathematical proofs that are communicated through spoken language and gesture, hypothesizing that specific features of participants' speech patterns may be important for constructing valid proofs. Speech patterns can be analyzed at many levels for a variety of features. For example, we could quantitatively assess how often people use pronouns, adjectives, or words relating to different topics, such as work or socializing, when they are speaking. We could also quantify how much overlap there is between different words or sentences in a block of continuous speech, or what tense a person tends to use. Indeed, the recent emergence of software-based text-mining tools has allowed us to better understand characteristics of human speech that are associated with

learners' making gains in conceptual understanding (see Jeon & Azevedo, 2007; Williams & D'Mello, 2010). When people are asked to verbally construct mathematical justifications, some of these quantitative measures of their speech may be indicative of a well-constructed, logically sound mathematical argument. By understanding the verbal properties of valid proofs, we can better grasp how to scaffold learners to adopt successful proof practices.

Hypotheses and Predictions

This study investigates how gestures, speech patterns, and the production of valid proofs are associated with one another. In doing so, this study seeks to deepen our understanding of whether—and how—gestures provide additional information, beyond speech, about the validity of students' verbally communicated proofs. This leads us to formulate two hypotheses: **Hypothesis 1** states that speech and gesture may constitute equivalent but distinct manifestations of the same underlying cognitive processes that result in participants' generating valid proofs. In other words, while both speech and gesture may each predict valid proofs separately, together they provide redundant measures of the same underlying construct relating to valid mathematical reasoning. Alternatively, **Hypothesis 2** claims that speech and gesture may indicate distinct underlying cognitive processes involved in proof production. For instance, gestures may reveal distinct aspects of the learner's mental action-based simulations that are not conveyed in their speech patterns. In order to test these competing hypotheses, we formulated the following research questions. **Research Question 1** asks, Is the use of dynamic gesture correlated with the validity of one's proof? This enables us to investigate whether gesture does, in fact, provide information about proof validity. **Research Question 2** asks, Are particular speech patterns correlated with the validity of one's proof? This allows us to capture what characteristics of speech tend to indicate logical and valid mathematical reasoning. **Research Question 3** asks, Are dynamic gestures correlated with the use of particular speech patterns during the proof process? If speech and gesture are highly correlated with one another, but accounting for both does not improve the model's prediction of the validity of participants' proofs, then this would provide evidence for Hypothesis 1. In order to allow for the falsification of Hypothesis 1 in favor of Hypothesis 2, **Research Question 4** asks, Do speech and dynamic gestures contribute *uniquely* to models for predicting proof validity? An affirmative response would indicate support for Hypothesis 2, which stipulates that gesture and speech reveal distinct underlying processes of proof practices.

Methods

Undergraduates ($N = 120$) (M age = 19.2 years, 51% female) enrolled in a psychology course at a large Midwestern university were prompted to verbally justify two tasks from distinct mathematical domains: one relating to planar geometry, the other to an inference about parity in the number system. Participants read each conjecture (order was counterbalanced) out loud and were prompted to think aloud as they verbalized their justifications. The first task was, "Mary came up with the following conjecture: 'For any triangle, the sum of the length of any two sides must be greater than the length of the remaining side.' Provide a justification as to *why* Mary's conjecture is true or false." The second task was, "An unknown number of gears are connected together in a chain. If you know what direction the first gear turns, how could you figure out what direction the last gear turns? Provide a justification as to *why* your answer is true."

Coding of Proof Accuracy

Participants were asked to think aloud as they provided their justifications, and videotapes of the sessions were uploaded into Transana, a software program that allows for transcribing and analyzing video data (Woods & Fassnacht, 2012). After transcription, the videos were split into 240 separate transcripts (one for each task for each participant). Based on the verbal transcript and video, each justification was each coded as *Valid* (1) or *Invalid* (0) based on Harel and Sowder's (2005) definition of valid proofs. Examples of spoken valid and invalid proofs for each of the two tasks are shown in Table 1. Overall, 40.8% and 50% of participants provided valid proofs for the gear and triangle tasks, respectively. Inter-rater reliability for coding the validity of participants' justifications was high, with a Kappa value of 0.84 for a 20% subset.

Coding of Speech

The 240 transcripts were analyzed using the LiWC (Pennebaker, Booth, & Francis, 2007) and Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004) text analysis software programs. LiWC comprises various dictionaries of words assigned to different topic categories, such as social process words (e.g., words relating to family or friends) and cognitive process words (e.g., words describing causation or certainty). LiWC's output consists of the percentage of words the participants used from each dictionary. LiWC provides a measure of the content of a text, whereas the second text-mining program, Coh-Metrix, analyzes the quality of a text. Coh-Metrix contains 108 different indicators of text readability (for a full list, see <http://cohmetrix.memphis.edu>), which are broadly organized into categories such as syntactic simplicity, word concreteness, and deep cohesion. Coh-Metrix's output provides continuous quantitative measures of the degree to which these characteristics are

present in a text. Therefore, using both software programs in conjunction allows for analysis of the technical aspects of the language and readability gathered from Coh-Metrix, as well as the content and topic of the language that participants use from LiWC, providing a more holistic picture of the nature of participants' speech. We included all speech categories from LiWC and Coh-Metrix in our analyses, with the exception of categories specific to the number of paragraphs or number of sentences, as we were using natural speech instead of written text, and these demarcations vary greatly by transcription norms.


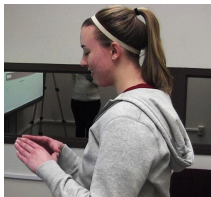

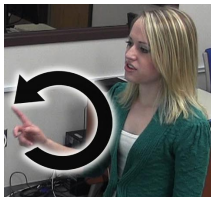
Table 1: Examples of spoken valid and invalid participant proofs from each task

Proof Type	Example
Triangle, Valid	"Mary's conjecture is true, because if the one side is long--is longer than the sum of the other two sides then the other two sides won't be able to touch at the top and it won't be a triangle."
Triangle, Invalid	"That isn't true. Uh it's false because you could have a triangle where one side is very long and the other two sides are shorter, um very short, and so they add up to a length that is shorter than the longest side."
Gear, Valid	"Um obviously the gear after the first one turns in the opposite direction and the next one turns in the opposite direction and so on and so on, so I guess if there's an odd number of gears it will turn in the same direction as the first gear and if there's an even number of gears it'll turn in the opposite direction."
Gear, Invalid	"Um I feel that all the gears should turn the same way because it's a chain reaction so it should turn in the same direction as the first gear."

Coding of Gestures

The 240 video clips were also coded for whether participants generated dynamic or static gestures while engaging in proof activities. For both the triangle and gear tasks, we coded a gesture as *Static* if participants indicated a single entity with their hands, but coded it as *Dynamic* if they indicated multiple entities that were connected together via a depicted relationship. For the triangle task, an example of a *Static* gesture would be indicating a side of a triangle or a full triangle, whereas a *Dynamic* gesture would be fluidly depicting several different sized triangles in a row to show a transformation, or showing a single triangle that grew in size. For the gear task, an example of a *Static* gesture would be portraying a single gear moving in one direction, whereas a *Dynamic* gesture would be two gears moving in the same or in opposite directions, or a chain of gears turning in alternate directions. Note that the distinction is not between gestures that "move" versus those that "don't move," since gestures that don't move can still simulate action or perceptual change. Instead, the distinction lies in whether participants used gestures to represent multiple mathematical objects that were related to one another. Table 2 provides images of participants creating each of these gesture types. A complete clip was coded as *Dynamic* (1) if it contained any instances of dynamic gestures. A clip was coded as *Not Dynamic* (0) if it contained only static gestures or no gestures. Our Kappa reliability value for a subset of 30 video clips (12.5% subset from a total of 240 videos; 15 gear, 15 triangle) was 0.85. This was for the categories of "No Gestures," "Static Only" and "Dynamic," which were assigned holistically to the entire clip (both attempts).

Table 2: Examples of dynamic and static gestures

Task	Example of Dynamic Gesture	Example of Static Gesture
Triangle	 <p><i>Participant uses both hands to make two sides of a triangle and fluently moves from making a flattened triangle to a normal triangle.</i></p>	 <p><i>Participant uses both hands to create a full triangle that does not move or change.</i></p>
Gear	 <p><i>Participant uses both hands to show two gears moving in opposite directions.</i></p>	 <p><i>Participant uses right index finger to show a single gear turning in one direction.</i></p>

Results and Discussion

To examine relations among gesture, speech categories, and proof validity, we calculated Pearson correlation coefficients and examined those correlations that were both in the same direction (positive or negative) and significant for both the gear and triangle tasks. The following three sections detail the results of those analyses.

Dynamic Gestures and Valid Proofs

In order to answer Research Question 1, we investigated the correlation between dynamic gesture and proof validity. Results showed that participants who used dynamic gestures were more likely to produce valid proofs than those who did not, both on the triangle task ($r = 0.454$, $p < .001$) and on the gear task ($r = 0.255$, $p = .005$)—that is, across two distinct mathematical domains. We also looked at the correlations between static gesture and valid proofs; static gestures were significantly negatively correlated with valid proofs for the triangle task ($r = -0.183$, $p = .045$), but were not significantly related to the validity of proofs for the gear task. However, a prior exploratory study showed that the relationship between dynamic gestures and valid proofs holds across six different mathematical tasks (Walkington et al., this volume). By answering Research Question 1 affirmatively, we do not directly falsify either of our competing hypotheses, but it was important to establish the relationship between dynamic gesture and proof validity prior to answering our other research questions.

Speech Categories and Valid Proofs

To answer Research Question 2, we investigated the correlation between proof validity and the speech categories derived from LiWC and Coh-Metrix. In prior analyses of this data set (Pier et al., 2014), we found that four speech categories were predictive in regression models of valid proofs across the two tasks: *Present tense* verbs, lexical diversity (as measured by the *type-token ratio* for content words), *discrepancy* words, and *temporal connectives*. Of those four categories, use of the *present tense* and *type-token ratio* were each negatively correlated with valid proofs, while use of *discrepancy* words and use of *temporal connectives* were each positively correlated with valid proofs. Table 3 describes each of the speech categories that were significantly correlated with valid proofs.

Table 3: Speech categories significantly correlated with valid proofs

Speech Category	Description of Speech Category	Triangle Task	Gear Task
<i>Present tense</i> ^L	Verbs in the present tense	$r = -0.280^{**}$	$r = -0.328^{***}$
Lexical diversity (<i>type-token ratio</i>) ^C	Ratio of the number of unique words (<i>types</i>) divided by the number of times that word occurs (<i>tokens</i>); higher ratios indicate more unique words	$r = -0.280^{**}$	$r = -0.500^{***}$
<i>Discrepancy</i> words ^L	Words in LiWC <i>Discrepancy</i> dictionary, e.g., “could,” “would,” “should”	$r = 0.264^{**}$	$r = 0.202^*$
<i>Temporal Connectives</i> ^C	Words connecting clauses that indicate time order, e.g., “before,” “then,” “after”	$r = 0.265^{**}$	$r = 0.240^{**}$
Note. $*p < .05$, $**p < .01$, and $***p < .001$. ^L is a measure from LiWC, and ^C a measure from Coh-Metrix.			

Following the methodology used by Pier et al. (2014), we systematically investigated the 20 transcripts that scored highest and the 20 that scored lowest on each category in order to determine which aspects of the transcripts seemed to be associated with high or low scores on each of these linguistic categories. These analyses revealed that the negative correlation between proof validity and present tense was due to participants’ use of self-conscious statements such as, “I don’t know,” or “I don’t understand.” The negative correlation with type-token ratio indicated that participants with high type-token ratios verbalized more varied, unrelated words, lacking continuity of ideas. In contrast, participants with a low type-token ratio repeated a consistent set of words related to reasoning through the conjecture. Finally, using both discrepancy words and temporal connectives stemmed from more conditional (i.e., “if...then”) statements indicating participants were making logical inferences—one of the features of valid proofs per the proof scheme from Harel and Sowder (2005).

Dynamic Gestures and Speech Categories

In order to answer Research Question 3, whether gesture use is correlated with speech patterns during the proof process, we investigated the correlations between dynamic gesture and speech categories from LiWC and Coh-Metrix. Although we also calculated correlations between static gesture and speech categories, none of these were both statistically significant and in the same direction (positive or negative) for both tasks. Table 4 describes each of the speech categories that were significantly correlated with dynamic gestures. Four speech categories were negatively correlated with dynamic gestures (*present tense*, *quantifiers*, *insight* words, and *cognitive processes* words), whereas three speech categories were positively correlated with dynamic gestures (*deep cohesion*, *all connectives*, and *temporal connectives*). Thus, not only are dynamic gesture and speech each

individually correlated with valid proofs, but they are correlated with one another; furthermore, many of the same categories that were found in prior work to be significantly predictive of proof validity in speech are also significantly correlated with dynamic gesture (i.e., *present tense* and *temporal connectives*) (Pier et al., 2014). However, in order to evaluate our competing hypotheses, we need to address Research Question 4: whether speech and gesture explain *unique* variance in determining proof validity, or whether they instead explain overlapping, redundant variance in proof validity.

We again systematically investigated the transcripts with the 20 highest and 20 lowest scores for each speech category. We found that when participants frequently mentioned their own mental state or cognitive processing, they were not likely to produce dynamic gestures relevant to the proofs, since they seemed to focus more on their thinking than on the content of the proof. Such patterns resulted in negative correlations between dynamic gestures and the categories of *present tense*, *insight* words, and *cognitive processes* words.

Table 4: Speech categories significantly correlated with dynamic gesture

Speech Category	Description of Speech Category	Triangle Task	Gear Task
<i>Present tense</i> ^L	Verbs in the present tense	$r = -0.188^*$	$r = -0.207^*$
<i>Quantifiers</i> ^L	Words referring to general quantities, e.g., “all,” “more,” “greater”	$r = -0.184^*$	$r = -0.200^*$
<i>Insight words</i> ^L	Words in LiWC <i>Insight</i> dictionary, e.g., “think,” “know,” “understand”	$r = -0.214^*$	$r = -0.337^{***}$
<i>Cognitive processes words</i> ^L	Words in LiWC <i>Cognitive Processes</i> dictionary, e.g., “ought,” “know,” “cause”	$r = -0.228^*$	$r = -0.321^{***}$
<i>Deep cohesion</i> ^C	Incidence of causal and intentional connectives, indicating higher coherence	$r = 0.257^{**}$	$r = 0.217^*$
<i>All connectives</i> ^C	Incidence of all connective words, i.e., conjunctions	$r = 0.267^{**}$	$r = 0.225^*$
<i>Temporal connectives</i> ^C	Words connecting clauses that indicate time order, e.g., “before,” “then,” “after”	$r = 0.248^{**}$	$r = 0.214^*$

Note. $^*p < .05$, $^{**}p < .01$, and $^{***}p < .001$. ^L is a measure from LiWC, and ^C a measure from Coh-Metrix.

Additionally, when participants used temporal connectives and connectives more generally to cohesively link together different statements, they tended to produce dynamic gestures. Both dynamic gestures and verbal statements with connective words typically express relationships between entities; thus, reasoning about relationships relevant to the proof content was manifested in the production both of cohesive speech and dynamic gestures. This drove the positive correlations between dynamic gesture and the speech categories of *deep cohesion*, *all connectives*, and *temporal connectives*.

Stepwise Regression Analysis

To answer Research Question 4, we assessed whether speech and gesture each explained unique variance in the models’ predicting proof validity. We ran two stepwise logistic regression analyses using the *lmer()* function (Bates & Maechler, 2010) in the *R* software environment. In each model, participant was a random effect and task (i.e., gear, triangle) was a fixed effect. The outcome was whether the participant generated a valid proof for the task, coded as a 0 or 1. Interactions were tested for significance, but none were present in the final model. In the first analysis, we added predictors in the following order: Dynamic gesture, speech categories significantly associated with dynamic gestures, and speech categories significantly associated with valid proofs. Predictors were tested for inclusion in the model using the *anova()* function, which uses a chi-square reference distribution to test for significant reductions in deviance. In the second analysis, we added predictors in the following order: Speech categories significantly associated with valid proofs, speech categories significantly associated with dynamic gestures, and dynamic gestures. Thus, in one analysis we added terms for dynamic gestures into the model first, and in the other analysis we added dynamic gestures last into the model. Both analyses resulted in the same final model, so we only discuss the results in terms of one stepwise regression analysis. Furthermore, we included controls for the word count of each proof, but this factor was not a significant covariate.

Table 5 provides the results of the stepwise regression. Three of the speech categories were significant predictors of valid proofs: *discrepancy words* ($z = 4.21$, $p < .001$), *type-token ratio* of content words ($z = -3.08$, $p = .0020$), and *cognitive processes* words ($z = -3.29$, $p = .001$). Also, use of dynamic gestures was a significant, positive predictor of valid proofs ($z = 3.00$, $p = .003$). Adding in the speech category predictors reduced the overall deviance of the model by 24.1% ($\chi^2(3) = 74.75$, $p < .001$), and adding the dynamic gesture predictor reduced the deviance by an additional 7.0% ($\chi^2(1) = 16.44$, $p < .001$). Overall, language and gesture predictors combined accounted for approximately 30% of the variance in whether participants generated valid proofs.

The regression analyses thus answer Research Question 4 by demonstrating that dynamic gestures uniquely predict the validity of proofs, which provides support for Hypothesis 2. This means that while both

speech and gesture independently predict whether or not a participant verbalized a valid proof, there is distinct information in each modality. In other words, gesture and speech are not redundant, but instead appear to provide unique but related insight into the cognitive processes involved in successful proof construction. Thus, consideration of *both* speech and gesture is important to understanding the depths of students' understandings—and misunderstandings.

Table 5: Stepwise regression analysis results

	Estimate	Std. Error	z value	p value	Sig.
(Intercept)	18.28	8.42	2.17	0.0300	*
<i>Discrepancy</i> words	3.17	0.75	4.21	<.001	***
<i>Type-token Ratio</i> (LDTTRc)	-49.25	15.97	-3.08	0.0020	**
<i>Cognitive processes</i> words	-0.94	0.29	-3.29	0.0010	**
Dynamic Gestures	8.92	2.98	3.00	0.0027	**
Note. * $p < .05$, ** $p < .01$, and *** $p < .001$.					

Conclusion

The findings in this study suggest that gesture and speech are not merely interchangeable windows into cognitive processes. Instead, gesture and speech each provide *unique* insight into student proof generation. From an embodied cognition perspective, these findings underscore the idea that since gesture and speech convey some distinct information, both are important components of cognitive processing to consider when evaluating student performance on a task, such as valid proof construction. In other words, although a student may not be able to specifically articulate a valid mathematical proof, his or her gestures can shed light on potential key mathematical insights that he or she possesses but is not yet able to verbalize. However, as with any study, our conclusions are limited by our measurement techniques; there may be other components of language that are important to proof construction that our tools were not able to capture, or other ways of classifying gesture that would allow for even more explanatory power. In addition, examining these patterns across more mathematical tasks would certainly be useful. Nonetheless, these findings suggest that attending to student gestures could assist teachers in diagnosing student misunderstandings when formulating mathematical proofs.

In addition, given that learners typically struggle to grasp the discursive, logical structure of explicit proof statements even with continued instruction, interventions seeking to change students' gestures may offer a feasible alternative channel for learning. Teachers could model and encourage students to use dynamic gestures in the classroom, while concurrently accentuating the importance of language-based aspects of proof. Therefore, understanding how gesture can uniquely explain proof validity can inform potential educational interventions.

Overall, this work indicates that gestures have explanatory power about students' mathematical knowledge and skills above and beyond that available from speech alone. Understanding the complex interplay between gesture and speech as reasoning unfolds is critical to supporting students in engaging in mathematics learning. By creating environments in which students engage in justification practices using both gesture and speech, we can help students to "learn and become" as they take part in mathematical practices.

References

- Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology*, 25, 468-523.
- Alibali, M. W., & Kita, S. (2010). Gesture highlights perceptually present information for speakers. *Gesture*, 10(1), 3-28.
- Alibali, M., & Nathan, M. (2012). Embodiment in mathematics teaching and learning: Evidence from students' and teachers' gestures. *Journal of the Learning Sciences*, 21(2), 247-286.
- Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Church, R. B., Johnson, C. V., Jacobs, S. A., & Knuth, E. J. (2014). How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction*, 32(1), 65 – 100.
- Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617-645.
- Bates, D., & Maechler, M. (2010). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-35. <http://CRAN.R-project.org/package=lme4>
- Boncoddio, R., Williams, C., Pier, E., Walkington, C., Alibali, M., Nathan, M., & Dogan, F. (2013). Gesture as a window to justification and proof. In Martinez, M., & Castro Superfine, A. (Eds.), *Proceedings of the 35th annual meeting of PME-NA* (pp. 229-236). Chicago, IL: University of Illinois at Chicago.
- Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics*, 38(1-3), 85-109.

- Gee, J. P. (2007). *Social linguistics and literacies: Ideology in discourses*. New York: Routledge.
- Göksun, T., Goldin-Meadow, S., Newcombe, N., & Shipley, T. (2013). Individual differences in mental rotation. *Cognitive Processing (Special Issue on Spatial Learning and Reasoning)*, 14, 153-162.
- Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, 20(3), 267-72. doi:10.1111/j.1467-9280.2009.02297.x
- Graesser, A., McNamara, D., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, (36), 193-202.
- Gresalfi, M. S., & Cobb, P. (2006). Cultivating students' discipline-specific dispositions as a critical goal for pedagogy and equity. *Pedagogies*, 1(1), 49-57.
- Harel, G., & Sowder, L. (1998). Students' proof schemes. *Issues in Mathematics Education*, 7, 234-283.
- 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.
- Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, 15(3), 495-514. doi:10.3758/PBR.15.3.495
- Jeon, M., & Azevedo, R. (2007). Analyzing human tutorial dialogues for cohesion and coherence during hypermedia learning of a complex science topic. In D.S. McNamara & J.G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 153-170). New York, NY: Routledge.
- Nathan, M. J. & Johnson, C. V. (2010, August). The relation of situation models to gesture production when learning from a scientific text. *Poster presented at the Twentieth Anniversary Meeting of the Society for Text & Discourse*, Chicago, IL.
- Nathan, M.J., Srisurichan, R., Walkington, C., Wolfgram, M., Williams, C., & Alibali, M. (2013). Building cohesion across representations. *Journal of Engineering Education*, 77-116.
- National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM. Retrieved from <http://www.nctm.org/standards/content.aspx?id=26792>.
- National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). *Common Core State Standards: Mathematics*. Washington D.C..
- Nemirovsky, R., & Ferrara, F. (2009). Mathematical imagination and embodied cognition. *Educational Studies in Mathematics*, 70(2), 159-174.
- Newcombe, N. S. & Shipley, T. F. (2012). Thinking about spatial thinking: New typology, new assessments. In J. S. Gero (Ed.), *Studying visual and spatial reasoning for design creativity*. New York, NY: Springer.
- Pennebaker, J. W., Booth, R. J., & Francis, M.E. (2007). *Linguistic Inquiry and Word Count: LIWC [Computer software]*. Austin, TX: LIWC.net.
- Pier, E. L., Walkington, C. A., Clinton, V., Boncoddio, R., & Nathan, M. J. (2014). Exploring computerized text analysis to predict the validity of students' proof construction. *Paper presented at the 2014 American Educational Research Association Annual Meeting and Exhibition*. Philadelphia, PA.
- Radford, L. (2003). Gestures, speech, and the sprouting of signs: A semiotic-cultural approach to students' types of generalization. *Mathematical thinking and learning*, 5(1), 37-70.
- Schoenfeld, A. (1994). What do we know about mathematics curricula? *Journal of Mathematical Behavior*, 13(1), 55-80.
- Shapiro, L. (2011). *Embodied Cognition*. New York: Routledge.
- Walkington, C., Nathan, M., Alibali, M., Pier, E.L., Boncoddio, R., & Williams, C.C. (2013). *Projection as a mechanism for grounding mathematical justification in embodied action*. Paper presented at the 2013 American Educational Research Association Annual Meeting and Exhibition. San Francisco, CA.
- Williams, C., & D'Mello, S. (2010, January). Predicting student knowledge level from domain-independent function and content words. In *Intelligent Tutoring Systems* (pp. 62-71). Springer Berlin Heidelberg.
- Williams, C., Walkington, C., Boncoddio, R., Srisurichan, R., Pier, E., Nathan, M., & Alibali, M. (2012). Invisible proof: The role of gestures and action in proof. In L.R. Van Zoest, J.-J. Lo, & J.L. Kratky (Eds.), *Proceedings of the 34th annual meeting of PME-NA* (pp. 182-189). Kalamazoo, MI: WMU.
- Woods, D., & Fassnacht, C., (2012). Transana v2.52. <http://transana.org>. Madison, WI: The Board of Regents of The University of Wisconsin System.

Acknowledgements

Thank you to Virginia Clinton and Dan Klopp for their contributions. This work was made possible through grants from the National Science Foundation (DRL-0816406) and the Institute of Education Sciences (R305B100007).

“I Want to be a Game Designer or Scientist”: Connected Learning and Developing Identities with Urban, African-American Youth

June Ahn, Mega Subramaniam, Elizabeth Bonsignore, Anthony Pellicone, Amanda Waugh and Jason Yip
University of Maryland, College Park, 4105 Hornbake Building, South Wing, College Park, MD 20742
Email: juneahn@umd.edu, mmsubram@umd.edu, ebonsign@umd.edu, apellicone@gmail.com, awaugh@umd.edu, jyip@sesame.org

Abstract: Understanding identity, including how young people come to aspire to become someone, is vital to address the underrepresentation of minorities in science, technology, engineering, and mathematics (STEM). We report on a two-year, research project where we designed, implemented, and conducted case study research in an after-school program for inner city, middle school students. The program utilizes the school library, new media activities, and science fiction to engage young people to imagine STEM as relevant in their lives. We focus our analysis on two African-American boys, Damian and Jamal, who are best friends and avid gamers. Despite their similar backgrounds, they show starkly divergent identity trajectories while participating in our program. We highlight how they experienced different connected-learning activities and social positioning over time, and how these experiences related to Damian’s developing aspiration to become a game designer or scientist, contrasted with Jamal’s struggle to imagine a future in STEM.

Introduction

Increasing the representation of minority groups in science, technology, engineering, and mathematics (STEM) fields is a high profile issue in education (Chen & Weko, 2009). However, helping individuals develop an aspiration to pursue STEM related fields is a complex endeavor. Identity-development experiences that help individuals define themselves and encounter positive positioning by others may provide important mechanisms to develop such aspirations. Research has shed light on how young people develop identity in a variety of formal and informal learning contexts (Barton et al., 2013; Nasir & Hand, 2008; Polman & Miller, 2010; Polman, 2006). We build on these studies to examine how young adolescents develop identity in a world in which they have increasing access to new media, online networks, and information.

In this paper, we report on a longitudinal, design-based research (DBR) project (Collins, Joseph, & Bielaczyc, 2004) in which we (a) designed and implemented an after-school program called Sci-Dentity for urban, middle school students that utilized science storytelling and new media to engage youths in STEM ideas, and (b) examined the students’ developing sense of personal identity over an 18-month time period. We developed the project within the ethos of recent efforts to understand young people’s new media participation, interest-driven practices, and connected learning experiences (Ito et al., 2013; Pinkard & Austin, 2010). Students in our after-school program engage in various storytelling and new media production activities that examine STEM in popular culture. We iteratively designed the program to allow youths to pursue their own interests and we leveraged this context to examine their experiences and identity development over time.

We focus our analysis on two African-American, adolescent males, Damian and Jamal (pseudonyms). They are best friends who share a deep affinity toward video games and technology. They attend the same middle school and are among a select group of students who have participated in our program consistently for the last two years (as 6th and 7th graders). Despite their similar backgrounds, Damian begins to assert an aspiration to be a game designer or scientist by the end of his 7th grade year, while his friend Jamal continues to struggle to imagine his future. In this paper, we outline how theories of identity development, new media literacy, and connected learning help us to understand their participation in our program and their evolving aspirations. Our findings contribute to current discussions about learning, such as Connected Learning models that foreground the potential for young people to leverage technology to produce, participate, and connect with resources – informational and social – to pursue their interests and develop career goals (Ito et al., 2013).

Theoretical Framework

The driving motivation for our research is to shed light on the ways in which school libraries, storytelling activities, and new media participation can help young adolescents identify with STEM and develop aspirations to pursue a future in a STEM field. We build from prior scholars who have examined (a) learners’ identity development and (b) ideas about how individuals learn with new media and across informal environments.

Identity Development

We leverage Barton and colleagues’ (2013) definition of identity as how a person figures themselves, and is perceived by others, in a given moment. How a person *figures themselves* can be observed as a series of stories

or narratives that are told about the self, such as the communities one affiliates with, what actions one is capable of, and what futures one can envision (Sfard & Prusak, 2005). Thus, an initial point of understanding is observing how the youths in our program voice their aspirations and interests. Figured selves are also enacted in specific contexts of practice that influence what actions and discourse can occur and are valued (Polman, 2006). The concept of positioning describes how a person authors him or herself in relation to the people with whom they interact, and the activities and community in which they participate (Polman & Miller, 2010). A key element of understanding an individual is how they are *perceived, recognized, and socially positioned by others*.

Recognition by others has been shown to play a major role in how female students connect with science and how they decide to pursue science careers (Carlone & Johnson, 2007). Learners are influenced by how a teacher or adult facilitator views and subsequently treats them, and what others see as their capabilities (Barton et al., 2013; Polman, 2006). In prior research, such identity-developing experiences have been analyzed as *moments* or scenes (Polman, 2006) that can illuminate how local and seemingly ephemeral interactions leave imprints on how individuals perceive their opportunities, position in society, and aspirations toward future activities. In our research we seek to examine the moments of interaction that our participant youths experience, and leverage these moments to glean insight into their developing identities over time. Identity develops through the history of accrued moments that span across contexts and time (Barton et al., 2013). Thus, we also turned an analytical eye toward understanding how our focal learners' moments (or scenes) linked together over time to relate to their own articulated aspirations and self-image.

Connected Learning, New Media Literacy, and Identity

We examine our youths' developing sense of self in a Connected Learning context, in which young people have increased access to a wider ecology of information, technology, and interest-driven learning communities (Ito et al., 2013). For example, informal learning programs and online communities provide ways for young people to learn important skills, cultivate relationships, and develop their own identities in the process. Theoretically, these capabilities should provide more pathways for young people to develop deeper identification with a personal interest, develop expertise and skill, and connection to career and life goals. However, there is a need to understand how these pathways develop and how learners actually make these connections between interests, learning opportunities, and formal academic or career goals. Much of the early work has focused on understanding successful learners, but attention is also needed to articulate the challenges that arise for learners in these contexts.

The intersection of new media literacy and identity provides a framework to understand and articulate these ideas. Skills such as searching for information, assessing credibility of content, and understanding complex cultural practices such as remix, are vital to fully participate in connected, interest-driven learning experiences. Research is needed to better understand the consequences for learners who are still developing, or face obstacles in learning these literacy skills. There is also increased understanding that successful learners are able to effectively mobilize and activate their ecology of resources in order to pursue self-interested learning (Barron, 2006; Ito et al., 2013). For example, successful learners can search for information and connect to content in ways that further their knowledge. They can develop relationships with peers or mentors, in order to develop skills and guide their activities. However, there are open questions still concerning how to support struggling learners who might not yet have these skills to mobilize their ecosystem of resources. One potential link is understanding how literacy skills and identity moments intersect in the cycle of experiences that a learner may face. For example, in order for an individual to fully participate in an environment or learning community, they need to have some modicum of literacy skills. To gain skills, one needs incremental and sustained participation. However, to sustain participation, one also needs to self-identify with the community in the first place. How do these inter-related factors combine in cycles of experiences for learners, and subsequently relate to their developing sense of future aspirations?

Research Questions

To delve deeper into this broader question, we focused on three research questions that were informed by the identity development and connected learning literature:

- R1a: How did our youth participants position themselves, and how did others position them, throughout their experiences in Sci-Identity?
- R1b: How did these experiences relate to each other, and inform future interactions over time?
- R1c: How did youth participation in the interest-driven, new media based activities in Sci-Identity relate to their identity trajectories?

Methodology

Context of the Study and Design of Sci-Identity

The context of this study is an after-school program called *Sci-Identity* that is run in middle school libraries, in partnership with school librarians, and focuses on using science-based stories to help adolescents think about the relevance of science in their everyday lives (Subramaniam et al., 2013). The students read young adult books, or watch videos that incorporate science in stories. In addition, they engage in storytelling activities using traditional (e.g. writing) and new media (e.g. videos, infographics etc.) modes of production. Participants in the program also network and share their work in a private social network site (<http://sci-identity.org/>).

We implemented Sci-Identity in a large, urban school district where approximately 90% of students come from minority groups and nearly 77% qualify for free and reduced meal (FARM) programs. Damian and Jamal, our two focal learners, attend a school where nearly 100% of students qualify for FARM and 98% of the student population is African-American. Over an 18-month period, the research team planned and facilitated weekly after-school sessions at two urban middle school libraries. At various iterations of the project, the project team included seven graduate student assistants and ten masters student volunteers that helped with the program. This team attended the after-school sessions, worked individually with the students, and participated in data collection and weekly team debriefing meetings where we discussed what occurred each week, emerging issues, and redesigned activities based on our findings. The project was designed iteratively, with periodic redesigns of the general after-school program based on research and analysis.

Data Collection

We collected several forms of data that spanned a period from January 2012 to June 2013. Each week, sessions at each school were video recorded. All members of the research team, including student assistants and volunteers, wrote weekly observation notes; took pictures, video, and audio recordings; and logged individual interactions with the youth participants. Our weekly team meetings were also audio recorded and included in our corpus of data. All artifacts and projects created by the students (e.g. stories, sketches) were collected, as well as log data of their activity on the social network site (e.g. record of logins, posts, comments etc.). Finally, students and librarians in our program were interviewed at the end of each school year, May 2012 and May 2013, which coincided with the end of 6th and 7th grade. In total, our data sources included nearly 60 hours of video and audio recordings, approximately 250 pages of written observation notes, and over 50 interviews.

Data Analysis

We analyzed our corpus of data using a case study framework (Yin, 2003). Each member of the research team chose several focal learners in our program. For each focal learner, the researcher reviewed each week of data over the 18-month period, and chronologically catalogued: (1) specific learner interactions that were captured in video and audio recordings, with salient interactions transcribed for dialogue and actions, (2) mentions of the learner in written observation notes by the project team, (3) artifacts such as posts to the social network site and their project work, and (4) identity-related information culled from our interview transcripts.

Our chronological data trail allowed us to both triangulate data sources to develop our claims and check the trustworthiness of our interpretations in a systematic way. For example, we could match scenes on video recordings with observation notes of the same interaction made by the project team, and notes from team debriefing sessions that mentioned that scene. We could combine multiple perspectives of the same scene – from video, from the perspective of an actor in the scene, or group understanding of that scene in team meetings. This constant comparison process enabled us to examine how our focal learners positioned themselves in different interactions, how others perceived them, and how they experienced our program activities. Interview data from the focal learners provided insight into their aspirations, perceptions of self, their participation in the program, and their home life. Interviews with librarians provided an additional window into how our focal learners were perceived in school. Each researcher coded data for instances of participation, positioning, and literacy. Each researcher also annotated the data for emerging themes, particularly if they were recurring and related longitudinally across time. Thick case descriptions for each learner's identity trajectory were crafted through this process, and enabled us to make comparisons across learners as well.

Findings

Many of Damian and Jamal's moments in the program occurred together or in close vicinity to one another. Likewise, we present their cases through intertwined vignettes that are representative of their closely linked experiences, while illuminating the contrasts in their identity trajectories over time.

New Beginnings, New Friendships

We began our after-school program with the students in January 2012. In the first session, students watched a video about storm chasers (scientists who follow and study tornados), and did a short story exercise based on the

video. Immediately, we observed how Damian and Jamal positioned themselves and how our own project team positioned them. Several facilitators immediately recognized Damian as a young man seeking to make connections and being proactive about charting his own path. One facilitator noted, “Damian is a researcher in the group. Our first week with him, I noticed that he was searching a lot of information concerning Storm Chasers” and another added, “Damian was extremely proactive, even searching other websites for information to enhance his story”. These initial observations – a willingness to engage in activities, search for information online, and move off on his own based on the task given – became consistent ways that Damian positioned himself throughout the program.

In contrast, Jamal had a rather inauspicious beginning. One facilitator wrote in her observation notes, “Jamal had his head on the desk. Not sure why he was even there. He did not seem the least bit interested in the story writing. I tried working with him individually, but it was like pulling teeth. His story was uninspired, short and full of spelling errors. I’ll be surprised if we see him again.” Other project volunteers confirmed this observation, noting that the other students treated his disengaged behavior as if it was the norm. This characterized a recurring theme for Jamal, one in which he positioned himself as outwardly disengaged, while adults responded by positioning him as apathetic to the program and showing low literacy skills. Despite this initial perception, Jamal was one of the few participants who attended the program fairly consistently, for the entire two-year period of this study. What explains this contradiction?

Damian and Jamal began a close friendship during this time period that revealed aspects of their motivations and personalities. Damian was a new student to the middle school and had few existing relationships with peers. Perhaps it was this situation that related to our consistent observations of Damian seeking social connections with both peers and adult facilitators over time. In our second week, video observations showed Damian developing a friendship with Jamal. Jamal confirmed these observations in our interviews with him, and he explained how he became friends with Damian, “He came here in the sixth grade, and I just started talking about games. That’s how we became friends.” Jamal’s self identity as kind and helpful also relates to this developing friendship. In our interviews with students, we asked them to imagine writing a biography and explain how they would describe themselves. Jamal provided insight into how he saw himself, “I don’t really know. I’m kind of fun sometimes. I’m always there to mostly help. Like when a new student’s here and you know, he gets bullied and stuff, I’m always there to help. I’m easy to be friends with. I like to play games and fix stuff.” The school librarian corroborated these observations of Jamal saying, “So, in his class, he’s very well-behaved. Like, he’s one of the better students and kind of just has to deal with a bunch of ruckus going on around him all the time, because he’s very respectful and he’s not going to give anybody a hard time. [But] I think his problems come with engaging in the work.”

Utopian and Dystopian Futures

The students’ primary 6th grade project was to develop a short story based on a discussion of utopian and dystopian future societies. The students read and discussed *The Hunger Games* (a popular young adult series that had a major movie release at the time). A requirement of the project was for students to search for and find an interesting scientific idea, around which they could craft a plot and characters. Damian and Jamal, like many of the other learners, struggled when using the library iPads to search for information related to their interests. In one video scene, we observed Jamal struggling to search for information and remarking, “everything is blocked” through the school district’s Internet filter. A facilitator asked Jamal what he was interested in and he replied “weapons”. This interaction led to a series of guided Internet searches about weapons, ballistics, and somehow to information about the game World of Warcraft. The two boys became very excited at this finding, leading to Damian jumping into the exchange:

- Facilitator: ... Perhaps you can do a search about what it takes to make video games and use that in your story... have you ever seen Tron?
- Damian: I love Tron! That’s a video game.
- Facilitator: Yea it is a video game... Someone actually made a story about the game Tron.
- Damian: Like a virtual world.
- Facilitator: Yea, so I would find a fact about virtual worlds and write a story about that.”
- Jamal: I want shooting games.

In this scene we noticed how the facilitator was initially helping Jamal through his difficulty in conceptualizing what to search for and how to conduct a search for information. However, Damian jumped into this activity and subtly took advantage of an opportunity to connect with this facilitator. We also observed how the adult facilitator positioned the boys’ interests in video games as a valid avenue to engage in the task. Damian made a connection, observing that the story of Tron was centered on the ramifications of designing and living in a virtual world. We later learned that Damian was an avid Minecraft player, a game where players can build a virtual world themselves. Damian realized that his interest in virtual worlds was a valid way by which he could

connect his interest in games with the exercise of imagining futures. In contrast, Jamal exhibited behavior that became a recurring pattern for him, where he was a silent participant who seemed passive but secretly listened in on the conversation and tried to make his own connections. Jamal chimed in at the end and voiced his interest in researching shooting games, revealing that he had been following the conversation silently. Both Damian and Jamal continued to follow this interaction pattern in future sessions.

In later weeks, we observed Damian searching for information on Minecraft. Video observations show Damian explaining to a facilitator that at home he was making a “cobblestone generator that mixes lava and water and makes cobblestones” in the game. Ultimately, Damian developed a story in which life was entirely lived in a virtual world, and incorporated his view of the societal ramifications of this future. His path to this project illuminated his propensity to connect ideas. His initial interest was in video games and Minecraft, but he was able to readily connect his out-of-school interest to the after-school activity of combining science (programming, virtual worlds) with stories (Tron) and apply it to his own imagination.

In contrast, Jamal faced obstacles that related to his self-doubts about his formal literacy skills. As a result, he consistently positioned himself in public as someone who was disengaged with formal academic work while privately attempting to engage behind the scenes. In one session the students worked in groups to think about utopian and dystopian futures, and how science and technology would play a role in those imagined futures. Jamal was a leader during group discussions, often contributing ideas such as zombies that the other students took up enthusiastically. However, a facilitator wrote in her observation notes that Jamal did not seem interested in writing his own story, “When he starts writing, his stories and ideas are always really interesting, but I think he is hesitant in people reading the materials or criticizing his work.” Video recordings show Jamal in this session, with his head down not knowing how to start. He also steals glances and listens in on his peers talking around him, searching for clues as to how to make progress, but does not readily accept help from adult facilitators. This pattern of interaction occurred regularly. Jamal would often engage when thinking about ideas or discussing topics of interest to him, but when asked to create formal products he often reverted quickly back to a posture of public disengagement.

Connections and Disconnections

In the 7th grade, we guided the students to consider identity by imagining themselves as superhero characters in science fiction. The learners created infographics about their characters, linking science and life events to the stories. Damian particularly struggled to focus in this activity. He spent much of his time using an online avatar creator to visually depict his superhero character, and joked around with Jamal who was often not engaged with the project. However, he also spent time trying to connect with facilitators that led him to new discoveries.

Damian’s superhero character’s name was Wesker, who had a magic helmet and hammer, and came from another planet. We learned later in the year that Damian was interested in the movie Thor, which was about a comic book character that had a similar storyline. Eventually, Damian focused on exploring potential planets that Wesker could have come from, and in one session, he suddenly exclaimed out loud “Hey I found a new planet like Earth!!!!” Damian proudly posted a news article to a section of our social network site called the “Brain”, where we encouraged learners to share science inspirations that could be used in stories. The article described NASA’s Kepler space telescope project that set out to discover planets that resided in “habitable zones” similar to Earth. This breakthrough energized Damian and in future sessions, we consistently observed him engaged in learning more about planets discovered through the Kepler satellite. In one session we saw Damian, working with a facilitator, leveraging his new discovery to engage with scientific ideas. Damian was excited, exclaiming, “Awww!! There’s more planets! ...I found another planet!!” He began to describe to the facilitator the facts about the planet, such as the number of candidate planets, the search for Earth sized worlds, and how satellites detect new planets by the dimming of star lights as a planet crosses a star.

While Damian found connections with facilitators and ideas, Jamal struggled. He often sat away from others during whole group activities or discussions, but still within the vicinity of his friend Damian and another female friend Chanda, who was close to the two boys. Often Jamal would not engage with tasks unless Damian or Chanda urged him on. For example, Jamal was not particularly interested in creating an infographic for his character, Dark Batman, who he envisioned as a more evil version of the comic book character. However, he readily engaged in describing his character when we encouraged Damian to “interview” Jamal using their iPad (instead of writing in a document). These patterns related to Jamal’s self-perceptions about his academic ability (not writing, but participating in other new media forms of storytelling). We also came to understand that his main motivation was to have a place to hang out with friends.

We learned through interviews with the school librarian, that Jamal was in the “lower academic track” in his school. His friends, including Damian, were all in the high academic track. In Jamal’s interview at the end of 6th grade, he noted that he had very little time to hang out with Damian and their friends during school time. They had different classes, different lunch periods, and Sci-Identity was one of the times that they were able to hang out together. These details of Jamal’s school life suggested that the opportunity to be with friends, whom he is otherwise disconnected from, was a main motivation that drove him to continually attend the program.

Much of his 7th grade data described Jamal's friendship, joking around, and social drama with Damian and Chanda. The school librarian also confirmed this observation saying, "He wants to be around his friends, which are Damian and Chanda, but he doesn't to me really feel like doing anything. He just kind of wants to be there and not be missing anything."

The baggage of Jamal's perceived lower-academic positioning also emerged in other instances. For example, in late Spring 2013, we had one session where nearly all of the students did not attend. We learned that the students had to practice for the school play, but only those in the high academic track were in the play. This situation left Jamal as one of the only youth in Sci-Identity that afternoon and served as a reminder of his academic status in school and his disconnect from his friends. We noticed a consistent pattern of interaction and positioning with Jamal that harkened back to a previous comment by the school librarian that he was respectful, but just had to deal with the "ruckus" of his school environment. We observed that Jamal would often shut down at predictable moments. When his peers would get rowdy and the school librarian would shout at the students, he often just put his head down and tuned out. When the librarian or a facilitator would attempt to push Jamal to continue writing or searching for ideas online, he would disengage and respond with reasons for his lack of progress (often saying that he did not know how to spell). He also shied away from asking adults for help, instead preferring to listen-in on other students' conversations, to glean clues about what he should do. These series of experiences were related to Jamal often positioning himself at the outset of a session as cool and disengaged. In moments where we observed him striving to engage, his momentum was often halted by a myriad of factors: distractions from the sometimes hectic atmosphere of the school, lack of confidence in his literacy skills, and reticence to approach adults or peers for help (e.g. mobilizing his ecosystem).

Finding a Spark: Video Games

At the end of 7th grade, the students defined their own interests and the after-school sessions were focused on guiding interest-driven learning. Jamal and Damian (along with several boys) decided to learn how to design video games. We leveraged this interest to focus on game design tasks and talking about computing related careers. This project seemed to ignite Jamal. The school librarian also noticed a change in him and began to position Jamal differently: "I feel like at the beginning of the year we couldn't even get him to sit and listen or pay any attention to what we were supposed to be doing. And by the end of the year, he started to come up with the video game. Once we started that video game project, he really came up with his own ideas. So, he definitely improved in that aspect." Towards the end of 7th grade, we began to see a spark of interest with Jamal despite the fact that he continued to struggle with his self-positioning and literacy skills.

Damian also latched onto this project, having been one of the loudest voices to advocate a focus on video games. We introduced a tool called *GameMaker* that allowed the learners to design and develop simple games by loading sprites and coding simple scripts. Damian was immediately interested in learning how to use GameMaker and it was clear that the concept of actually building a video game himself ignited a passion in him. We observed the two friends deeply engaged in interest-driven learning activities and they spent much time deciding on the details of their game design. However, it was interesting to observe how members of our own team positioned these activities at times. One facilitator wrote in her notes, "Damian was working with Jamal on their game, but in fact spent the entire time looking up music to script to their game, which has not really been started. On the one hand, it is great to see them engaged in something, on the other, there was not a lot of product at the end. That said, they were very thoughtful about the kind of music they wanted." While the boys were very engaged in the project, there were some tensions with facilitators viewing activities such as looking up music for the game as non-productive.

Ultimately, the school year ended before the students could fully develop their video games. The experience illuminated some of the tensions related to interest-driven learning environments. We touched on a passion for these two friends (games) that spurred engagement in other interest-driven learning practices such as using technology to research information (e.g. find music to fit their game narrative), design products, and learn about tools to produce artifacts (e.g. GameMaker). In contrast, this interest-driven process also required much more time and personal guidance from facilitators. The boys (and other learners doing this project) were excited about their game designs, but were somewhat disappointed when the school year ended before they could become fluent enough with GameMaker to realize their game designs as actual games. A spark – an interest – was a vital first step, but formal practices such as design and coding also need time to develop and learn. The boys were disappointed that summer vacation would bring a lull in the program. However, they both cited this project as the most memorable and interesting of their two-year experience in Sci-Identity and voiced a strong desire to continue the project in their 8th grade year.

Discussion

This study contributes additional understanding of Connected Learning contexts and the role of identity development in such learning environments (Ito et al., 2013). Past research has identified factors that contribute to a successful connected learning experience. For example, successful learners begin with an interest and utilize

new media to work on active projects that promote further learning within those interests. They participate in learning programs and online communities to glean support from peers and mentors. And, finally they find ways to connect their learning in these experiences to formal endeavors such as schoolwork or progressing toward a career aspiration. Much of the early research describes success cases and focuses on the importance of configuring access to an ecosystem of resources – technology, people, institutions etc. – to enable this type of technology-mediated, interest-driven process.

This study lends additional detail and considerations to this framework. The cases of Damian and Jamal highlight the local processes – identity and literacy experiences – within a connected learning environment that relate to their successes and challenges over time. Damian’s story shows a young man who was well served by our program. He was in the high academic track in school and was already deeply engaged in a well-defined interest (Minecraft). Damian brought this background to Sci-Identity, and he accumulated positive experiences over time that further reinforced his interests and sense of self. Despite struggles he might have had with new media literacy skills or social distractions, he found ways to continually relate his experiences over time. Like more successful new media literacy learners, Damian developed an aptitude to productively mobilize his ecology of available resources. He sought help from facilitators and connected to peers, bridged his ideas across informal and formal contexts, and used technology to aid in these tasks. From an identity perspective, Damian positioned himself to achieve these goals and others positioned him in positive ways. We argue that this cycle of mutually reinforcing experiences related to Damian beginning to develop aspirations toward STEM. In his 7th grade interview, Damian voiced aspirations to become a “game designer or scientist”. He also shared details that spotlight his deepening identity trajectory. He disclosed that he knew of particular game design companies in his home state that he aspired to work for and had also done independent, online research to identify colleges that had game design degree programs.

Jamal’s case highlights a young man who could fall between the cracks, even in interest-driven programs that value alternative pathways to learning. Jamal brought his background to Sci-Identity, as a student that was insecure about his academic positioning in school. These factors related to his self-positioning in the program. Outwardly, he did not ask for help, but instead took to silently listening in and observing his peers, in efforts to get clues about how to engage in tasks. When able to tackle tasks that he was literate in – recording videos, designing a game – he showed progress. When we asked him to work in formal ways – doing research or writing – he shut down. We observed a learner who identified with our program because it gave him a connection to his friends. Jamal’s story is of a well-intentioned learner, who progresses in fits and starts, due to a cycle of obstacles that accrue over time. These experiences related to Jamal’s struggle to connect his learning across contexts and imagine his future. When asked about his future aspirations he replied, “Hoping to mostly fix cars or to customize cars or become I don't really know. That's mostly what I think I can do.” Interestingly, Jamal had never voiced an interest in cars for the 18-months he had been in the program, although it could have been a valid avenue to engage with STEM ideas. Jamal also had few plans for his educational future, stating that he should get a good education but probably would not go to college. When probed further, we learned that his grandfather and his socioeconomic success were the salient identity stories in Jamal’s mind. Jamal stated that his grandfather was the person he most looked up to, “... he's successful. He's not like that big time rich person, but he is successful. He has his own house. He has three cars that he was able to buy. He has a big TV. He was able to do all that stuff and retire too.”

Design Considerations and Conclusions

The results of this study have several implications for practice and research. In the context of our DBR project, the case studies introduce several considerations that we aim to explore in the next iteration (the students’ 8th grade year). A major challenge for interest-driven learning environments is designing experiences to continually promote each individual learner’s unique development of interest, expertise, and skill at every given moment in time (Edelson & Joseph, 2001). For Damian, we need to design activities that push him to further his expertise in formal technical skills, interact with individuals who can further his imagination of possible futures in game design or computing, and link his future educational plans to his developing interest. Damian needs further positive experiences that continue to add onto his positive identity trajectory and future aspirations. For learners like Jamal, we are faced with a complex task of how to break through his cycle of obstacles. We must find ways for him to engage in literate ways that do not lead to shut down, but also guide him towards more confidence in formally recognized practices and dispositions. Furthermore, there are other learners in Sci-Identity who each have their own unique interests to address. For example, some learners are avid storytellers, and other learners in the program, like Jamal, are still searching for interests to pursue and facing their own unique obstacles.

In our next DBR iteration, we see intriguing opportunities to further integrate ideas of interest-driven learning, within the unique affordances of Connected Learning environments, while serving the specific goal of helping under-represented youths develop stronger identities in STEM. For example, Edelson & Joseph’s (2001) Interest-Driven Learning framework provides insights into the types of contextual motivation strategies that could be employed in our weekly sessions. The Connected Learning literature underscores the importance of

putting into place a larger ecosystem of resources and experiences – beyond just our after school setting – that could help our learners to begin explicitly linking their experiences across contexts and expand their imagination. Finally, research in learners’ identity-development attunes us to examine the cycle of experiences that our youths experience, connected to their broader social, economic, and cultural settings, that relate to their deepening sense of self over time. Integrating these separate research streams may help us better understand complex challenges such as how to promote a learner from an interest to a deep commitment to pursue a STEM career. As one example, DiSalvo and Bruckman (2009) found that there is a correlation between video game interests and a student’s decision to pursue computer science in college, but the relationship is very small and the majority of avid gamers never translate their interest in games to actually pursue computer science. Our findings offer a potential response to this perplexing relationship. The careful design and accumulation of interest-driven activities, enhanced through connected learning contexts, that promote identity-conscious experiences, may be required in concert to reduce the gap between a learner’s *figured self* with their *potentially realizable self*. We posit that the cases of Damian and Jamal are applicable to other learners. By being attuned to the contextual factors at play, at key moments during the identity development of learners like Jamal and Damian, we can perceive and design for those moments that present opportunities to spark more productive trajectories and to sustain those already rocketing forward.

References

- Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development*, 49, 193 – 224.
- Barton, A. C., Kang, H., Tan, E., O’Neill, T. B., Bautista-Guerra, J., & Brecklin, C. (2013). Crafting a future in science: Tracing middle school girls’ identity work over time and space. *American Educational Research Journal*, 50(1), 37–75.
- Carlone, H. B., & 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.
- Chen, X., & Weko, T. (2009). *Students Who Study Science, Technology, Engineering, and Mathematics (STEM) in Postsecondary Education* (No. 2009161). U.S. Department of Education, National Center for Education Statistics. Retrieved from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2009161>
- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, 13(1), 15–42.
- DiSalvo, B. J., & Bruckman, A. (2009). Questioning video games’ influence on CS interest. In *Proceedings of the 4th International Conference on Foundations of Digital Games* (pp. 272–278). Retrieved from <http://dl.acm.org/citation.cfm?id=1536561>
- Edelson, D. C., & Joseph, D. M. (2001). Motivating active learning: A design framework for interest driven learning. Retrieved from <http://www.designbasedresearch.org/reppubs/edelson-joseph.pdf>
- Ito, M., Gutierrez, K., Livingstone, S., Penuel, W. R., Rhodes, J., Salen, K., ... Watkins, S. C. (2013). *Connected learning: An agenda for research and design*. Irvine, CA: Digital Media and Learning Research Hub.
- Nasir, N. S., & Hand, V. (2008). From the court to the classroom: Opportunities for engagement, learning, and identity in basketball and classroom mathematics. *The Journal of the Learning Sciences*, 17(2), 143–179.
- Pinkard, N., & Austin, K. (2010). Digital youth network: Creating new media citizens through the affinity learning model. *International Journal of Learning and Media*, 2(4).
- Polman, J. L. (2006). Mastery and appropriation as means to understand the interplay of history learning and identity trajectories. *The Journal of the Learning Sciences*, 15(2), 221–259.
- 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.
- Sfard, A., & Prusak, A. (2005). Telling identities: In search of an analytic tool for investigating learning as a culturally shaped activity. *Educational Researcher*, 34(4), 14 – 22.
- Subramaniam, M., Ahn, J., Waugh, A., Taylor, N. G., Druin, A., Fleischmann, K. R., & Walsh, G. (2013). The role of school librarians in enhancing science learning. *Journal of Librarianship and Information Science*, 0961000613493920.
- Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Acknowledgments

We would like to thank all of the school librarians, our students, and graduate research assistants who played an integral role in this project. This material is based upon work supported by the National Science Foundation under Grant No. 1124176. 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.

Reflecting on Educational Game Design Principles via Empirical Methods

Oswaldo Jiménez, University of the Pacific, 3601 Pacific Ave, Stockton CA, ojimenez@pacific.edu

Abstract: In designing environments for children to use for learning, there are many design decisions that are made by the game's creators that can affect their effectiveness. One such example that many creators determine is the amount of story to embed in an educational game. To learn more about the importance of story in educational games, 77 fourth grade students in one elementary school were randomly assigned to play one of three versions of an educational video game with varying levels of story. The goal of the videogame is to give students practice with fractional-whole operations. In addition to logging interactions students made during gameplay, pre and post-tests that capture students' fractional knowledge in a classroom environment are reported and discussed. Results indicate that while story may not seem to be a critical factor in improving learning, its benefits and impact on learning may be more nuanced and complex.

Introduction

With one of the conference aims being about learning and becoming in practice, and foregrounding ways that learning processes are situated in practice, the research reported herein contributes to better understandings and practices involved in children's engagement with educational games. There have been many commercial and educational video games that have been studied or created by scholars. These scholars have shown that the games have diverse benefits, such as increased self-efficacy (Nelson, Ketelhut, Clarke, Bowman, & Dede, 2005), authentic scientific inquiry, (Barab, Sadler, Heiselt, Hickey, & Zuiker, 2007), and learning to use evidence to warrant claims (Steinkuehler, 2006). However, the creation of educational video games, merely in viewing it as engineering a complex software program, is both an expensive (Squire, 2003) and difficult endeavor to undertake (Naone, 2007). While some games have shown learning differences, (Barab et al., 2007, Jiménez, Arena, & Acholonu, 2011), there are relatively few studies demonstrating the learning benefits of educational games using traditional measures (Honey & Hilton, 2011).

Because educational video games are so difficult to make, researchers and proponents of making games ponder how to create more effective educational video games. One way is to better understand the impact that certain aspects or features of educational games have in helping students learn. One area where variable amounts of resources are often placed is in a game's story and story structure. Stories have many links to learning that may be important for games, ranging from promoting role-playing (Williamson & Slivern 1991), to influencing people's judgments (Strange & Leung, 1999), to impacting individuals' comprehension and memory (Mandler & Johnson; 1977; Thorndyke, 1977).

Studies have demonstrated a potential link between stories and learning as well as links between stories and educational games. Researchers have demonstrated links between having a story in an educational video game and children's motivation to work on educational content (Malone, 1981; Parker & Lepper, 1992; Cordova & Lepper, 1996, Barab, 2007). Others have argued for a more refined link between stories and video games, arguing against some of this same research, which suggests that developers be very careful about integrating such stories/fantasy contexts with a game (Habgood, Ainsworth, & Benford, 2005). Scholars have further argued that context is largely irrelevant stating that stories chosen for their own games could easily be replaced with something else (Habgood & Ainsworth, 2011, p.176). If this is the case, story could be absent from a game altogether or could be replaced with something unappealing. Determining whether a story is important to an educational game informs game designers' future development and can help create better educational experiences for students. To further investigate this factor, one approach is to research an educational game that can be modified to have different amounts of story in the game, and use results to make decisions on how to change a game to have an efficacious amount of story.

Does Story Matter in Educational Games?

This paper reports on this type of inquiry. A computerized fractions card game that demonstrated increases between pre and post-test classroom measures (Jiménez, Arena, & Acholonu, 2013) was borrowed and used to create three versions. The three versions leveraged the game and story used in the aforementioned research. The game was appropriate for this kind of study because it dealt with a difficult learning domain, fractions. Fractions is one of the most difficult concepts for children in elementary education (Petit et al., 2010) and also one of the first concepts that causes children to stray away from a STEM pathway (Nunes, 2006).

Past research was leveraged to select and vary the critical components in each experimental condition. For example, prior studies on the use of characters and their impact on story recall and comprehension (Bower,

1978; Thorndyke, 1977) led the authors of this study to select characters as a feature that could be varied across conditions. Characters such as protagonists are a critical component of a story and a feature that could be modified to change the amount of story. Therefore one version, the *Characters version*, places more prominence on the characters in the game. Another version, the *Abstract version*, is completely devoid of any story structure, leveraging story grammars to take out all critical components, including the use of concrete characters. The final version of the game, which I refer to as the *Original version*, mirrors closest to what students played in the original fractions game cited. Nonetheless, all three versions share the same underlying code base. Only the text and images used in the software differ between them.

In the Characters version of the game, care was taken to promote and embellish the story by making the characters more salient for children. That is, rather than having generic characters as the manipulatives, the students worked with the characters present on the cards throughout the game. Figures 1 and 2 detail the differences present in the three conditions across the most important game screens.



Figure 1. The three conditions on the card playing screen: the Characters condition (left), the Original condition (middle) and the Abstract game condition (right).

This Characters version is displayed on the left of both Figure 1 and Figure 2. The only crucial difference is that the manipulatives in this condition were made so that they would look exactly like characters that they represented in the cards. For example, rather than seeing eight general figurines which are used to represent everyone, each figurine shown in the game would correspond to one of the members of the Johnson Family shown on the card. These figurines would also be shown during the calculation and would also be shown as *stinky* or *fresh*. Now that each character is salient, when players select characters to become “stinky”, those specific characters would also be stinky on the other screen. The other conditions also give the students the opportunity to select which manipulative that they want to become stinky; however, one would think that choice would become more trivial since all of the manipulatives look alike.

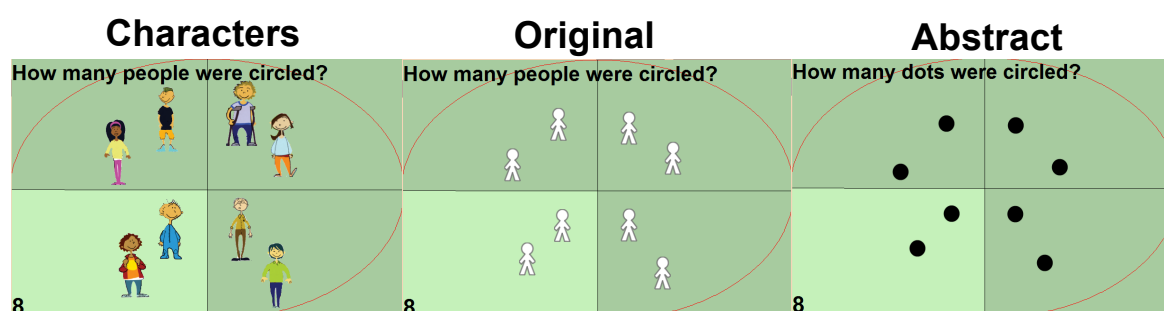


Figure 2. The three conditions after having provided help on answering $\frac{3}{4}$ of four. The Characters condition (left), the Original condition (middle), and the Abstract game condition (right)

In the Abstract condition, which is shown on the right of Figures 1 and 2, one notices that the game has been completely stripped of the story structure. Therefore, all of the *Tug-of-War* story images were replaced with abstract images. When replacing the images and text, careful consideration was given so that the software would not be written any differently in the Abstract condition. Only the images and text displayed to the user were changed. The people changed to dots, both in images and language; cards like the stink bomb were changed to be called “Reduce” cards. All of the new animations were still kept in the game, except different images were used to explain what happened. All of the story elements were replaced by abstract representations so as to make the game playable without any insinuation of story. Taking the literature on story into consideration, one might hypothesize that the story conditions would be able to capitalize both on helping the user learn as well as engaging more with the material. However, it could be that traditional game mechanics are powerful enough for students to drive the rest of the learning and memorization in the game.

There are researchers who propose that the abstract is better, citing articles that argue against using idealized images and for using lines and shapes for student work (Goldstone & Wilensky, 2008), in learning applications such as agent-based modeling (Wilensky, 2002). There are also arguments against using a version of the game where the characters are more salient. The argument against characters may be that the characters become so salient that they distract the players from the content itself (Son & Goldstone, 2009). These complex nuances play out in games and need to be further discussed and researched. The goal of the experiment described below was to understand and detail the roles that story plays in games and how they impact and interplay with games. In order to explore this impact, a protocol was employed that assigned children to play one of the three different versions of the game.

The Study

The study was conducted once the versions of the game were developed. Students were exposed to one of three versions of the game. Afterwards, they were compared on how much they had learned and how much fun they had. The method and results of that study are presented below.

Participants

77 fourth grade students participated in this study. All of these students attended one elementary school in the San Francisco Bay Area. Over 90% of the students in the study were classified as English language learners. These students were a part of classrooms in the same school, although one of the classrooms was a mixed fourth- and fifth-grade classroom. On a previous standardized benchmark test that the students took, only 27% of students tested well enough to be classified as having mastered the mathematics content taught so far at their school, with 20% struggling to learn the material. The remaining 53% were assessed as making progress towards mastery.

Materials

Both the pre- and new post-tests were an expanded version of the tests used in earlier studies of the fractions card game. This test expanded on questions that relied more heavily on students' understanding of part-whole relationships (such as "what is $\frac{1}{2}$ of 8" and less on their understanding of chance, which is known to be a very difficult subject to learn (Garfield & Ahlgren, 1988). The questions that were devised in this test belonged in one of five categories. These categories were questions where students had to compare fractions to one another, questions involving decimals, complex word problems, and fractions questions. The fractions questions were further broken down into two categories: fractions they had most likely seen in the game and fractions they could not have encountered in the game. The maximum score that could be obtained on the test was 41. Most questions were given one point; however, large problems were divided into sub-problems, and each of those sub-problems was given a point.

As mentioned previously, the post-test added a fun survey as well as questions about the game to the end of the measurement. I also added additional questions about the story and playing the game. Due to student absenteeism, 73 out of 77 students completed both the pre and post-tests.

Method

Rather than assigning each classroom to a game condition, all 77 students were randomly assigned to a condition. This random assignment was stratified by previous performance on a standardized test taken earlier in the year, and controlled for their performance so that no condition had a significant difference in standardized test scores. To accommodate wishes expressed by parents about their children not being videotaped, some of the students were switched out of conditions with other students who had equivalent standardized test scores.

Once students were assigned to a condition, they were then ranked by their previous performance and randomly assigned by rank to a playing group, which I call a *pod*. This was done to ensure that all of the pods were of mixed academic ability. Once in a pod, students were then placed in mixed-ability pairs since mixed-ability pairs have been shown to help with student learning in math board games (Guberman & Saxe, 2000). The only indicator that was used for these mixed-ability groups was a standardized test score that the students had taken earlier in the year. Each condition played in one of the three classrooms, which meant that some students would play in the classroom where they had instruction. To avoid having any effects based on being in the same rooms or having the same partners, students would also switch classrooms every two weeks over the five sessions. This ensured that all students played in their classrooms at least once. When they would switch classrooms, students were also given new partners and new play groups, in another random assignment, based on their standardized test scores. Subjects switched partners and pairs to limit any particular subjects from having a bad (or good) partner throughout the study. In instances where the random assignment left a student with the same partner they had earlier, partners were then manually reassigned.

The teachers administered a pre-test assessment in their classrooms. Pre-tests were followed by one hour a week sessions in the game condition for five weeks. The students started each session in their original

classrooms, and then were led to their new classroom based on condition. Once in the new classroom, students were assigned partners. Students then played the game. Playing time varied from 35-45 minutes during each session, with student pairs playing against each other.

After each session, the classroom researchers met and discussed problems that they noticed children were experiencing with the software. Simple changes to the software were made each week to help assuage any concerns or problems that had arisen. Some changes made mainly improved the language in the game, such as shortening messages given to students. Other features made the administration of the game easier, such as preventing students from starting games with other groups. Each week, the single software code base would be updated and the images and text would be replaced to create the three versions. Students would then play with the latest versions of the software.

Students played with random cards during the first four sessions, which meant they played the game as intended. In the fifth session, rather than giving everyone a random set of cards, each pod was given a fixed deck of cards arranged in the same order on that last day. While the students thought the cards were random, they were manipulated to be the same for everyone. This was done in order to gather enough data to help answer a few questions about students' gameplay. A few of the students noticed and told researchers that they had received the same set of cards as before when their games had to be restarted. However, when the researchers were asked if they felt the students knew the cards were not random, and rather the same for everyone, they answered "No". In addition, on the fifth session researchers videoed two pods during their gameplay, one from the Characters condition and one from the Abstract condition. Two cameras were used to record each pod from two different angles. Video from each pod is approximately 30-45 minutes in length.

One week after having finished the five sessions of playing the game, the students were given a post-test by their classroom teachers in their original classrooms. This post-test looked exactly like the pre-test, except it added a few items at the end which were meant to assess the amount of fun that children had with playing the game, and their experiences with the game.

Results

Table 1 below shows the average gain in scores for each condition, with the standard deviation of those scores in parentheses. Average gain is reported as the difference in student scores from pre-test to post-test (e.g. post-test score – pre-test score). When collapsed across all conditions, students had an average gain score of 5.5, which indicates about a 13% improvement. Moreover, the 5.5 gain score from pre-test to post-test was significant when compared to zero, $t(72) = 7.86$, $p < .001$, $d = 1.85$, with a large effect size. Regardless of condition, student's scores on average increased from pre-test to post-test.

Table 1: Average gain scores from pre to post by condition, (standard deviation in parenthesis)

	Abstract Group	Original Group	Characters Group
Gain Score	4.00 (5.35)	5.59 (6.74)	6.89 (5.87)

A cursory examination of the gain scores in Table 1 shows a small increase in the directions that I expected. The Characters group had a higher average score than the Original group, who in turn had a higher average score than the Abstract condition. A one-way ANOVA with planned comparisons between the Abstract and Characters conditions and Abstract and Original conditions was performed to investigate whether or not those differences were significant by condition. The results from that ANOVA show that the small increase by condition is insignificant $F(2, 70) = 1.479$, $p = .24$. The planned comparison between the Original and Abstract conditions also was not significant $t(45) = .9$, $p = .37$. However, there was a marginal difference between the Character and Abstract conditions $t(50) = 1.72$, $p < .09$. The marginal difference between the Abstract and Characters conditions led to analyzing test scores by type of test questions, which were devised *a priori*. To reiterate, the five types of questions on the assessment were: questions that involved comparisons, questions involving decimals, questions involving complex word problems, questions involving fractions students would have seen in the game, and questions involving fractions students would not have seen in the game. I refer to each of these types of questions as Comparisons, Decimals, Word Problem, Old Fractions and New Fractions respectively.

All questions were categorized into one of these five groupings, and a subscore was calculated for each of the five categories for each student. The averages of these subscores along with the standard deviations for each condition are reported in Table 2. The table's first three subscores (Comparison, Decimal and Word Problem questions) show almost no gains and little difference by condition. Nevertheless, there seem to be trends in the Old Fraction and New Fraction questions in line with trends seen on the overall post-test. A one-way MANOVA then was run to determine if these trends were significant by condition. The MANOVA determined that the overall fit of group to these gain scores was not significant $F(2, 70) = .90$, $p = .54$.

Table 2: Sub-test gain scores by condition (standard deviation in parenthesis)

	Abstract Group	Original Group	Characters Group
Comparison gain	0.83 (1.66)	0.41 (1.71)	0.41 (1.80)
Decimal gain	0.38 (0.77)	0.68 (1.13)	0.56 (1.12)
Word Problem gain	0.63 (1.01)	0.68 (1.78)	0.78 (1.12)
Old Fraction gain	1.58 (3.20)	2.32 (3.56)	2.70 (3.29)
New Fraction gain	0.58 (2.10)	1.50 (2.74)	2.44 (2.39)

The next step was to look at the univariate scores for Old Fractions and New Fractions. The univariate test demonstrates that Old Fraction gain subscore by condition was not significant $F(2, 70) = .73, p = .49$. Nevertheless, the New Fraction gain subscore does have a significant difference by condition $F(2, 70) = 3.78, p < .03$. Performing a planned comparison on this gain demonstrates that while the difference between the Original and Abstract conditions was not significant, $t(45) = -1.29, p = .20$, there was a significant difference between the Abstract and Characters condition with a medium effect size $t(50) = 2.75, p < .01, d = .78$. The Characters condition had a higher average gain on the new fractions when compared to the Abstract condition.

Investigating the Relationship between Learning and Fun across Conditions

In addition to investigating the learning, it was also important to investigate the association between fun and learning across all three conditions. Figure 3 displays the gains that individual students had on the overall fractions test by condition and the gains that students had on the New Fractions subscore. The height of each digit on both graphs represents the gain that a particular student had from the pre and post assessments. The particular digit represented on the graph is that particular student's fun score. Upon an initial examination of both graphs, the Characters condition exhibited a pattern of having higher fun scores associated with higher learning gains. Nonetheless, this pattern does not seem apparent in the Abstract or Original conditions.

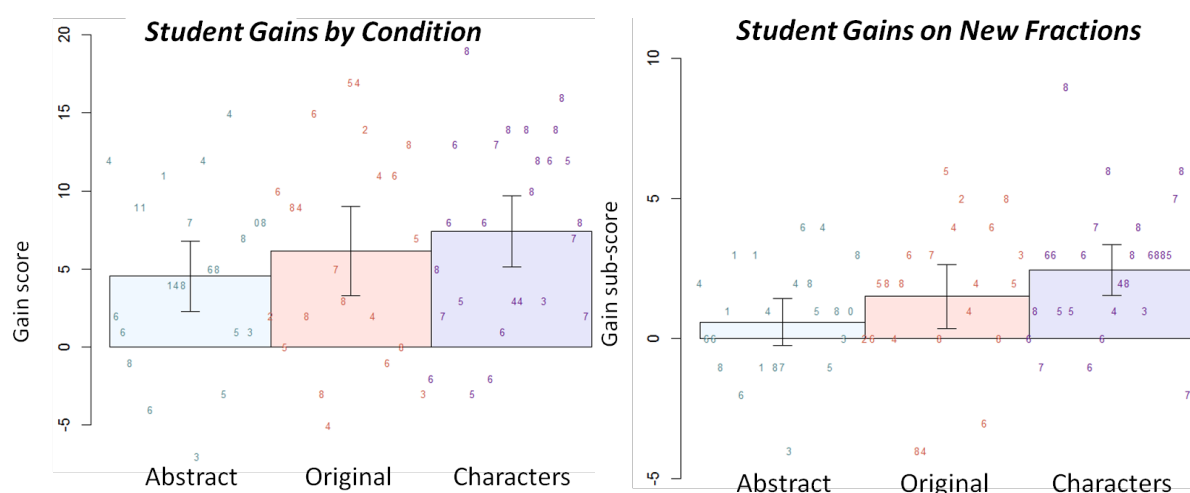


Figure 3. Student Gain Scores by Condition (Overall: left, and New Fractions: right). The height of each digit represents the gain score for a particular student, and the digit used represents the student's fun score rating

Table 3 below gives the correlations between the gain score (or subscore) and students' reported rating of fun. Positive correlations would provide evidence that there is a positive association between fun and learning. In a positive correlation, students who reported having more fun would be linked to stronger gains from pre-test to post-test. The correlation table does not display a positive correlation for either the Abstract or the Original conditions, but rather a negative one. Furthermore, the Abstract condition has a stronger correlation than the Original condition, but in the *negative* direction. Nonetheless, the Abstract condition exhibited no significant correlations in either the fraction gain scores $z(22) = -.89, p = .37$, or the sub-gain scores $z(22) = -.68, p = .50$, which means that there is not a likely association between the gain score and fun rating. Because the Original condition was closer to zero than the Abstract condition, no significance tests were performed.

Table 3: Correlations of fun ratings with gain or new fractions sub-gain scores by condition

	Abstract	Original	Characters
Correlation (r) of Overall Gain with Fun	-.14	-.08	.45
Correlation (r) of New Fractions with Fun	-.11	-.09	.34

While the Abstract and Original conditions did not exhibit any correlations, there does seem to be a positive correlation between students' self-reported ratings of fun and the gains that they made from pre-test to post-test in the Characters condition. The Characters condition exhibits significant positive correlations between the fun ratings and gain scores $z(25) = 2.92, p < .01$ as well as the New Fractions sub-gain scores $z(25) = 2.17, p < .03$.

The fact that there exists a positive correlation in the Characters condition, but no correlation in the other two conditions warrants further review. To make sense of what was happening, the following section details analysis of in-game measures during gameplay.

Investigating In-Game Measures

In addition to measures taken outside of the game, there were also internal measures that were created and stored in the software. These measurements were logs detailing activity from each user, such as when a question was answered without first getting it wrong and whether or not a student asked for help. To determine if the play by students differed by condition, a MANOVA was performed to investigate any differences by condition on several measures that were determined to be of possible use in earlier studies, such as the number of times they asked for help, the number of questions solved, and the average amount of time students took to answer a question. Performing the MANOVA demonstrates an overall fit of the data to the general model $F(2, 70) = 1.99, p < .01$, which would indicate that the type of play students exhibited seemed to differ by condition. Looking at the univariate scores, however, one notices that there is only one variable that carries a significant difference, the average number of questions a student would answer correctly on their first attempt in the last session $F(2, 70) = 4.72, p < .02$. By the last session, the effects that students have in playing the game should be different, which may explain the differences reported in Table 4 below.

Table 4: Select in-game results by condition

	Abstract	Original	Characters
Number of 1 st Attempts Correct in Last Session	6.00 (2.80)	8.32 (4.34)	9.15 (3.95)
Number of Incorrect Moves	8.13 (3.46)	7.86 (4.89)	7.59 (3.21)
Percentage of Incorrect Moves to Correct Moves	24.7 (9.1)	24.4 (9.1)	20.4 (6.6)

Performing a planned comparison between the Abstract and Original conditions yields a significant difference between them, $t(45) = 2.10, p < .05$, as well as when comparing the Abstract condition with the Characters condition $t(50) = 3.00, p < .01$.

One line of reasoning about the Characters condition or story conditions was that if the students understood the story and its rules, they would be less likely to make errors. If they were less likely to make errors, then the software would have logged less of those errors on average. Table 4 above reports the average amount of times that a student would incorrectly make an incorrect move across all sessions. The MANOVA's results displayed no significant differences across the three conditions $F(2, 70) = .12, p = .88$. However, thinking about the question critically, solely investigating the number of incorrect moves was not as appropriate a measure as the *percentage* of incorrect moves to total moves, as students who played more would naturally make more incorrect moves. Analyzing how this measure differed by condition yielded a larger F-statistic but demonstrated no significant difference $F(2, 70) = 2.13, p = .13$. Performing planned comparisons between the Original and Abstract conditions yielded no significant difference $t(45) = .12, p = .90$; however, there was a marginal difference between the Abstract and Characters conditions $t(50) = 1.84, p < .07$. There is a marginal trend towards students having less of a tendency for making illegal moves. Based on the earlier analysis that correlated student's fun with learning, it may be that interest is driving students to make fewer errors. When we look specifically at whether such a measure correlates with the amount of reported interest, we do get correlations in the correct direction for both the Original ($r = -.26, t(20) = -1.18, p = .25$) and Characters conditions ($r = -.17, t(24) = -.82, p = .42$), though neither of these are significant, while the Abstract condition shows no correlation between the percentage of incorrect moves and the amount of interest that students have ($r = 0$). While the in-game measures have generated insight into the student's behaviors in the game, the data logging techniques did not catch any of the social behaviors that the children exhibited *while playing the game*. In the future I hope to analyze and report qualitative findings generated from the video captured in the Abstract and Characters conditions.

Discussion

The main question this experiment sought to answer is whether or not story is a necessary feature in developing effective educational games. Exploration of the experimental study's results indicates a complex scenario for the role that story may play in educational video games. The initial results seem to indicate support for

arguments indicating that story may not be necessary. This conclusion could be inferred based on the experimental results that demonstrated that students increased from pre-test to post-test regardless of condition.

While the initial results from the final study could be used to argue that students can learn from a game without story, it is sensible to recognize the marginal differences that were present overall between the Abstract and the Characters conditions. In particular, the Characters condition displayed marginal trends in having larger gains from pre-test to post-test when compared to the Abstract condition. These results, combined with the high gains across all conditions, lay the foundation for a general viewpoint about the role of stories in educational games: a story may not be necessary for creating a successful educational video game, but its presence may help in increasing student's gains from pre to post.

The role of a story becomes clearer with respect to learning when looking specifically at the types of questions asked on the tests. Once the pre and post-tests were broken down by the type of fraction question, the results indicated that most of the types of questions did not differ by condition. Nevertheless, the results did indicate that one question category did differ by condition – new fractions questions. Students in the Characters condition had significantly higher gains when answering new fractions questions when compared to students in the Abstract condition. This means that students in the Characters condition had higher gains from pre-test to post-test when answering fraction questions they had not seen in the game, like “What is $\frac{2}{5}$ of 15?”. Not only does the story with characters marginally allow students to learn more, those students were better able to answer fractions questions they had not seen before. Because students in the Characters condition were able to answer questions they had not seen before, one may conclude that the Characters condition was better preparing students to think about fractions in the future. This finding coincides with earlier research with commercial video games and preparation for future learning (Arena, 2012). While the research presented in Arena's work establishes that video games can be used as preparation for future learning in a domains such as history, I would argue that the results from this experimental study both support Arena's research and further reveal the role that story has as being a potential critical aspect for that preparation. While game play may be strong enough to drive learning in a specific domain, having story may help students be better prepared to learn new material related to that domain in the future.

Another finding from the study was the high correlation present between the amount of fun that the students reported, and the amount that a student's score increased from pre-test to post-test. While such a correlation displayed no correlation in the Abstract and Original conditions, the Characters condition exhibited a strong correlation between fun and learning. Therefore, the results reveal a strong positive association between the amount of fun a Characters condition student had and the amount that they learned. One possible speculation for this comes from previous research done on transportation (Green & Brock, 2000). That research highlighted the role that the amount that a student was absorbed in a narrative would affect that amount that they were persuaded by such a narrative. Green & Brock's research may be one way to explain the high correlation – students who enjoyed and were more absorbed by the story were more persuaded and thus motivated to learn the material.

The results suggest that children who played in the story with characters condition got marginal learning benefits overall. Having a slightly better understanding of the role that story has in games can help the educational community think about how many resources might be devoted to the creation of a story and its incorporation into a game. If a designer does a good job in creating a powerful story that most people enjoy, that designer may have an easier time engaging students to work with the material. It may better prepare students for future learning, as well as providing them more entertainment.

References

- Arena, D. (2012, June 5). Commercial Video Games as Preparation for Future Learning..
- Barab, S. A., Sadler, T. D., Heiselt, C., Hickey, D., & Zuiker, S. (2007). Relating Narrative, Inquiry, and Inscriptions: Supporting Consequential Play. *Journal of Science Education and Technology*, 16(1), 59–82.
- Bower, G. H. (1978). Experiments on story comprehension and recall*. *Discourse Processes*, 1(3), 211–231. doi:10.1080/01638537809544437
- California Department of Education (Safe and Healthy Kids Program Office), & WestEd (Health and Human Development Department). (2012). *California Healthy Kids Survey*. Retrieved from <http://chks.wested.org/resources/chks-elem-bilingual-1213.pdf>
- Cordova, D., & Lepper, M. (1996). Intrinsic Motivation and the Process of Learning: Beneficial Effects of Contextualization, Personalization, and Choice. *Journal of Educational Psychology*, 88(4), 715.
- Garfield, J., & Ahlgren, A. (1988). Difficulties in Learning Basic Concepts in Probability and Statistics: Implications for Research. *Journal for Research in Mathematics Education*, 19(1), 44–63. doi:10.2307/749110
- Goldstone, R. L., & Wilensky, U. (2008). Promoting Transfer by Grounding Complex Systems Principles. *Journal of the Learning Sciences*, 17(4), 465–516. doi:10.1080/10508400802394898

- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*; *Journal of Personality and Social Psychology*, 79(5), 701–721. doi:10.1037/0022-3514.79.5.701
- Guberman, S. R., & Saxe, G. B. (2000). Mathematical Problems and Goals in Children's Play of an Educational Game. *Mind, Culture, and Activity*, 7(3), 201–216.
- Habgood, M. P. J., Ainsworth, S. E., & Benford, S. (2005). Endogenous fantasy and learning in digital games. *Simulation & Gaming*, 36(4), 483–498.
- Habgood, M. P. Jacob, & Ainsworth, S. E. (2011). Motivating Children to Learn Effectively: Exploring the Value of Intrinsic Integration in Educational Games. *Journal of the Learning Sciences*, 20(2), 169–206. doi:10.1080/10508406.2010.508029
- Honey, M., & Hilton, M. (2011). *Learning Science: Computer Games, Simulations, and Education*. Washington, D.C.: National Academies Press.
- Jimenez, O., Arena, D. A., & Acholonu, U. (2011). Tug-of-War: a Card Game for Pulling Students to Fractions Fluency. In C. Steinkuehler, C. Martin, & A. Ochsner (Eds.), *Proceedings of the Games, Learning, & Society Conference 7.0* (pp. 119–127). Madison, WI: ETC Press.
- Jimenez, O., Arena, D., & Acholonu, U. (2013). Educational Video Games: The Devil is in the Details. *Paper presented at the annual meeting of the AERA*, San Francisco, CA.
- Malone, T. W. (1981). Towards a Theory of Intrinsic Motivation. *Cognitive Science*, 4, 333–369.
- Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9(1), 111–151. doi:10.1016/0010-0285(77)90006-8
- Naone, E. (2005, July 12). Virtual Labor Lost: The failure of a highly anticipated, multiplayer game shows the limist of academic virtual worlds. *Technology Review*. Published by MIT. Retrieved from <http://www.technologyreview.com/web/19817/page1/>
- Nelson, B., Ketelhut, D. J., Clarke, J., Bowman, C., & Dede, C. (2005). Design-based research strategies for developing a scientific inquiry curriculum in a multi-user virtual environment. *Educational Technology*, 45(1), 21–27.
- Nunes, T. (2006). Fractions: difficult but crucial in mathematics learning. *Teaching and Learning Research Programme (TLRP) Research Briefing*. Retrieved from www.tlrp.org/pub/documents/no13_nunes.pdf
- Parker, L. E., & Lepper, M. (1992). Effects of fantasy contexts on children's learning and motivation: making learning more fun. *Journal of Personality and Social Psychology*, 62(4), 625.
- Petit, M. M., Laird, R. E., & Marsden, E. L. (2010). *A focus on fractions: Bringing research to the classroom*. New York: Routledge.
- Son, J. Y., & Goldstone, R. L. (2009). Contextualization in Perspective. *Cognition and Instruction*, 27(1), 51–89. doi:10.1080/07370000802584539
- Squire, K. D. (2003). Video games in education. *International Journal of Intelligent Games & Simulation*, 2(1), 49–62.
- Steinkuehler, C. (2006). Massively Multiplayer Online Video Gaming as Participation in a Discourse. *Mind, Culture, and Activity*, 13(1), 38–52.
- Strange, J. J., & Leung, C. C. (1999). How Anecdotal Accounts in News and in Fiction Can Influence Judgments of a Social Problem's Urgency, Causes, and Cures. *Personality and Social Psychology Bulletin*, 25(4), 436–449. doi:10.1177/0146167299025004004
- Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9(1), 77–110. doi:10.1016/0010-0285(77)90005-6
- Wilensky, U. (2002). Modeling nature's emergent patterns with multi-agent languages. In *Proceedings of EuroLogo* (p. 1–6). Retrieved from <http://ccl.northwestern.edu/papers/MEE/>
- Williamson, P., & Silvern, S. (1991). Thematic-fantasy play and story comprehension. *Play and early literacy development*, 69–90.

Acknowledgments

I would like to thank Shelley Goldman, Dan Schwartz, and Kristen Pilner Blair, whose suggestions and support have undergirded this study from its inception. I would also like to thank Dylan Arena and Ugochi Acholonu for helping create the game's initial ideas. Finally, I would like to thank the research facilitators from the AAA lab, as well as the teachers and students who participated. This work was supported by a grant from the National Science Foundation (NSF#0354453). Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Exploring Group-Level Epistemic Cognitions within a Knowledge Community and Inquiry Curriculum for Secondary Science

Alisa Acosta, Michelle Lui and James D. Slotta
Ontario Institute for Studies in Education, University of Toronto
252 Bloor Street West, Toronto, Ontario M5S 1V6 Canada
alisa.acosta@utoronto.ca, michelle.lui@utoronto.ca, jslotta@oise.utoronto.ca

Abstract: This paper explores Epistemic Cognition (EC) at the collective level within the context of an inquiry curriculum for high school biology. The “EvoRoom” curriculum was 10 weeks in duration, with two major units in evolution and biodiversity as well as a rich media “immersive simulation” activity and a field trip to a local zoo. All activities were designed according to the Knowledge Community and Inquiry (KCI) model, which guides the design of collective inquiry curricula where students make progress as a knowledge community with explicit connections to learning expectations. The paper applies Chinn’s EC framework to the design and enactment of EvoRoom and, by extension, to the KCI model. Findings reveal a shift in students’ perceptions of “sources of knowledge,” however students’ “justification of knowledge” was seen to be compromised in cases where students satisfied their original epistemic stance in favor of group consensus.

Introduction

Substantial research has investigated the use of technology enhanced learning environments to support collective inquiry – a form of learning in which students work together as an entire class to create and advance knowledge. For example, Scardamalia and Bereiter (2006) have advanced a theoretical perspective of knowledge building, employing an innovative technology environment called “Knowledge Forum” where students contribute and edit inquiry notes. Other researchers have explored various forms of online discussion tools (e.g. Hmelo Silver, 2010), wiki-based environments (e.g. Peters & Slotta, 2010; Najafi, Zhao & Slotta, 2011) and handheld data collection activities (e.g. Metcalf *et al.*, 2010) in exploring different aspects of students’ learning and engagement in collective inquiry.

One aspect of collective inquiry that has not yet received much study is that concerned with students’ epistemological knowledge – their beliefs about knowledge and learning, and how those beliefs influence their participation in the curriculum activities. Substantial research has addressed this topic for student inquiry in the wider area of computer supported collaborative learning, where small groups or individuals engage in structured inquiry tasks. For example, several lines of work have examined students’ beliefs about the nature of science, and science learning (Lin *et al.*, 2013) or the students’ understanding of collaboration (Najafi & Slotta, 2010). However not much attention has been given to the collective aspects of epistemology. It is important to study how students’ epistemic cognition influences collective inquiry. For example, if students have little experience learning as a community, or if their understandings about learning are not aligned with those of the designed curriculum (e.g. if they are expecting to learn content with the aim of individual achievement), then the outcome of the enacted curriculum may diverge widely from what was intended. Thus, an understanding of epistemic cognition in relation to collective inquiry is vital to our theoretical models, as well as to our design of curriculum and technology environments

Drawing upon the theoretical framework for epistemic cognition, developed by Chinn *et al.* (2011), the present study is guided by the following research questions:

1. How does the design of collaborative inquiry activities affect the ways in which knowledge is justified and shared within a knowledge community?
2. How are students’ epistemic stances influenced by the nature of the collective inquiry design?

Theoretical Foundations

“Epistemic cognition” (EC) is a term used to describe any explicit or tacit cognitions that pertain to epistemological matters, such as knowledge, beliefs, truth, sources, justification, evidence, understanding, and explanation (Chinn *et al.*, 2011). A number of studies have shown that students’ epistemic beliefs are an important predictor of achievement in a variety of learning domains, including information processing (Garner and Alexander, 1994), reading comprehension (Rukavina & Daneman, 1996), test performance (Schommer *et al.*, 1992), argumentation (Kuhn, 1991), and the ability to synthesize information from multiple sources (Strømsø and Bråten, 2009). Other studies have revealed the role of epistemic beliefs in affecting chosen learning strategies (Ryan, 1984; Schommer *et al.*, 1992), motivation and behaviour (Pintrich *et al.*, 1993), and attitudes such as learned helplessness (Qian and Alvermann, 1995). In the domain of school science, Windschitl and

Andre (1998) demonstrated that students with more sophisticated epistemic beliefs (e.g. the belief that knowledge is complex, cumulative, and context-dependent) exhibited greater learning gains when engaged with constructivist pedagogies compared to individuals with less sophisticated beliefs (e.g. the belief that knowledge is simple, quick, and certain). Kardash And Scholes (1996) revealed that students' beliefs about the certainty of knowledge affected the ways they handled and presented contradictory evidence in scientific research, and Qian and Alvernann (1995) further showed that students' beliefs about the simplicity and certainty of knowledge impacted the levels of conceptual change they experienced in school science. It is therefore important to consider students' epistemic beliefs and their interaction with learning when developing theoretical models of learning or specific curricular interventions.

Several scholars have attempted to identify specific dimensions of epistemic cognition to provide general models or frameworks (see, for example, Perry, 1970; Schommer, 1990, 1993; Hofer and Pintrich, 1997). In 2011, Chinn, Buckland and Samarapungavan extended the work of Hofer and Pintrich (1997) by developing an expanded framework for epistemic cognition. Their model included the addition of several new components and subcomponents of EC, as well as the specification of a finer grain size of cognitions within each of these dimensions in order to account for contextual and situational differences in learning processes. The five dimensions of Chinn *et al.*'s 2011 EC framework are:

1. Epistemic aims and epistemic value
2. Structure of knowledge and other epistemic achievements
3. Sources and justification of knowledge and related epistemic stances
4. Epistemic virtues and vices
5. Reliable and unreliable processes for achieving epistemic aims

Chinn *et al.* (2011) recognize that their description of this EC framework maintains a focus on individual cognitions, and the authors further suggest that an opportunity exists for future research to explore these cognitions at the level of groups, which would include students engaged in collective inquiry within a constructivist learning environment. To this end, this paper explores one of the above dimensions, "sources and justification of knowledge and related epistemic stances," in relation to an innovative pedagogical model for collaborative learning in secondary science called "Knowledge Community and Inquiry" (KCI).

As defined by Slotta and Najafi (2010), a knowledge community is one where members (a) collectively develop a shared knowledge base (b) establish characteristic practices for knowledge creation or advancement, and (c) share in discourse for idea sharing, critique and improvement. From an epistemological perspective, the knowledge community approach represents a key shift from the notion of self-as-learner, where an individual is potentially in competition with peers, to one of collaboration and cooperation in which shared knowledge advancement is favored over individual gains. In the context of K-12 classrooms, two widely researched examples of the knowledge community approach are Fostering Communities of Learners (FCL) (Brown & Campione, 1996), and Knowledge Building (Scardamalia & Bereiter, 2006). Although both of these examples have been implemented in K-12 contexts, current school structures and content-heavy curriculum demands can often make these approaches inaccessible to course instructors – particularly at the secondary level, and particularly in content-heavy domains like science. Knowledge Community and Inquiry (KCI) was developed for secondary science as a means of blending the core philosophies of the knowledge building approach with the structural and scripted affordances of scaffolded inquiry (Slotta & Peters, 2008; Slotta & Najafi, 2012). The KCI model includes five major principles, each including a set of epistemological commitments, pedagogical affordances, and technology elements. Together, these guide the creation of inquiry activities, peer interactions and exchange, and cooperative knowledge construction (Slotta et al, 2013).

Until the present study, the epistemic elements of KCI had not been explicitly tested or evaluated. To assess students' epistemic cognitions within KCI, we first evaluated the design of a KCI biology curriculum unit in terms of its epistemic elements, which was done by evaluating the design in terms of the stated epistemic commitments of KCI. We then then evaluated the enactment of the curriculum in terms of actual epistemic cognitions observed in student interactions. The curriculum was a ten-week Grade 11 Biology unit that met the Ontario Ministry requirements for evolution and biodiversity and included activities that were situated in part within a unique immersive environment called "EvoRoom" (Lui & Slotta, 2012). It should be noted that while the KCI model served as an important referent and guide for design decisions, the design of the EvoRoom curriculum was not explicitly concerned with the role of epistemic cognition within KCI. While such elements are essential to the KCI model, they were not at the forefront of concern for the EvoRoom research project, which was focused on designing activity sequences that engaged students with the relevant biology content, as well as a collective immersive simulation environment (Lui & Slotta, 2013). The present research examines the role of epistemic cognition within the EvoRoom curriculum design by taking an 'epistemological pass' at the most recent design iteration. As noted above, this paper presents the findings for one of Chinn *et al.*'s (2011) dimensions of EC, "sources and justification of knowledge and related epistemic stances."

Methodology

Research Setting and Design Considerations

The term ‘EvoRoom’ refers to a 10-week multi-locational curriculum, but also to an actual room that was constructed using smart classroom technologies to provide an immersive simulated rainforest environment. Students’ movements and interactions within this smart classroom environment – where they go in the room, and with whom – are carefully orchestrated and dependent upon real-time observations that students make using tablet computers. The broader 10-week curriculum was designed to fulfill the requirements for evolution and biodiversity in Grade 11 Biology. At the time of this study, the EvoRoom curriculum was undergoing its second design iteration. The curriculum design included activities across a number of different contexts, including at home, within the students’ regular classroom, the smart classroom, and on a field trip to the local zoo. The main components of the EvoRoom curriculum were organized around the iterative use of an online portal that served as a knowledge base for the community. Periodic inquiry and knowledge construction activities were blended with traditional classroom lectures in the following sequence:

1. Pre-Activity: Epistemic Orientation (Week 0)
2. Online Learning Portfolio/Knowledge construction (Ongoing)
3. EvoRoom Evolution Activity (Week 2)
4. Zoo Field Trip Activity (Week 8)
5. EvoRoom Biodiversity Activity (Week 10)

This paper will focus on the Zoo Field Trip Activity and the EvoRoom Biodiversity Activity. These two activities were chosen as the focus of this study because they employed the same technology platform, Zydeco, to support student inquiry (Kuhn et al., 2010). However these two activities were quite distinct, in terms of their epistemic nature (i.e. their purpose within the knowledge community) and use of collaborations. In regard to the Chinn et al’s (2011) “justification of knowledge” dimension, these affordances were directly linked to the design of the Zydeco environment, which was developed by another research team with the aim of supporting evidence-based justifications (Zhang & Quintana, 2012). Hence, KCI might not, on the basis of its own principles, have emphasized such justificatory elements, but its choice of Zydeco as an observational data collection environment enabled this form of EC to be prominent within the enactment.



Figure 1. *Top-left:* Students using Zydeco to collect data within the EvoRoom smart classroom; *Bottom-left:* Zydeco data collection screen used by students to collect and tag multimodal data artifacts (e.g. text notes, photos, video notes, audio notes) and contribute them to a shared evidence base; *Right:* Zydeco explanation screen where students generated knowledge claims and supported them using both evidential justification (i.e. artifacts from the shared evidence base) and non-evidential justification (i.e. reasoning).

Co-Design Team

In order to ensure that the curriculum design (i.e., all activities, tools, materials and interactions) was suitable for high school biology, a co-design approach was used (Roschelle, Penuel & Shechtman, 2006). The co-design team consisted of three researchers (two graduate students and one faculty member), three programmers, and the classroom teacher. Throughout the EvoRoom design process, the teacher was highly involved in the development of the orchestrational scripts and technology elements that went into the design. The teacher met weekly with two researchers over a two-year period, providing valuable feedback with regards to tool development and the overall curricular goals for the evolution and biodiversity units.

Participants

The participants for the current design iteration consisted of two sections of Grade 11 Biology (n=56) from a high-achieving secondary school within a large and ethnically diverse urban setting.

Data Sources

Data sources for both the Zoo Field Trip Activity and the EvoRoom Biodiversity Activity included students' digital learning artifacts contributed through the Zydeco platform, as well as researcher observations and field notes from both activities. Additionally, an open-ended survey was administered to students both before and after the full 10-week EvoRoom curriculum unit, with two items pertaining to sources and justification of knowledge.

Zoo Field Trip Activity Sequence

The zoo field trip was situated between the two immersive experiences in the EvoRoom smart classroom. Prior to the actual field trip, students were given a full period of training on how to use the mobile app Zydeco (Cahill et al., 2010; Kuhn et al., 2010), which was used to collect evidence and observations whilst in the field at the zoo. On the day of the zoo field trip, students were divided into groups of three or four, and each group was given two mobile devices: an iPod touch and an iPad or iPad mini. It should be noted that all sections of Grade 11 Biology – including course sections outside of the sample group – participated in this mobile activity and used Zydeco to contribute evidence to the shared knowledge base.

At the zoo, groups were assigned to a particular species group (e.g. birds, fish, primates, reptiles and amphibians, plants and insects, and other mammals), as well as a designated geographic region of the zoo (e.g. African rainforest, African savanna, Australasia, Indomalaya, Eurasia, and Americas). Their task was to collect evidence, using a variety of multimodal formats (e.g. text notes, audio notes, video notes, photographs), in order to take a position on three issues: 1) the unifying principles underlying biodiversity; 2) human impacts on biodiversity, and 3) what makes an effective educational exhibit. The data from all groups was pooled into a shared evidence base. Students returned to school for the final period and worked in the computer lab using the Zydeco web platform. Using a 'claim–evidence–reasoning' structure, students had the opportunity to draw upon the full set of evidence gathered by their peers from the zoo in order to generate and support their knowledge claims. Students were given one week to complete this individual, summative assignment.

EvoRoom Biodiversity Activity Sequence

During the final week of the EvoRoom curriculum, students participated in a second activity within the smart classroom. Students were assigned to one of four sessions (A to D) in the EvoRoom, and each session consisted of four groups of approximately three or four students. Leading up to this activity, students had learned about biodiversity throughout their regular classroom activities, and groups were asked to make predictions on the online learning portfolio as to how their assigned climatic 'scenario' would impact biodiversity (e.g. high temperature, low temperature, earthquake, tsunami, high rainfall, low rainfall, high temperature, low temperature). When students entered the smart classroom, the screens around the room depicted the present-day Borneo-Sumatran rainforest. After making some initial observations, each of the four walls was transformed to represent one of the climatic scenarios that had been assigned. Within their groups, students used the mobile app Zydeco to collect evidence from each of the four walls in order to identify which rainforest best represented their assigned scenario. The multimodal evidence collected using Zydeco was tagged and aggregated in real-time on the front IWB. The resulting aggregate of evidence served as a reference for a full-class discussion, facilitated by the teacher, and provided visual clues as to which scenario was depicted by which station.

Following the evidence-gathering stage, students worked in their groups to generate claims as to which of the four walls most likely represented their climatic scenario. Using the claims–evidence–reasoning structure within Zydeco, groups took turns presenting their findings to their classmates. After all four scenario solutions were revealed, the teacher facilitated a deeper whole-class discussion related to human impacts on biodiversity.

Analytic Approach

For the purposes of this analysis, Chinn *et al.*'s third dimension was broken into three constituent parts: "Sources of knowledge," "justification of knowledge" and "epistemic stance." "Sources of knowledge" was evaluated using the EvoRoom pre/post survey, in which students were asked to identify their main sources of knowledge in school science. Responses were open-ended, and the frequency of each response was recorded using a tally system, with students often reporting multiple sources of knowledge within a single answer. "Justification of knowledge" for both the Zoo Field Trip Activity and the Biodiversity Activity was examined using a qualitative, descriptive analysis of students' Zydeco contributions, with researcher field notes facilitating a comparison of the way knowledge was justified in each of these contexts. Additionally, a post-survey item asked students to identify the knowledge negotiation strategies that were used throughout the Biodiversity Activity.

An analysis of epistemic stance at the group level was conducted for the Biodiversity Activity only, since knowledge claims for the Zoo Activity were submitted individually. Students responded to a post-survey item asking whether their group reached a consensus on their final claim statement. Responses to this survey item were then cross-referenced with the “solutions” to the Biodiversity Activity, indicating whether or not their final claim statement was actually correct. In cases where there was consensus within the group and the claim statement was correct, responses were coded as “true certainty.” In cases where there was consensus within the group but the claim statement was incorrect, responses were coded as “false certainty.” Finally, in cases where a group consensus was not reached, responses were coded as “uncertain.”

Analysis and Findings

Sources of Knowledge

The sources of knowledge students identified in this pre-survey (n=43) mainly consisted of authoritative sources (89%) such as the course textbook (34%), the teacher (32%) and other authoritative sources such as online resources or publications (23%). Only a small percentage of students identified themselves (3%) or their peers (8%) as sources of knowledge. Throughout the EvoRoom curriculum activities, students were asked to draw from a variety of knowledge sources in order to make contributions to the shared knowledge base, and also to regard the shared knowledge base as their community resource, representing the pooled ideas and knowledge artifacts contributed by their peers. In the EvoRoom Post-Survey, students were asked to identify the sources of knowledge they used throughout the EvoRoom curriculum unit. Here, there was a dramatic shift in the sources of knowledge that students identified. Results indicated that 33% of responses included authoritative sources (e.g. textbook, teacher, other external sources), 28% of responses identified their peers and/or the knowledge community as a source of knowledge (e.g. peer discussion, shared knowledge base, aggregate displays), and 38% of students identified themselves as sources of knowledge (e.g. through primary observations, prior learning/memory, reasoning/logic) (see Figure 2).

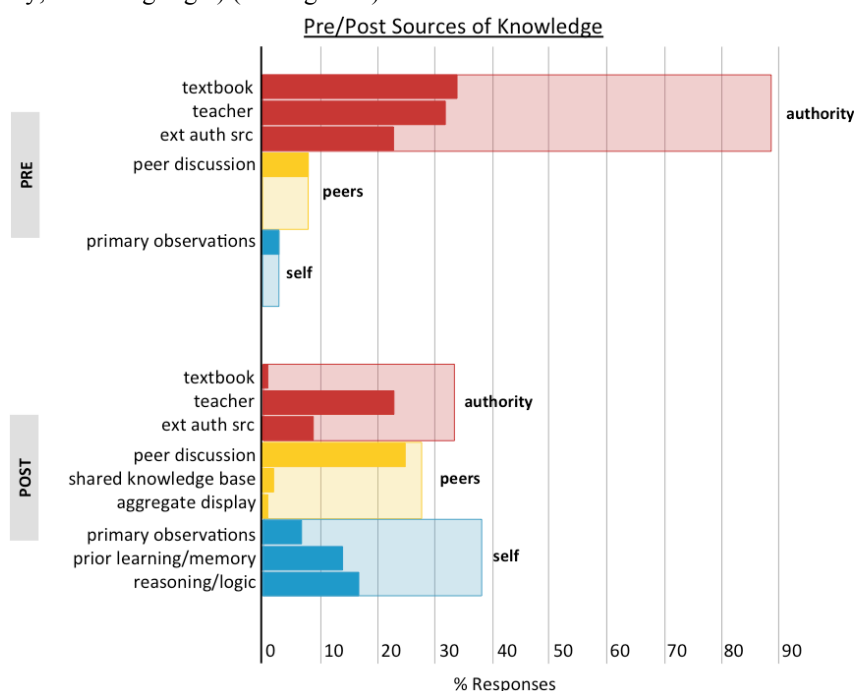


Figure 2. Pre and post survey results showing students' sources of knowledge in 'traditional' school science (pre) and in the EvoRoom curriculum (post). Pre-survey results indicate a heavy reliance on authoritative sources of knowledge (89%), whereas post-survey results show a more even distribution between authority (33%), peers (28%) and the self (38%) as sources of knowledge.

Justification of Knowledge

For both the Zoo Field Trip and the Biodiversity Activity, justification of knowledge was built into the design of the Zydeco app, as students had to provide both evidence and reasoning to support their claims statements. However, the way that this justification was enacted was quite different between these two contexts. For the Zoo Field Trip, although students collaborated to collect data, pool their evidence, and draw from this shared evidence base, students' final knowledge claims were ultimately completed individually. Questions were open-ended, and students had to be conscientious about their choice of evidence and reasoning in order to perform

well on this summative assignment. The Biodiversity Activity, however, had three key differences: (1) students completed their knowledge claims in groups, (2) there was a “right” and “wrong” answer, and (3) this activity was not explicitly for marks. In contrast to the Zoo activity, field observations revealed that use of evidence artifacts to support claims statements in the Biodiversity Activity was almost an afterthought, with group consensus taking priority over evidential justification. Here, the majority of knowledge negotiations among group members occurred verbally, with Zydeco claim statements reflecting the product of these negotiations. The EvoRoom post-survey revealed that the negotiation strategies students described as occurring throughout this activity included taking turns “reasoning out loud” (41%), using the process of elimination (18%), collecting additional evidence (11%), using argumentation/debate (11%), listening in to other groups’ decisions (9%), and bringing the decision to a group vote (9%).

An additional finding for the Biodiversity Activity was related to the quantity of evidence that students used in their claims statements compared to their ability to correctly identify their climatic scenario. As shown in Figure 3 below, Session C used the highest percentage of the data artifacts that they collected as supporting evidence in their claims statements (93%). At the same time, Session C was the only session for which all four groups correctly identified their climatic scenario. In Session A, which had the second highest percentage of evidence artifacts used towards claims statements (57%), one of the four groups correctly identified their climatic scenario. However in Sessions B and D, which used 38% and 29% of evidence artifacts, respectively, none of the groups were able to correctly identify their climatic scenario. These findings highlight the importance of evidentiary justification in supporting collective knowledge negotiations.

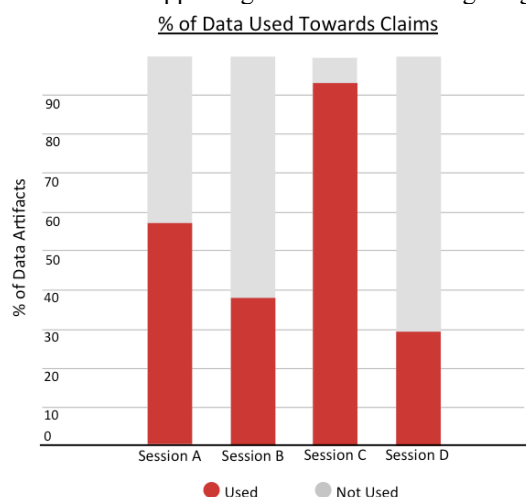


Figure 3. The percentage of data artifacts within each Biodiversity Activity session that were used as supporting evidence throughout the generation of group knowledge claims.

Epistemic Stance

Epistemic stance refers to the position one takes with respect to a knowledge claim (e.g. certainty, uncertainty, entertaining an idea, utilizing an idea as a working hypothesis, withholding judgment on an idea, etc.). For activities in which students collaboratively negotiate knowledge, such as in the generation of knowledge claims within the Biodiversity Activity, there are opportunities for students to take a variety of epistemic stances with regards to shared ideas or decisions. In each of these cases, it is possible that a student may be engaged in collaboratively generating a knowledge contribution, however he/she may have an epistemic stance that is in contrast or conflict with that of her collaborating peers. Satisficing one’s true epistemic stance in order to appease group members (e.g. by establishing consensus within the group by means of a “vote”) would detract from the justificatory rigor of the inquiry.

The Zydeco Biodiversity claims were analyzed in order to identify possible instances of satisficing within group responses. Following the Biodiversity Activity, students were given a survey in which they were asked to report whether or not their group reached a consensus about which rainforest depicted their scenario. 39% of respondents indicated that their group had consensus throughout the duration of the activity, 45% indicated that they came to a consensus after engaging in some knowledge negotiations, 11% indicated that they came to a consensus by “voting” or satisficing their response, and 5% did not reach a consensus. These responses were then cross-referenced with the group claims statements that were submitted through Zydeco. Responses were coded for group Epistemic Stance using the following categories:

1. **True certainty** – there was consensus in the group and the claim was correct
2. **False certainty** – there was consensus in the group but the claim was incorrect
3. **Uncertainty** – there was no consensus in the group

Results indicated that “False certainty” occurred in 58% of cases and “True certainty” occurred in 37% of cases. As noted above, only 5% of students reported that they did not reach consensus and their claim statements were therefore uncertain. Within the sub-set of students that reported satisficing their responses, only 25% of these groups were successful in correctly identifying their climatic scenario.

Discussion

Findings for “sources of knowledge” demonstrated that engaging students in activities where they rely on the ideas of their peers, as well as their own observations, can result in a noticeable shift in their epistemic cognitions within this dimension. However, the analysis for “justification of knowledge” revealed a distinct contrast between the way knowledge justifications occurred for individual knowledge claims (i.e. the Zoo Field Trip Activity) versus group knowledge claims (i.e. the Biodiversity Activity). Although the technological scaffolds within Zydeco were identical in both cases, these two activities were pedagogically quite distinct. The “claims statements” that students generated for the Zoo Field Trip Activity were in response to open-ended questions related to a driving question about the unity of biodiversity. Conversely, the claims statements within the Biodiversity activity were in response to a closed-ended, right-or-wrong question (asking which of the four walls of the EvoRoom depicted a particular climatic scenario). Here, group consensus was favored over justificatory rigor, and the addition of supporting evidence to claims statements commonly occurred after a group decision had already been reached. As a result, 58% of students were “falsely certain” that their claim was correct (i.e. their group had reached a consensus, however their claim statement was ultimately incorrect). This lack of evidentiary justification to support group claims statements is indicative of the satisficing of epistemic stances within the group. However, as indicated in Figure 2, groups who used more evidence to support their claims statement were more likely to reach “true certainty” (i.e. where their group had reached a consensus and their claim statement was correct).

In light of these findings, the following design priorities are recommended with respect to “justification” and “epistemic stance”:

1. Although technological scaffolds to support evidentiary and non-evidentiary justification were built into the design of the shared knowledge base, the technology environment should also capture knowledge *negotiations* amongst members of the knowledge community. This would not only assist in preventing student satisficing, but would also to provide a history of the idea-growth *process* within the knowledge community.
2. To support the justification of knowledge in collaborative contexts, inquiry questions should be open-ended and should promote explanatory coherence over ‘correctness.’
3. From a pedagogical perspective, activities should be designed that help students to understand the importance of justificatory rigor and how the satisficing of their epistemic stance in group learning activities might compromise the integrity of the inquiry.

References

- Brown, A. L. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist*, 52(4), 399–413.
- Cahill, C., Kuhn, A., Schmill, S., Pompe, A., and Quintana, C. (2010). Zydeco: Using mobile and web technologies to support seamless inquiry between museum and school contexts. Proceedings of the 9th International Conference on Interaction Design and Children (pp. 174-177). June 9-12. Barcelona, Spain. Association for Computing Machinery Inc.
- Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: arguments from philosophy and psychology. *Educational Psychologist*, 46, 141–167.
- Garner, R., and Alexander, P. A. (eds.). (1994). *Beliefs About Text and Instruction With Text*, Erlbaum, Hillsdale, NJ.
- Hmelo-Silver, C. E. (2006). Design principles for scaffolding technology- based inquiry. In A. M. O’Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative reasoning, learning and technology* (pp. 147–170). Mahwah, NJ: Erlbaum.
- Hofer, B. K., and Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88–140.
- Kardash, C. A. M., and Scholes, R. J. (1996). Effects of preexisting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology* 88: 260–271.
- Kuhn, A., Cahill, C., Quintana, C., and Soloway, E. (2010). Scaffolding science inquiry in museums with Zydeco. Proceedings of the 28th International Conference on Human Factors in Computing Systems (pp. 3373-3378). April 10-15. Atlanta, USA. Association for Computing Machinery Inc.
- Kuhn, D. (1991). *The skills of argument*. Cambridge, England: Cambridge University Press.

- Lin, T.J., Chai, C.S., and Tsai, C.C. (2013). High school students' scientific epistemological beliefs, motivation in learning science, and their relationships: A comparative study within the Chinese culture. *International Journal of Educational Development*. 33, 37-47.
- Lui, M., and Slotta, J.D. (2012). Designing immersive environments for collective inquiry. *Proceedings of the Tenth International Conference of the Learning Sciences – Volume 2* (pp. 12-14). July 2-6. Sydney, Australia: International Society of the Learning Sciences, Inc.
- Lui, M., & Slotta, J. D. (2013). Immersive simulations for smart classrooms: exploring evolutionary concepts in secondary science. *Technology, Pedagogy and Education*, (ahead-of-print), 1-24.
- Metcalfe, S., Dede, C., Grotzer, T., and Kamarainen, A., (2010, May). EcoMUVE: Design of virtual environments to address science learning goals. Paper presented at the American Educational Research Association (AERA) Conference, Denver, CO.
- Najafi, H., & Slotta J.D. (2010). Analyzing equality of participation in collaborative inquiry: toward a knowledge community. *Proceedings of the 9th International Conference of the Learning Sciences*, Chicago. Volume 1. 960–967. International Society of the Learning Sciences (ISLS).
- Najafi, H., Zhao, N. & Slotta, J.D. (2011). Facilitating Knowledge Communities in Science Classrooms through Scripted Collaboration. *Proceedings of the Ninth International Computer-Supported Collaborative Learning Conference*, Hong Kong. Volume 1. 996-997. International Society of the Learning Sciences (ISLS).
- Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart and Winston.
- Pintrich, P. R., Marx, R. W., and Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research* 63: 167–199.
- Qian, G., and Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology* 87: 282–292.
- Rukavina, I., and Daneman, M. (1996). Integration and its effect on acquiring knowledge about competing scientific theories from text. *Journal of Educational Psychology*, 88, 272–287.
- Roschelle, J., Penuel, W. R., & Shechtman, N. (2006). Co-design of innovations with teachers: Definition and dynamics. *Proceedings of the International Conference of the Learning Sciences* (pp. 606-612). Bloomington, IN.
- Ryan, M. P. (1984). Monitoring text comprehension: Individual differences in epistemological standards. *Journal of Educational Psychology* 76: 248–258.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. *The Cambridge handbook of the learning sciences*, 97–115. Retrieved from http://ikit.org/fulltext/2006_KBTheory.pdf
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology* 82: 498–504.
- Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology* 82: 498–504.
- Schommer, M., Crouse, A., and Rhodes, N. (1992). Epistemological beliefs and math text comprehension: Believing it is simple does not make it so. *Journal of Educational Psychology*, 84: 435–443.
- Slotta, J.D., and Najafi, H. (2010). Knowledge communities in the classroom. *International encyclopedia of education*, 8, 189-196.
- Slotta, J.D. and Najafi, H. (2012). Supporting Collaborative Knowledge Construction with Web 2.0 Technologies. In N. Lavigne (Ed.), *Emerging Technologies for the Classroom: A Learning Sciences Perspective* (pp. 93-112). Springer.
- Slotta, J.D., and Peters, V. L. (2008). A blended model for knowledge communities: Embedding scaffolded inquiry. *International Perspectives in the Learning Sciences: Creating a learning world. Proceedings of the Eighth International Conference for the Learning Sciences* Utrecht. pp. 343-350. International Society of the Learning Sciences (ISLS).
- Slotta, J. D., Tissenbaum, M., & Lui, M. (2013, April). Orchestrating of complex inquiry: three roles for learning analytics in a smart classroom infrastructure. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 270-274). ACM.
- Strømsø, H. I., and Bråten, I. (2009). Beliefs about knowledge and knowing and multiple-text comprehension among upper secondary students. *Educational Psychology*, 29, 425–445.
- Windschitl, M., and Andre, T. (1998). Using computer simulations to enhance conceptual change: The roles of constructivist instruction and student epistemological beliefs. *Journal of research in science teaching*, 35(2), 145–160.
- Zhang, M. and Quintana, C. (2012). Scaffolding strategies for supporting middle school students' online inquiry processes. *Computers in Education*. 58(1), 181-196.

High School Students' Parameter Space Navigation and Reasoning during Simulation-Based Experimentation

Hee-Sun Lee, University of California, Santa Cruz, 218 ISB, Physics Department, hlee58@ucsc.edu
Amy Pallant, Robert Tinker and Paul Horwitz, The Concord Consortium,
Email: apallant@concord.org, rtinker@concord.org, phorwitz@concord.org

Abstract: Students formulated their own questions for a virtual spring/mass system and collected and analyzed data in the InquirySpace environment featuring probes, computational models, and data visualization software. We investigated how students navigated and reasoned with the parameter space defined by a set of manipulative variables related to a virtual spring/mass system. We analyzed logging data of 31 high school student groups and a student group's Screencast video and found that (1) students' investigations followed stages: exploration, crude initial investigation, refined investigation, and data analysis, (2) some logging events acted as markers for these stages, (3) students used more extreme parameter values during exploration than collecting data to answer their questions, and (4) students' discourse was mostly centered around their parameter space navigation and analysis.

Introduction

Inquiry-based science learning has been emphasized in recent science education reform documents in the last fifteen years. In the original National Science Education Standards (NRC, 1996), inquiry was stated as “a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results” (p.23). Since then, what constitutes inquiry-based learning has gone through several revisions, but what remains consistent is to promote student-initiated inquiry, the type of inquiry that seeks students' active participation, contribution, reflection, and learning.

At the heart of empirical inquiry is experimentation where evidence is generated for scientists to answer their questions, connect to theory, elucidate or hypothesize mechanisms behind phenomena, and develop arguments. Scientific experiments have traditionally involved physical apparatus. However, the advancement in computer technologies has provided scientists with an additional tool for exploring complex scientific topics. In the literature, students' experimentations with both physical apparatus and simulations have been investigated. With physical experimentation, research has focused on whether, how, or to what extent students can design and conduct experiments (Hackling & Garnett, 1992; Kanari & Millar, 2004). Students' experimentation skills include recognizing multi-covariate relationships (Amsel & Brock, 1996), dealing with experimental errors (Allchin, 2012), addressing variability in the data (Masnick, Klahr, & Morris, 2007; Petrisino, Lehrer, & Schauble, 2003), applying statistical reasoning (Lubben et al., 2001), treating anomalous data (Chinn & Brewer, 1993), and revising hypotheses, experiments, and questions after reflecting on evidence (Schauble, 1996). Studies have found that students have difficulties in recognizing, identifying, and controlling variables (Toth et al., 2000). When studying students engaged in simulation-based experiments, McElhaney and Linn (2011) found that students' experimentation patterns can be characterized as intentional, random, and exhaustive based on the number of trials attempted by students with or without experimental coherence and found that these patterns resulted in what parameter values students explored during their simulations.

In this paper, we argue that the type of reasoning that enables students to conduct an empirical inquiry is much broader than student reasoning investigated in any of these studies such as controlling variables or identifying outliers in the data. The purpose of this paper is to characterize student reasoning in a much broader sense to capture students' planning, operationalizing, navigating, and reflecting on multiple experimental runs to generate evidence to answer their questions. Since each experimental run can be summarized based on a set of parameters, we name this type of student reasoning *parameter space reasoning* (PSR). In this paper, we focus on PSR involved in simulation-based experimentations. Research questions are (1) how did students navigate and analyze the parameter space in their experimentation with a simulated spring/mass system?; and (2) what aspects of parameter space reasoning were demonstrated in the discourse of one typical student group through the course of a simulated experimentation?

Theoretical Framework: Parameter Space Reasoning

A parameter is referred to herein as a measurable factor that defines a system or determines the conditions of a system's function. In science, a parameter space is defined as all possible combinations of values related to a set of parameters that define a system. A different set of parameters is used for a different system or a phenomenon under investigation. Parameter space also depends on the conceptualization of an experimenter or a modeler

who studies a system. An experimental run can be described as a point in an n -dimensional parameter space where the number of parameters related to a system is n . Some parameters are salient to determine a given outcome while others are not. The range of a parameter can be plotted on one axis (typically, x) and the outcome of a system can be plotted on another axis (y , for example). In a spring-mass system, an outcome variable of interest such as period can be estimated from a sinusoidal graph between the distance from the center and time, instead of directly measuring it. We call a graph that shows changes in a variable over time as a time-series graph. If all other remaining parameters are kept constant, then a point representing (parameter space value, outcome value) can effectively summarize the result of an experimental run. We call this plot a parameter-outcome graph. It is quite possible that experimenting with different regions of the parameter space yields different results, e.g., a spring too stretched to lose its elasticity. Therefore, it is important to explore the adequate range of the parameter space to test a model or a relationship in experimentation.

An investigation includes multiple experimental runs, each of which is defined uniquely by a set of values chosen for the run (as one run sets a single value for each parameter, two values of the same parameter cannot be investigated simultaneously in a single run). The region of parameter space examined by students can be traced. Before experimental runs, students need to plan for which variables to vary and how to vary them. At the end of each experimental run, students must make decisions on their next step such as redoing the same experiment, varying parameter values, checking their equipment, and eventually concluding their investigation. After all the data are collected, students need to think about the quality of the data and recognize relationships between the parameters they manipulated and system outcomes they measured. PSR captures this array of cognitive processes as they relate to empirical inquiry, and entails, among other things, the ability to compare an experimental run to other runs that differ in the value of at least one parameter. Table 1 lists PSR in three phases of an investigation: parameter setting, data collection, and data analysis and describes the reasoning in each investigation step and how PSR can be observed.

Table 1: Characterization of parameter space reasoning (PSR) during student investigation.

Investigation phases	Investigation steps	PSR occurs as students:
Parameter setting	<ul style="list-style-type: none"> Formulate a question with parameters and outcome variable Identify parameters and outcome variables for an experimental setup Know how to vary parameters and measure outcome variables 	<ul style="list-style-type: none"> Set parameters for an investigation based on the question and the setup Conduct test runs to build a mental model between the phenomenon under investigation and the data to be acquired Describe which variables will be varied and how
Data collection	<ul style="list-style-type: none"> Select a parameter set, carry out a run and measure the outcome variable Reflect on the quality of the run. Determine whether to rerun or stop data collection 	<ul style="list-style-type: none"> Make multiple runs purposefully to answer the question Determine when to rerun, modify a run, or stop data collection Select data for analysis
Data analysis	<ul style="list-style-type: none"> Calculate a way to characterize a run with a single value, in order to compare runs Create a parameter-outcome graph Use a time series graph to obtain an outcome value Identify patterns in a parameter-outcome graph Reflect on quality of data Answer the question using evidence generated from the investigation 	<ul style="list-style-type: none"> Calculate and incorporate outcome into analysis Create and explain parameter-outcome graphs Explain a point in the parameter space in connection to a time series graph of a run Recognize the shape and important features parameter-outcome graphs (linear, nonlinear, periodic, etc. or break points where the nature of shape changes) Identify and treat outliers Identify and treat noise in the data and noise sources Communicate conclusions using evidence

Methodology

InquirySpace (IS) Learning Environment

The IS environment works with both physical and simulated systems. Figure 1 shows the IS environment for a simulated spring/mass system. Students can conduct multiple experimental simulation runs by varying parameters such as gravity, spring constant, starting position, mass of the ball, and damping. When students finish a simulation run, they can view their data in a table and export the data for analysis after clicking the "Analyze" button. For instance, Figure 1 shows that students conducted four simulation runs by varying the

spring constant parameter. A column for the period of oscillation of the system was added in the table in Figure 1, and students inserted period values estimated from the time series graph (bottom left). In Figure 1, students created two graphs: (1) time vs. distance (i.e. time series graph) shown at the bottom left of the screenshot and (2) spring constant vs. period (i.e., parameter-outcome graph) shown at the bottom right.

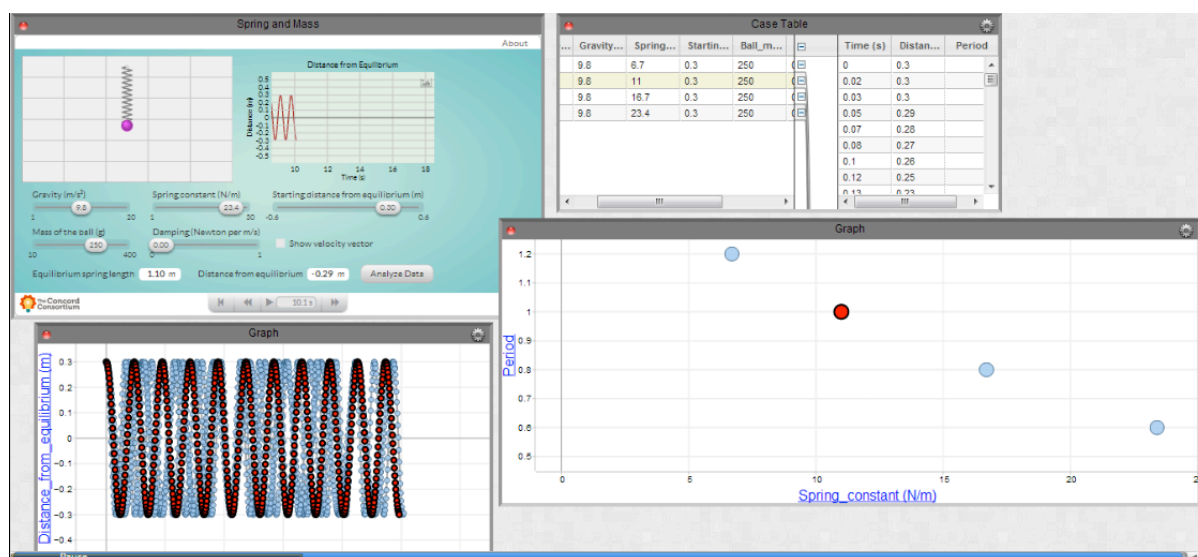


Figure 1. InquirySpace learning environment.

Data Collection and Analysis

The simulation-based spring/mass experimentation took place over two class periods in four ninth grade physics classrooms in an urban high school. Eighty-two students worked in 32 small groups. Each group chose its own question. 92% of the students spoke English as a first language, 52% were female; 20% self-reported to have used computers regularly for school learning. The school is a public charter high school where 98% of the students are from minority populations and 77% receive reduced or free lunch. According to the teacher, the IS curriculum sequence provided the opportunity for students to work in computerized lab experiments for the first time. The teacher was very structured and organized and made class performance expectations clear at the beginning of each class and was implementing the IS curriculum for the first time.

The IS curriculum sequence consisted of three investigations. First, all groups worked on the same hands-on investigation, using probes connected to the data collection and graphing environment, to answer how the period was affected by the mass of the ball. In the next investigation, groups were encouraged to explore how other variables in the physical spring/mass system impact period. The third investigation was conducted in the simulation-based spring/mass system. In this paper, we focus on the third investigation where groups had more choices for independent variables such as spring constant and gravity than were available to them with the physical spring/mass system. Each investigation was guided by accompanying worksheets and was carried out in the order of Explore, Plan, Create a Screencast Video about Plan, Experiment, Analyze, Explain, and Summarize conclusions on a Screencast Video. In this study, we focus on logging events recorded in the data analytics component of the IS environment and Screencast Videos that summarized conclusions of the third investigation. For the exploration stage, the worksheet said "Play with the model until you see what the model does to the spring." Throughout the worksheets, it was clear to students that they were investigating the impact of one variable on another and they were encouraged to go back to collect data if they were confused. The teacher told students to collect data from at least four simulation runs.

We used logged events to track students' parameter space navigation. A total of 5,277 events were logged for 31 student groups during the simulation-based mass/spring experiment. One group's logging data were lost. Fifty-three different syntaxes were used for logging events. Among the logged events, 62.6% were associated with simulations such as starting models and exporting models to the data analysis interface; 16.0% were related to data analysis involving tables, graphs, and data points; 17.9% involved creating and deleting components such as table, graph, and model; 2.8% were related to logging in and out. Each logged event was time-tagged, allowing duration and temporal order analyses. In order to investigate how students explore and analyze the parameter space to answer their investigative questions, we used the logging data to develop an event map for each student group over the period of the group's investigation and plotted key logging events such as beginning of exporting model results, creating period in a table, creating a time-series graph to estimate the period, and creating a parameter-outcome graph that students needed to use to answer their questions. These key logging events signal overall progression of their experimentation from exploration to refined

experimentation to data analysis. We also counted the number of simulation runs that were exported vs. not exported and compared students' parameter space navigation patterns between the two. Students' screencasts made at the end of the investigation were used to determine which independent and dependent variables students chose to investigate and how many and what values were chosen for the parameter-outcome graph. In order to determine whether PSR occurred as conceptualized in Table 1, we transcribed a 45 minute video of a student group's investigation and selected all segments where students reasoned for, within, and about the parameter space defined by their experimentation.

Findings

Student groups selected their own independent and dependent variables for their investigations. For the independent variable, 18 groups selected spring constant, 11 groups selected gravity, one group (T5) used mass, and one group (T9) used damping. All but one group selected period as a dependent variable. T9's investigation of the relationship between damping and the amount of decrease in amplitude was unique among all investigations. Students' investigations were carried out in the order of exploration, crude initial data collection, refined data collection, and data analysis. Identification of these stages in the logging data was facilitated by the presence of key logging events that signaled changes in students' experimentation focus. For example, students did not export their simulation results until they were serious about analyzing the data. Thus, *exported model* became an important logging event for students' moving from the data exploration stage to the data collection stage. Another important logging event was *created attribute period* because period was an outcome variable and needed to be estimated from the time-series graph. This event signaled students' moving from the crude initial data collection stage to the refined data collection stage. The syntax, *changed plot horizontal/vertical axis [variable name]* indicated that students were creating either a time series graph or a parameter-outcome graph. If students changed the horizontal plot axis with time and the vertical plot axis with distance from the starting position, they were making a time series graph to make measurements on period, the outcome variable for most student groups. If they assigned the horizontal axis to their independent variable and the vertical axis to their dependent variable, they were creating a parameter-outcome graph that was necessary evidence for their conclusion.

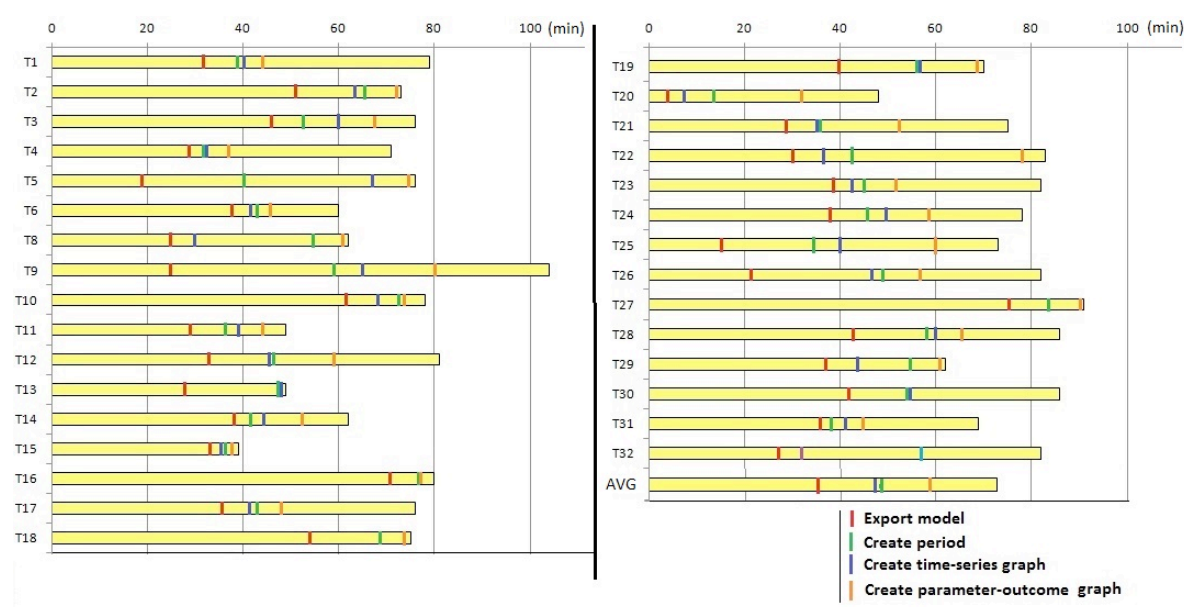


Figure 2. Event map for experimentation

As shown in Figure 2, on average, students' investigation lasted 72.5 minutes, ranging from 39 to 104 minutes. The four main logging events are marked in Figure 2. Creating a time-series graph ensued after several data runs were exported at the average of 47.3 minute mark and immediately followed by creating a column for period at the 48.6 minute mark. Some groups created the period column in the table before they created the time-series graph while other groups did it in the opposite order. All but one group (T17) were able to create a time series graph. On average, groups created their parameter-outcome graphs towards the end of their investigations at 59.8 minutes. Three groups (T13, T30, and T32) failed to create parameter-outcome graphs. Among the 28 groups who created the parameter-outcome graphs, two groups created incorrect graphs. For example, T5 created a graph of gravity vs. period, instead of mass vs. period, showing gravity was the same for all four simulation runs even though the group wanted to investigate the relationship between mass and period. T27 created a graph of gravity vs. elapsed time, rather than gravity vs. period. Nine groups were able to plot

parameter-outcome graphs on their first attempt while others needed to try two or more times to put the correct independent variable on the x-axis and the correct dependent variable on the y-axis. Groups that created the parameter-outcome graphs earlier used the remaining time to refine the graphs such as connecting points or making sense of what the graphs represented.

When we examined parameter values students chose to investigate, we discovered that students did not appear to export their data for analysis in the increasing or decreasing order. Rather, students were more concerned about having a set of parameter values that could cover a fairly good range of their chosen independent variable. There were several noticeable differences between exploration and formal data collection stages. According to Table 2, the average number of simulation runs student did not export was 45.5 while that of simulation runs students exported was 7.3. This indicates that students' exploration covered more of the parameter space than they actually analyzed. Moreover, in this exploration stage, students tested all available independent variables before settling into one independent variable of their choice. This indicates students' interest and need for tinkering with the simulation model as recognized in other studies (Berland, Martin, Benton, Smith, & Davis, 2013). All student groups plotted four points in their parameter-outcome graphs as suggested by their teacher. However, many groups exported more than 4 simulation run results for data analysis. Our further inspection of this discrepancy indicates that (1) some groups had to repeat their runs the following day for analysis because they did not save the data, (2) some groups ran simulation runs with the same set of parameters multiple times and exported them, or (3) other groups refined values for intervals between parameter values to be even. Table 2 lists five parameters with ranges that were manipulated by students. We considered the bottom and the top 10% as extreme value ranges. We obtained percent frequencies for these parameters that were set in that extreme ranges for exported vs. not-exported simulation runs. When students were just exploring they used more extreme values than when they were collecting data to answer their questions. One exception to this rule was damping. Students set the damping parameter "0" for most of their simulation runs for data collection because, even though damping does not affect the period, it does decrease the amplitude over time, making estimating the period from the time-series graph more difficult.

Table 2. Navigated parameter space.

Parameters	Value range	Exported parameter sets (n = 235)		Un-exported parameter sets (n = 1,457)	
		Middle (%)	Extreme (%)	Middle (%)	Extreme (%)
Gravity (m/s^2)	0.8 - 19.8	77.5	22.5	40.0	60.0
Spring constant (N/m)	1.0 - 30.0	76.6	23.4	45.1	54.9
Amplitude (m)	-0.6 - 0.6	91.5	8.5	57.9	42.1
Mass (g)	10 - 400	82.1	17.9	48.3	51.7
Damping (N per m/s)	0.0 - 1.0	3.8	96.2	14.3	85.7

Example: A Group's Parameter Space Navigation

In this section, we describe the types of PSR that occurred when a group of students (Group T1) explored the parameter space. Group T1 consisted of one male and two female students. Figure 3 shows a time-lapsed event map for the first 45 minutes of their 79-minute long investigation. The curriculum investigation sequence consisted of exploration, planning, creating a Screencast plan video, data collection, and data analysis. T1 chose a question of how spring constant affected period and was able to complete four main key logging events during this time period. T1 conducted 25 simulation runs for exploration and 10 runs for data collection. They chose the spring constant values of 8, 10, 12, and 16 N/m. From the video transcripts, we identified six occasions where students were actively engaged in PSR as envisioned in Table 1. While not all features of PSR listed in Table 1 occurred, students' discourse shows that PSR played a central role in various stages of their investigation.

Making a Hypothesis

An important aspect of PSR is to setup a parameter and what it means in order to develop a hypothesis. After exploration, students wanted to come up with a hypothesis. In formulating a hypothesis involving a spring constant, students were confused as to what a spring constant meant. The excerpts below indicate students' conceptual clarification of the spring constant variable and how it would affect the period.

- S3: We are doing spring constant, right guys?
 S2: Less mass it is, the more the spring to be constant?
 S1: What?
 S4: I don't know.
 S1: so, the constant, he means this right?

- S3: Is that constant?
[Unsure what spring constant is, students attempting to get the teacher's attention]
- S2: We thought the greater the spring constant the faster it would go.
- Teacher: So, the higher the spring constant, and what?
- S2: the time...
- Teacher: What about the time? longer time or shorter time?
- S2: I don't get it.
- Teacher: Do you know what spring constant is?
- S2: No.
- Teacher: Do you know something's really stiff.
- S2: Right!
- Teacher: The spring constant is a measure of the stiffness ...
- S2: So, if the spring constant is high, then the faster the period to finish?
- Teacher: Maybe...that's the hypothesis you need to figure out.
- S1: You got the new hypothesis?
- S2: The higher the spring constant, the faster the period.

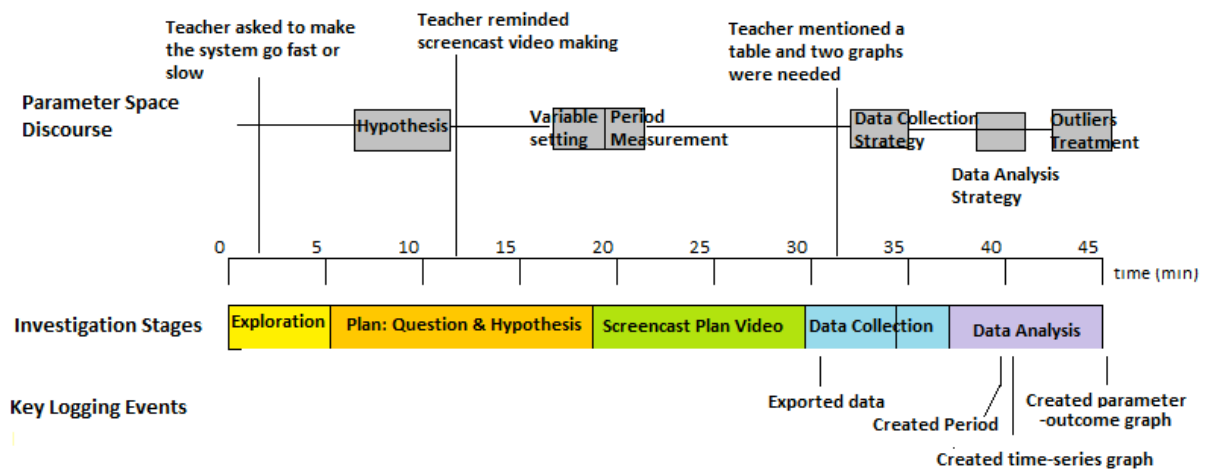


Figure 3. A student group's detailed event map

Setting Variables

About 12 minutes into the class, the teacher reminded students to make a Screencast video communicating their experimental plan. The teacher particularly asked students to focus on choosing a parameter to vary at least four times. This teacher's request had students think about how to manipulate parameter space using the simulation.

- Teacher: All you need to do is what variable to vary and what are the four numbers you are going to change variables to, to test it.
- S2: How am I going to change numbers?
- Teacher: Move the bar before you chose random numbers....like evenly spaced out
- S1: Which bar?
- Teacher: What are you doing? Gravity or spring constant?
- S2: spring constant.
- Teacher: spring constant bar....choose four numbers.
- S2: don't do odd numbers, do even numbers that are far apart.
[T1 chose 8, 10, 12, and 16 N/m for their experiments.]

Measuring Periods

Now, T1 had a hypothesis of "the higher the spring constant, the lower the period" and chose four spring constant values to test. As they started recording a Screencast video about their plan, they recognized that they did not know how to measure the period:

- S2: Our question is how does spring constant affect the period. The two variables that we are using today are spring constant and period.
- S1: OK, how do we measure the period?
[At first, confused...asking around...from distance "it's the difference between crests"....].
- Teacher: Remember when you drew a graph...
- S1: OK, then we need to find out the difference?

- S2: OK, we measure the spring constant and period. We do period by subtracting the distance between the two times, but we haven't started yet so we don't have it.

Data Collection Strategies

Immediately after their recording of the Screencast video, T1 started collecting data by changing spring constants from 8, 10, 12, to 16 N/m. Then, the teacher mentioned that a table and two graphs (time series and parameter-outcome graphs) were necessary. Then, T1 realized that they had not been saving data for analysis:

Teacher: You need to create a chart and two graphs.

S2: Oh.

Teacher: I see nothing related to charts. You have to analyze data.

S2: Do we need to start all over?

Teacher: yes. Do you know what we are saying? You need to push that button to send to the chart.

S2: Then, we got to do one by one again, then.

Data Analysis Strategies

T1 had to redo the same simulation runs so that they could save their data for analysis. This time, after each time they varied the spring constant value, they exported the data to the chart. The four simulation runs went fast. T1 then wanted to create a graph and first went for a parameter-space graph. After setting the x-axis as spring constant, they realized that they did not have period in their data. This realization allowed T1 to work on obtaining the dependent variable from time-series graphs.

S1: I need to get period. Where is the other graph [*time-series graph*] I was looking for.

[S1 deleted spring constant x-axis and changed it to time series graph for the spring constant 8 one]

S1: Is time x or y?

S2: Time? x.

[S1 then put the distance from the equilibrium variable on the y-axis to create the time series graph where students could obtain period from subtracting time difference between two crests]

Detecting and Treating Outliers

For the period of the spring constant of 8 N/m, they calculated period as 1.1 seconds. For the period of the spring constant of 10, 12, and 16 N/m, they calculated as 9.9, 9.0, and 7.9 seconds. The first period estimate was a correct one. However, students made mistakes for the other three. A student surmised that the next three periods were too big because she thought, according to their hypothesis, that the periods should be getting smaller. However, her observation went unnoticed. This subtraction mistake was caught when the group plotted their first parameter-outcome graph and immediately recognized something was not right:

S1: Does it supposed to look like that?

[students were examining the first point because it looks an outlier from the time-series graph. ...then went to the second point.]

S2: This is .9

[after examining the corresponding points in the time-series graph]

S1: That's not 9.9. It is .99

[They checked all the other points so that their final periods were 1.1, .99, .90, .79 for spring constants of 8, 10, 12, and 16 N/m. This gave a reasonable spring constant vs. period graph]

Conclusion and Significance and Connection to the Conference Theme

The time progression of students' investigations generally followed directions on the accompanying worksheets. This means that the suggested curriculum sequence was able to accommodate students' needs in simulation-based experimentation. However, we must emphasize that the teacher played an important role in moving and refocusing students' attention on important parts of their investigations, and clarifying questions and confusions students might have had. Students' parameter-value choices were made purposefully. For the exploration, students used extreme values across all available parameters to quickly grasp the general tendency of how each parameter affected the virtual spring/mass system. For the data collection purpose, prompted by their teacher, they chose a parameter to investigate and selected the range of values that could roughly determine the overall shape of the relationship between the parameter and the system outcome variable. Conversations within a student group centered on various aspects of parameter space reasoning: selection, definition, and measurement of independent and dependent variables, data collection and analysis strategies, and outlier treatment. In these efforts, visualizations of the raw or analyzed data appeared to be critical in prompting students' immediate responses, such as forming a hypothesis, modifying period estimates, and recognizing the next actions to take.

As shown in this study, during experimentation, students reasoned exclusively with the parameter space, rather than with mechanics of the spring and mass system. Our currently ongoing research efforts are focused on developing logging data analytics for teachers and researchers to use, defining converging evidence of learning that is unique to student-generated experimentation from multiple sources, and defining and validating PSR.

Inquiry-based investigations in science class have been promoted as an important pedagogical approach in science education reforms. We characterized and illustrated student reasoning associated with the parameter space that defines individual experimentation runs in the context of simulations. However, PSR—the ability to think about the simulation runs as simply data points at a higher-level—can get confounded with other “inquiry skills” such as the ability to ask “interesting” questions, the ability to come up with experimental designs likely to shed light on a question, the ability to control potentially confounding variables, and the ability to reliably distinguish signal from noise. The science education literature refers to “systematicity” (often mindlessly recommending the “one size fits all” strategy of varying only one parameter at a time) or suggests as a normative standard that students exhaustively cover every region in the parameter space no matter what question they are trying to answer. Such supposed universal markers for PSR are clearly too simplistic to capture the richness of the phenomena under study. As more and more learning technologies are integrated to form powerful learning environments, it becomes necessary to reconceptualize student reasoning in a more nuanced and multi-faceted manner. We believe that a learning environment such as InquirySpace can provide opportunities for students to be *learning and thus becoming in practice* while also enabling designers and researchers to study resulting student learning to a greater extent. In this study, we outlined what PSR might mean in a simulation-based learning environment and illustrated that students were indeed engaged in PSR frequently throughout planning, exploration, data collection, and analysis stages.

References

- Allchin, D. (2012). Teaching the nature of science through scientific errors. *Science Education*, 96(5), 904-926.
- Amsel, E., & Brock, S. (1996). The development of evidence evaluation skills. *Cognitive Development*, 11(523-550).
- Berland, M., Martin, T., Benton, T., Smith, C. P., & Davis, D. (2013). Using learning analytics to understand the learning pathways of novice programmers. *The Journal of the Learning Sciences*, 22, 564-599.
- 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.
- Hackling, M. W., & Garnett, P. J. (1992). Expert-novice differences in science investigation skills. *Research in Science Education*, 22, 170-177.
- Kanari, Z., & Millar, R. (2004). Reasoning from data: How students collect and interpret data in science investigations. *Journal of Research in Science Teaching*, 41(7), 748-769.
- Lubben, F., Campbell, B., Buffler, A., & Allie, S. (2001). Point and set reasoning in practical science measurement by entering university freshman. *Science Education*, 85(4), 311-327.
- Masnick, A. M., & Klahr, D. (2003). Error matters: An initial exploration of elementary school children's understanding of experimental error. *Journal of Cognition and Development*, 4(1), 67-98.
- Masnick, A. M., Klahr, D., & Morris, B. J. (2007). Separating signal from noise: Children's understanding of error and variability in experimental outcomes. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 3-26). New York: Lawrence Erlbaum Associates.
- McElhaney, K. W., & Linn, M. C. (2011). Investigations of a complex, realistic task: Intentional, unsystematic, and exhaustive experimenters. *Journal of Research in Science Teaching*, 48(7), 745-770.
- National Research Council (1996). *National science education standards*. Washington, DC: National Academy Press.
- Petrisino, A., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical Thinking and Learning*, 5, 131-156.
- Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, 32, 102-119.
- 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.

Acknowledgments

This material is based upon work supported by the National Science Foundation under grant No.1147621. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors acknowledge Edmund Hazzard who developed IS activities, William Finzer and Daniel Damelin who developed the simulation and logging technology used in this study, and students and teachers who participated in this study.

Learner Alignment with Expansive Framing as a Driver of Transfer

Diane P. Lam, Adam Mendelson, Xenia S. Meyer, Lloyd Goldwasser,
University of California, Berkeley, 4646 Tolman Hall, Berkeley, CA
dianelam@berkeley.edu, amendelson@berkeley.edu, xenia.meyer@berkeley.edu, goldwasser@berkeley.edu

Abstract: Recent theoretical perspectives on transfer highlight the active role of learners in transferring knowledge. In this study, we focus on the learner by exploring the possibility that when learners align their actions and beliefs with expansive framing, they are more likely to transfer knowledge from one context to another. We define expansive framing as the perspective that learning is connected to other times, places, participants, and topics, and that the role of students is to generate and share knowledge. We examine the degree to which high school biology students aligned themselves with this perspective. Using correlational and discourse analyses of students' performance on transfer assessments in relation to their perceptions of the expansive relevance of their class, we present evidence of a relationship between alignment with expansive framing and transfer. Students who perceived greater temporal relevance across past experiences, current learning, and future opportunities displayed the highest levels of transfer.

Introduction

Research on transfer of learning has a century-long history, dominated primarily by laboratory experiments and a cognitive perspective that views transfer as dependent on the characteristics of learners' mental representations (e.g., Gentner, 1983; Gick & Holyoak, 1983; Reeves & Weisberg, 1994; Ross, 1984). In recent years, a new set of perspectives has emerged including actor-oriented transfer (Lobato, 2012), preparation for future learning (Bransford & Schwartz, 1999), consequential transitions (Beach, 1999), dispositional and motivational views (e.g., Belenky & Nokes-Malach, 2012; Perkins & Salomon, 2012), and expansive framing (Engle, 2006; Engle, Lam, Meyer, & Nix, 2012; Engle, Nguyen, & Mendelson, 2011). A key feature of these perspectives is a greater focus on the role of the learner in transfer, and the importance of more than purely cognitive factors.

This paper focuses on one of these contemporary approaches: expansive framing, as an explanation for transfer. The underlying premise of this approach is that transfer is facilitated when learners perceive instruction as part of larger intellectual conversations that extend across contexts, and perceive themselves as capable of and accountable for contributing to those conversations (Engle, 2006; Engle et al., 2011). It is hypothesized that instruction that is expansively framed promotes these learners' perceptions and beliefs of intercontextuality between learning and potential transfer opportunities, which then triggers social and cognitive mechanisms that lead to transfer (Engle et al., 2012). Our prior research has focused primarily on expansive framing from the instructor's perspective and how he or she can frame a learning environment to promote transfer (Engle, 2006; Engle et al., 2011). In this paper, we turn our attention away from the discursive moves through which instructors can expansively frame instruction, and focus directly on the perceptions and beliefs of learners in an expansively framed biology classroom. This shift in focus is motivated by the observation that a teacher's pedagogical moves matter only to the extent that these moves are perceived and taken up by students and incorporated into their practices.

We investigated the degree to which students in this classroom *aligned* their own actions and beliefs with expansive framing, and compared their level of alignment with their performance on transfer measures. We found converging evidence of relationships between student alignment with expansive framing and transfer, supporting the premise that learners' perceptions and beliefs of the extended relevance of learning are fundamental for transfer. In our discussion we consider how the instructional approach used by the teacher may have influenced those crucial perceptions and beliefs.

Background and Theoretical Framework

An Emerging Focus on the Learner in Contemporary Transfer Research

The role of learners' perceptions and beliefs in transfer has been gaining increasing attention from transfer researchers in recent years. Perhaps the most fully developed learner-focused perspective on transfer is Joanne Lobato's (2012) *actor-oriented transfer* (AOT). Dissatisfied with the limits of traditional transfer measures, Lobato aims to "understand the interpretative nature of the connections that people construct between learning and transfer situations" independently of whether or not those connections are expected or deemed appropriate by researchers (p. 239). From this perspective, analyses of transfer provide insight about other ways in which students draw on prior learning.

James' (2008) study on transfer between writing tasks designed to be similar or dissimilar illustrates the benefits of taking an AOT approach. Student interviews showed that students considered many dimensions of the tasks that the researcher had not considered, and that different students identified different similarities and differences between tasks. Independently of which of two tasks each individual student completed, there were greater levels of transfer by students who perceived their assigned task to be similar to instruction than by students who perceived their task as different. This finding suggests that transfer depends on subjective perceptions of similarity more so than researcher-assumed objective similarities (Day & Goldstone, 2012; Pea, 1987).

Perkins and Salomon (2012) call for a focus on the learner by considering learner motivations and dispositions in relation to transfer. They claim that the bulk of transfer research focuses on "learners' ability to make the [researcher's] desired connection" while ignoring the learners' "motivation or disposition to do so" (p. 253). Belenky and Nokes-Malach (2012) illustrate the importance of these factors in their investigation of achievement goals in relation to transfer. Subjects' achievement goals were measured before and while engaging in different activities about statistical concepts. The researchers found that students who displayed higher levels of motivation to develop long-term competence and understanding of statistics, and higher senses of personal relevance and contribution while engaging in activities displayed higher levels of transfer. Together this body of work suggests that transfer is highly dependent on the individual learner, and his or her orientation towards what is being learned and transferred.

The Learner in Expansive Framing

We use the term *framing* to refer to the discursive moves through which interlocutors propose and align themselves with expectations regarding the nature of the interaction in which they are engaged (Bateson, 1972; Goffman, 1974; Tannen, 1993). Framing consists of primarily linguistic cues that index schema of types of interactions (Gumperz, 1982; Ochs, 1996; Tannen, 1993) and activate cognitive resources associated with those schema (Hammer, Elby, Scherr, & Redish, 2005). We refer to framing as *expansive* when it proposes *intercontextuality* (e.g., Floriani, 1994; Gee & Green, 1998), or connections between the current interaction and others (Engle, 2006; Engle et al., 2012; Engle et al., 2011). For example, a classroom teacher can expansively frame a lesson by presenting it as an opportunity for students to take on knowledgeable roles within communities in which they may participate throughout their lives. Doing so expands the scope of the lesson by proposing its relevance across time, space, and social configurations. It is thought that transfer is facilitated when learners perceive and believe in this expanded relevance (Brown, 1989; Greeno, Smith, & Moore, 1993; Laboratory of Comparative Human Cognition, 1983; Pea, 1987). When learners hold these perceptions and beliefs, we say that they are *aligned* with expansive framing, and we present data in this paper that indicate that this alignment is indeed related to transfer.

Table 1. Intercontextualities created by expansive framing and mechanisms for how each may lead to transfer.

Contextual aspect	Proposed intercontextuality	Transfer mechanism
<i>Time</i> : When is the lesson happening?	The lesson is part of an ongoing activity that started in the past and will continue into the future.	Students draw on prior knowledge during the lesson. They learn content expecting to be able to use it in the future.
<i>Place</i> : Where is it happening?	The lesson is relevant outside the classroom (e.g. to rest of school, homes, local community, places around the world, other institutions, etc.).	Students draw on experiences from other places during the lesson. They learn current content expecting it to be applicable in other places.
<i>Participants</i> : Who is participating?	The lesson is relevant to a broad community that extends throughout and beyond the classroom.	Students consider the relevance of interactions with others during the lesson. They learn current content expecting it to be of interest to others.
<i>Topics</i> : What is the topical scope of the lesson?	The lesson is part of larger and interrelated units, topics, and subject areas.	Students see connections between the lesson and other topics they have studied and will study.
<i>Roles</i> : How are learners positioned intellectually?	Learners are authors who are responsible for developing, sharing, and defending their own ideas.	Students feel accountable for using and sharing the ideas they author. They may also adopt the practice of generating and adapting ideas to attempt to solve novel problems.

Empirical work on expansive framing has identified five different aspects of learning contexts that instructors can frame expansively: time, place, participants, topics, and roles (Engle, 2006; Engle et al., 2011).

More recently, Engle and her colleagues (2012) proposed several explanations for how expansive framing can promote transfer. Drawing on this body of research, the following table presents the intercontextual links that are proposed when each of the five aspects is expansively framed, as well as explanations of how learners' alignment with these proposals is believed to promote transfer (Table 1).

Methods

In this paper, we investigate the idea that what ultimately matters for transfer is the degree to which learners align themselves with expansive framing. We focus on relationships between learners' alignment with expansive framing, captured through surveys and interviews, and their performance on learning and transfer assessments. We use correlational analyses of surveys and transfer assessments to investigate the relationship between learner alignment with expansive framing and transfer. Through qualitative analysis of interviews, we explore patterns in the expressions of alignment of students who demonstrated different levels of transfer.

Research Context: An Expansively Framed Biology Class

The data presented in this paper come from a larger, mixed-methods study on framing and transfer in the classroom of an experienced and award winning high school biology teacher who we refer to as Mr. Kent. We first began observing Mr. Kent's classroom in 2008, at which time we recognized many key aspects of expansive framing in his instruction (Engle, Meyer, Clark, White, & Mendelson, 2010). Through our prolonged collaboration with Mr. Kent, we believe that the degree to which his instruction embodies our framework has increased over time (Engle et al., 2012). This paper focuses specifically on Mr. Kent's second period biology class consisting of 32 freshmen and sophomores (22 girls, 10 boys) at a San Francisco Bay Area high school during the 2010-2011 school year.

Data Collection and Analysis

We draw from three data sources in this paper: (1) a series of written assessments, to measure student learning and transfer of knowledge related to osmosis and diffusion; (2) student surveys, to measure student alignment with expansive framing as well as their recognition of Mr. Kent's use of expansive framing in his instruction; and (3) student interviews, to capture more detailed student expressions of alignment with expansive framing.

Assessments of Learning and Transfer

Our analysis of transfer for this paper is based on four written assessments about osmosis and diffusion, a concept that is fairly difficult for students to master (Christianson & Fisher, 1999) and therefore to transfer (Bransford, Brown, & Cocking, 2000). The pre- and post-assessments were administered before and after the instructional unit on osmosis and diffusion and were identical for the purpose of recording changes in student knowledge of the focal content. The initial transfer assessment was administered several months after the post-test, and the follow-up assessment (which measured both transfer and retention) was administered nearly two years later, in 2013 (see Figure 1). Seventeen of the original 32 students from the focal class completed the follow-up assessment.

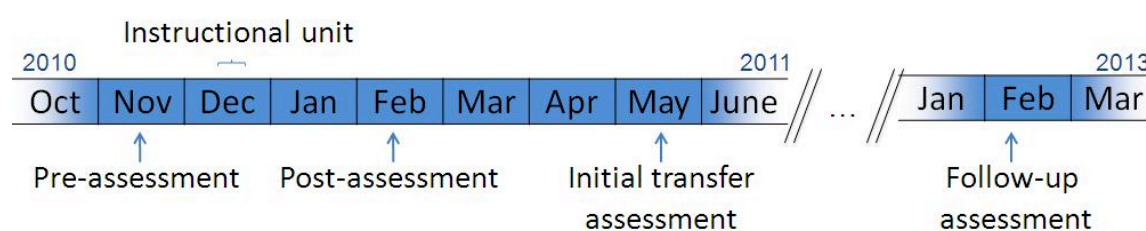


Figure 1. Timeline of written assessments in relation to the instructional unit on osmosis and diffusion.

Questions on the initial and follow-up transfer assessments presented students with a variety of scenarios about concentration gradients that were designed to be more or less analogous to what students had already learned in class (Gick & Holyoak, 1983). Some questions had a similar structure and referred to content that was discussed during the instructional unit. Other questions included content areas not yet addressed in class, and required students to recognize the role of concentration gradients in the problem before being able to solve it. Still others posed a different kind of task, such as drawing a diagram rather than selecting from multiple-choice answers.

Our analysis of transfer was based on knowledge elements related to osmosis and diffusion that were reflected in students' responses to assessment questions. We use the term "knowledge element" (KE) to refer to an individual fact or proposition that forms part of the normative understanding of a larger conceptual principle (Engle et al., 2011; Renkema 2004; van Dijk & Kintsch 1983). In our coding scheme, we identified seven of

KEs that together represented the targeted level of understanding of osmosis and diffusion. Students' relative mastery and application of KEs was assessed through their identification of the outcome of a problem scenario by selecting a multiple-choice answer, and their explanation of that outcome in a written response. Two of the authors coded the responses independently, and compared codings to measure interrater reliability. All differences were resolved by mutual agreement. Interrater reliability was 0.937, giving a Cohen's Kappa of 0.861.

As a baseline measure of each student's knowledge about osmosis and diffusion that was available for transfer, we adopted the total set of all KEs that were coded for each student on *any* of the questions of the post-assessment. That is, if a student's response was coded for a particular KE on *any* item on the post-test, that student was identified as having that KE available to transfer. We were then able to compare this baseline with students' answers to each of the questions on the initial and follow-up transfer assessments. We calculated each individual student's transfer by question as the number of KEs received for the particular transfer question divided by the number of total post-assessment KEs. The student's initial transfer score was then computed as the average of the nine ratios. The formula used to calculate transfer for each individual was

$$\text{Transfer score} = \frac{\text{Transfer (average of Knowledge Elements across questions)}}{\text{Post-test (all Knowledge Elements achieved)}}$$

Because the denominator was the same throughout, this overall transfer score also equaled the average number of KEs achieved throughout the transfer assessment divided by the number of post-assessment KEs. In this way, the computed transfer scores are the average number of KEs students applied to new scenarios in relation to what they displayed on the post-test.

Student Surveys

Early in the spring semester (February), Mr. Kent's students completed an online survey in class, designed to capture student recognition of their teacher's discursive moves related to expansive framing as well as students' own alignment with expansive framing across each of the five contextual aspects described earlier. In this paper, we focus primarily on responses to survey questions related to alignment because our framework suggests that alignment is the mechanism for transfer. The survey was comprised of 36 Likert-style questions and showed high internal consistency both within and across all five contextual aspects ($\alpha \geq 0.795$ in all cases). For each student, we created scores for each of the five aspects by averaging the responses to questions that addressed each aspect. These per-aspect framing scores were then compared with individual students' transfer scores through correlation analyses.

Student Interviews

Towards the end of the school year, we conducted individual interviews with 18 students to ask them to elaborate on their survey responses. These students were selected to constitute a representative sample of the class, based on their range of performance on our written assessments, survey responses, and demographics. Interviews were audio recorded and transcribed. We used discourse analysis to code interview transcripts for expressions of alignment with expansive framing and self-reports of transfer.

Findings

In this section we provide evidence of student transfer, and then explore possible relationships between transfer and student alignment with expansive framing as captured through surveys and interviews.

Evidence of Transfer in Mr. Kent's Class

Using the approach described above, we first compared student performance on the initial transfer assessment with their performance on the post-test. The average initial transfer score for the class was 0.43 ($SD=0.25$; $N=28$), indicating that on average students transferred 43% of the KEs that they had displayed on the post-test. This score demonstrates that transfer did, in fact, occur across the post-assessment and initial transfer test. We then used the same approach to compare results of the follow-up transfer question with the post-test. The average follow-up transfer score was 0.45 ($SD=.43$; $N=16$), or 45% of KEs displayed on the post-test. On the retention question students averaged 54% ($SD=.43$, $N=16$) of the knowledge elements they had displayed on the post-test.

Relationships between Transfer and Alignment with Expansive Framing

Student Surveys

Table 2 presents class-wide means and standard deviations for student alignment with expansive framing of each of the five contextual aspects in our framework. These data indicate that students perceived the expanded relevance of their learning across time, place, and topics to a degree between "sometimes" and "often." They

perceived expanded relevance across participants, and themselves as authors of biology knowledge to a degree slightly less frequent than “sometimes.”

Table 2: Results from student survey measuring student alignment with expansive framing

	Time	Place	Participants	Role	Topics	Average
Alignment	3.72 (0.72)	3.51 (0.73)	2.87 (0.70)	2.92 (0.75)	3.45 (0.99)	3.29 (0.55)

After calculating framing scores for each student for each of the five contextual aspects, we ran correlations between these scores and student scores on our initial transfer assessment and our follow-up assessment. Correlation analyses revealed alignment with expansive framing of time, per our survey, as having the strongest relationship with the results of our transfer assessments. We found strong and statistically significant correlations between student alignment with the expansive framing of time and initial transfer scores ($r[26]=0.376$, $p<0.05$), follow-up transfer score ($r[14]=0.589$, $p<0.01$), and retention scores ($r[14]=0.444$, $p<0.05$). In other words, those students who indicated on our survey that they perceived higher degrees of temporal relevance of their biology class, both in terms of biology being related to what they already knew and in terms of the future applicability of biology content, were the same students who had higher scores on our assessments of transfer.

We found one additional statistically significant correlation in these data, between alignment with expansive framing of role (i.e., perceiving oneself as an author of biology knowledge) and the retention question on the follow-up assessment ($r[14]=0.440$; $p<0.05$). That is, those students who indicated on our survey that they most perceived themselves as playing an active and accountable role in authoring biology knowledge were the same students who displayed the highest level of retaining what they had learned about osmosis and diffusion after nearly two years.

Student Interviews

In this section we present patterns from our interview data related to time, authorship, and transfer because of the importance and interrelatedness of these factors as suggested by our correlation analysis. Our goals in this section are to complement our survey data with concrete examples of students' expressions of alignment with expansive framing; and to complement our statistical analysis with descriptions of the patterns in our interview data. Here, we present our findings by three themes: 1) alignment with the expansive framing of time, 2) alignment with the expansive framing of role, and 3) self-reports of transfer.

With regard to perceiving the long-term temporal relevance of biology content, most students expressed some alignment with expansive framing, but there were some salient differences by level of performance on the transfer assessment. One difference was that students who displayed higher levels of transfer on the assessment spoke about the future relevance of biology in much more concrete terms than students who had lower transfer scores. For example, high transfer students were quite definitive about their anticipated future use of biology content in college, making comments such as, “for college, yeah, I’m gonna use this stuff” (Lindy). In contrast, while students with average to low transfer scores also spoke about the applicability of biology content to college, they tended to give hedged or conditional responses, such as: “Biology will probably be involved” (Deena); “I might use that later on in the future, maybe” (Bianca); and “If I go to college...it’ll help me” (Pedro).

In addition to the specificity with which they anticipated the future relevance of biology, some high transfer students also expressed the perception of continuity of course content throughout the school year. This sentiment was best illustrated by Elaine: “The only time it’s really brand new is in the beginning of the year. Like, we’ll um, we always refer back to certain stuff. (...) We always just kinda add something new, but it builds on what we already know.” Dorthie echoed this sentiment by referring to learning about cells as “just the beginning” of biology-related knowledge. We did not find similar comments in the interviews of students with lower transfer scores.

What we did find from some of the lowest transfer students, however, were examples of questioning and even rejecting the future relevance of biology. For example, Carolina expressed that she was unlikely to pursue a future for which biology would be relevant: “Like, be a scientist? Mmm, no I don’t think so...I don’t think it’s for me.” She later added “I think there are some thing [from class] we don’t need to know later on in life. Like I don’t think it’ll be useful for some of us students that aren’t gonna be needing that stuff later on.”

In terms of perceiving themselves as taking on roles as authors of biology knowledge, most of the students we interviewed expressed a belief in the value of sharing ideas with their classmates, a practice that was very much emphasized by Mr. Kent. However, their explanations of the value of this practice varied. Mid to low transfer students saw sharing their ideas primarily as a way to “learn from mistakes” (e.g., Luis, Deena, Carolina), an idea also emphasized by Mr. Kent. In contrast, high transfer students depicted interacting with their classmates as an active process of collaborative knowledge construction. According to Dorthie, sharing ideas with her classmates meant “if they have a different view they tell me why and (...) You can combine your

ideas to make a better answer.” Baara added, “We actually have to discuss [questions] amongst the whole class and we have to come up with an outcome on our own. (...) It’s not like we’re really dependent on our teacher all the time. We become more able to think for ourselves.” It is worth noting the way in which these high transfer students’ commonly used first person pronouns and possessives to position themselves as owning and taking responsibility for their ideas. Elaine used similar language when referring to “my understanding” of a biology-related phenomenon, as did Lindy in anticipating that she would “refer back to my notes” when taking future science classes.

The high transfer students also positioned themselves as capable of productively disseminating those ideas to others, both inside and out of class. Lindy anticipated “help[ing] my sister with biology,” and Baara believed that by sharing what she knew in the classroom she could “help out the class.” With the exception of Deena, who also claimed, “I could teach other people,” the mid and low transfer students did not make similar comments about knowledge dissemination. One low transfer student, Carolina, actually rejected some aspects of authorship, stating, “I don’t think that if you don’t talk about ideas -- like if you don’t say ideas you don’t go anywhere. I’m pretty sure you’ll go somewhere in life.”

Self-reports of transfer were relatively scarce in the student interviews, but even so some differences were observed between higher and lower performing students on the transfer assessment. First, such reports were slightly more common from high transfer students than from others. Second, and perhaps more importantly, was a qualitative difference: high transfer students reported actually applying biology knowledge to novel situations, while mid and low transfer students reported retelling facts learned in class. Of the high transfer students, Lindy had by far the highest number of self-reports of transfer. As a softball player, she reported, “[Mr. Kent] says ‘oh you need a lot of carbohydrates and nutrients and stuff, for, if you’re athletic’ and (...) I look online and try to find out what’s the daily stuff if you do sports and stuff. And I just look on that and see how many servings I need and everything.” Here we see Lindy’s practical application of what she has learned about nutrition in her biology class to her needs as an athlete outside of class. This particular report also exemplifies transfer in the form of preparation for future learning (Bransford & Schwartz, 1999) as Lindy used what she had learned about nutrition to search online for specific answers to her needs.

The only low transfer student we interviewed who made self-reports of transfer was Pedro, who, like Lindy, reported some transfer to his experiences as an athlete. However, in contrast to Lindy’s practical application, Pedro’s reports were limited to retelling facts and terms. For example, with teammates who also studied biology, he reported answering their definitional questions such as “What are carbohydrates?” saying, “I’ll just try to sound smart and like, tell them all of these terms I learned.” Pedro’s transfer is therefore limited to repeating facts that has learned, but does not necessarily display any sort of application of biology knowledge to solving novel problems.

In summary, our interview analysis revealed several characteristics that seem to generalize fairly well across those interviewees who displayed high levels of transfer on the assessment. These students not only perceived biology as generally relevant to the future, but actually imagined themselves in concrete future situations in which they anticipated using that content. They also perceived biology instruction as continuous across the school year, as earlier lessons connected to later ones. These students perceived themselves as active participants in the construction of biology-related knowledge. They took ownership of their ideas and productively shared them for the benefit of others. Finally, when they reported transferring biology knowledge to other aspects of their lives, they did so in ways that implied practical application to meet actual needs.

In contrast, lower transfer students either expressed the future relevance of biology content in much more conditional terms, or, in some cases, actually rejected that future relevance. They shared ideas with classmates in order to learn from mistakes, but did not position themselves as actively constructing and disseminating knowledge. Very few of them reported transferring their knowledge of biology, and those reports were limited to retelling facts more so than solving novel problems.

Discussion and Conclusion

Our analyses of students’ surveys and interviews in relation to their performance on our measures of transfer provide converging evidence of a relationship between alignment with expansive framing and transfer. Our strongest finding is that those students who reported perceiving more temporal relevance between past experiences, current learning, and future opportunities for application were the same students who displayed the highest levels of transfer. This finding is consistent with prior empirical and theoretical work that posits perceived relationships being central to transfer (e.g., Day & Goldstone, 2012; James, 2008; Lobato, 2012; Pea, 1987). Additionally we found higher levels of transfer, although perhaps only in the form of retention, from students who indicated that they perceived themselves as taking on active role in constructing and disseminating biology-related knowledge. This finding resonates with work that links transfer to students’ perceptions of autonomy and authority with regard to learning and knowledge (e.g., Belenky & Nokes-Malach, 2012; Perkins & Salomon, 2012).

Our findings also indicate that Mr. Kent's attempts to expansively frame his instruction contributed to his students' beliefs and perceptions of expansive relevance. While we found no statistically significant correlations between students' recognition of Mr. Kent's expansive framing and their performance on our transfer assessment, we did find a strong correlation between their recognition and their alignment with expansive framing ($r[28]=0.558$, $p<0.01$). The relationship between their recognition and their reported perceptions of temporal relevance was even stronger ($r[29]=0.632$, $p<0.001$). In other words, those students who showed the highest degree of alignment with the expansive framing of time – a factor that we've shown to be related to transfer – were the same students who were most aware of Mr. Kent's discursive positioning of the expansive relevance of learning in his classroom. This finding is consistent with suggestions that while learner perceptions may be at the heart of transfer, instruction can play a role in shaping those perceptions (e.g., Belenky & Nokes-Malach, 2012; Lobato, Rhodehamel, & Hohensee, 2012). Future research could better investigate the relationship between instruction and learner alignment with expansive framing by including pre- and post-measures of alignment, and by comparing alignment of students in classrooms that are more and less expansively framed by their instructors. Additionally, video analysis of classroom behaviors could provide a more intimate understanding of the degrees to which students discursively took up Mr. Kent's proposed framing (Clark, 1996). If students were to discursively reciprocate their instructor's expansive proposals, it is conceivable that higher levels of perceived relevance might spread throughout the class, thus establishing a classroom culture of transfer (Pea, 1987). We plan to explore this premise through future analyses of our current data and in future studies.

Given that our framework includes five contextual aspects that can be expansively framed (time, place, participants, topics, roles), we must also consider possible explanations for the fact that in our data we only found relationships between transfer and alignment with two of these factors (time and role). Is an alignment with the expansive framing of time and role more important for transfer than an alignment the other aspects of contextual framing? What relationships exist among the different contextual aspects? In a concurrent research project, we aim to disentangle the effects of an expansive framing of role or settings (time, places, and participants), separately and in conjunction with one another, on transfer. The study involves one-on-one tutoring with students, where the tutor frames the role or settings in either an expansive or bounded manner. Knowledge assessments are used to measure learning and transfer, and online surveys record student recognition and alignment with expansive framing. The strong correlation between recognition of and alignment with expansive framing that we presented in this paper suggests that we may see changes in student alignment with expansive framing in each of these experimental conditions, and we plan to investigate the relationship between these levels of alignment and levels of transfer.

Finally, we must acknowledge that our measures of transfer for this study are limited in only considering performance on researcher-designed assessments as opposed to also considering additional forms of transfer that students may have been capable of (Lobato, 2012). We did draw on Lobato's notion of Actor-Oriented Transfer in analyzing interviews, but because those interviews were structured around our surveys, they did not provide students with many opportunities for self-reported transfer. Similarly, we have not yet conducted sufficient video analysis of student behaviors in class to systematically identify other possible manifestations of transfer. It is possible that including a greater variety of measures of transfer may illuminate additional relationships between transfer and alignment with expansive framing.

References

- Beach, K. (1999). Consequential transitions: A sociocultural expedition beyond transfer in education. *Review of Research in Education*, 24, 61–100.
- Belenky, D. M., & Nokes-Malach, T. J. (2012). Motivation and transfer: The role of mastery-approach goals in preparation for future learning. *Journal of the Learning Sciences*, 21(3), 399–432.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. Washington DC: National Academy Press.
- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking Transfer: A Simple Proposal with Multiple Implications. *Review of Research in Education*, 24, 61–100.
- Brown, A. L. (1989). Analogical learning and transfer: What develops? In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 369–412). Cambridge, United Kingdom: Cambridge University Press.
- Christianson, R. G., & Fisher, K. M. (1999). Comparison of student learning about diffusion and osmosis in constructivist and traditional classrooms. *International Journal of Science Education*, 21, 687–698.
- Clark, H. H. (1996). *Using Language*. Cambridge, UK: Cambridge University Press.
- Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist*, 47(3), 153–176.
- Engle, R. A. (2006). Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. *The Journal of the Learning Sciences*, 15(4), 451–498.

- Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. . *Educational Psychologist*, 47, 215-231.
- Engle, R. A., Meyer, X. S., Clark, J., White, J., & Mendelson, A. (2010). *Expansive Framing and Transfer in High School Biology Class: Hybridizing Settings and Promoting Connections within a Larger Learning Community*. Paper presented at the National Association for Research in Science Teaching Annual International Conference.
- Engle, R. A., Nguyen, P. D., & Mendelson, A. (2011). The influence of framing on transfer: Initial evidence from a tutoring experiment. *Instructional Science*, 39(5), 603-628.
- Floriani, A. (1994). Negotiating what counts: Roles and relationships, texts and contexts, content and meaning. *Linguistics & Education*, 5, 241-274.
- Gee, J. P., & Green, J. L. (1998). Discourse analysis, learning, and social practice: A methodological study. *Review of Research in Education*, 23, 119-165.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. . *Cognitive Science*, 7, 155-170.
- Gick, M. L., & Holyoak, K. J. (1983). Schema Induction and Analogical Transfer. *Cognitive Psychology*, 15(1), 1-38.
- Greeno, J. G., Smith, D. R., & Moore, J. L. (1993). Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 99-127). Norwood, NJ: Ablex.
- Gumperz, J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
- Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning: Research and perspectives* (pp. 89-119). Greenwich, CT: Information Age Publishing.
- James, M. A. (2008). The influence of perceptions of task similarity/difference on learning transfer in second language writing. *Written Communication*, 25(1), 76-103.
- Laboratory of Comparative Human Cognition. (1983). Culture and cognitive development. In P. H. Mussen (Ed.), *Handbook of child psychology: Vol. 1. History, theory and methods* (pp. 295-356). New York: Wiley.
- Lobato, J. (2012). The actor-oriented transfer perspective and its contributions to educational research and practice. *Educational Psychologist*, 47(3), 232-247.
- Lobato, J., Rhodehamel, B., & Hohensee, C. (2012). "Noticing" as an alternative transfer of learning process. *Journal of the Learning Sciences*, 21(3), 433-482.
- Ochs, E. (1996). Linguistic resources for socializing humanity. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 407-438). Cambridge: Cambridge University Press.
- Pea, R. D. (1987). Socializing the knowledge transfer problem. *International Journal of Educational Research*, 11, 639-663.
- Perkins, D. N., & Salomon, G. (2012). Knowledge to go: A motivational and dispositional view of transfer. *Educational Psychologist*, 47(3), 248-258.
- Reeves, L. M., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, 115(3), 381-400.
- Renkema, J. (2004). Structured content. In *Introduction to discourse studies* (pp. 87-102). Amsterdam: John Benjamins.
- Ross, B. H. (1984). Reminders and Their Effects in Learning a Cognitive Skill. *Cognitive Psychology*, 16(371-416).
- Tannen, D. (1993). *Framing in Discourse*. New York, NY: Oxford University Press.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Game-Enabled Agency: Outcomes that Matter

Sasha Barab, Arizona State University, Tempe, Sasha.Barab@ASU.Edu
Anna Arici, Arizona State University, Tempe, Anna.Arici@ASU.Edu

Abstract: The Doctors Cure learning experience is a 3D immersive game that positions players as protagonists in a virtual world where they must use their understanding of persuasive writing and how to gain evidence from complex texts in their role of investigative reporter. We report on the underlying educational theory, the game itself, and, to justify its impact, share data from a comparison study with 8 7th grade classes assigned the control and 8 assigned the experimental conditions—about 450 total kids, 90% receive free-and-reduced lunch and 87% are Hispanic. While both the game-based intervention group and the control showed statistically significant learning gains on short-answer and multiple-choice items, the intervention group demonstrated statistically significant gains with a large effect size on an open-ended persuasive essay task. Differences in engagement and learning are credited to player agency, affordances of the embedded scaffolds, and the power of consequential outcomes.

Statement of the Problem

There is currently a sizable gap in innovative, engaging, high-quality products designed to help middle school students build writing skills in a manner consistent with the Common Core. Specifically, literacy skills in the United States are failing to keep up with growth in other countries (Thompson et al., 2012) with a third of US 8th graders being proficient in reading and writing (National Center for Educational Statistics, 2011, 2012). More generally, every 9 seconds, an American student drops out of school, more than a million every year. One of the primary reasons youth cite for dropping out of school is a lack of engagement, thanks to the perceived lack of relevancy of the school curriculum to their lives. We need new curriculum models that excite and inspire youth, not simply to remember or even apply academic content, but to foster in them a confidence and commitment to use academic concepts (like those articulated in the Common Core Standards - CSS) in their own lives.

This requires the creation of new curriculum based on innovative learning theory, and that is intended to position academic content, individual learners, and those situations in which the content has value as interrelated. Videogames in particular are being touted as providing a powerful learning technology with the potential to transform education (Barab, Gresalfi, & Ingram-Goble, 2010; Gee, 2003; Shaffer, 2009; Squire, 2006), with many educators, researchers, designers, and even industry partners working to develop new forms of game-based curriculum. This is, in part, because games offer a well-designed mix of challenges, rewards, and goals that drive motivation, time-on-task, and levels of engagement that can seamlessly move back and forth between formal and informal learning environments. Games can make learning engaging, social, and relevant.

At the core of our theory of change model is that learning is not simply the acquisition of literacy skills cognitively, but involves meaningful participation in which the learner applies their emerging competencies within a situation in which they have agency and consequentiality to accomplishing something they value (Barab, Gresalfi, & Ingram-Goble, 2012; Brown, Collins, & Duguid, 1989; National Research Council, 2002). They can give students real agency in ways that static textbooks simply cannot. Beyond addressing the engagement and relevancy challenges, games are also well positioned to address key foundational skills such as those specified in the Common Core Standards.

Theoretical Framework

The philosopher and educator John Dewey (1938) supported a transactive view of schooling, where learners are active change agents rather than passive observers, and through their actions and consequences, they transform the problem into a known. Technology has finally caught up to Dewey's vision, and we now have tools that make it possible to individualize learning and provide authentic tasks and roles to students, particularly in the field of game-based learning. Digital games are different from other media in that they are interactive, participatory and highly engaging. They enable players to step into different roles (e.g. scientist, explorer, inventor, political leader), confront a problem, make meaningful choices, and explore the consequences.

Through immersing players in virtual worlds designed to support deep and engage learning, there is real potential to position students (even in the context of schools) as agents-of-change who use real-world knowledge, skills, and concepts to make sense of a situation and then make choices that actually transform the play space and themselves; creating a place in which *what you know* is directly related to *what you are able to do* and, ultimately, *who you become* (Barab, Gresalfi, & Arici, 2009). Further, games have the potential to provide the learner a motive and motivation for pursuing this expertise, with consequential feedback on their

strategies, strengths, and even their own evolving identity (Barab, Gresalfi, & Ingram-Goble, 2010; Gee, 2003; Squire, 2006; Steinkeuhler, 2006).

Many of the strengths of game-based learning can be summarized in the theory of *Transformational Play*; a 3-fold theory that positions the *person* with intentionality, the *content* with legitimacy, and the *context* with consequentiality (Barab, Gresalfi, & Ingram-Goble 2010). Transformational play as a theory has emerged through our design and research on games for learning. The idea of transformational play highlights relations among the three interconnected elements of person, content, and context. In these games, learners become virtual protagonists who use the knowledge, skills, and concepts of the educational content to first make sense of a situation and then make choices that actually transform the play space and the player—they are able to see how that space changed because of their own efforts.

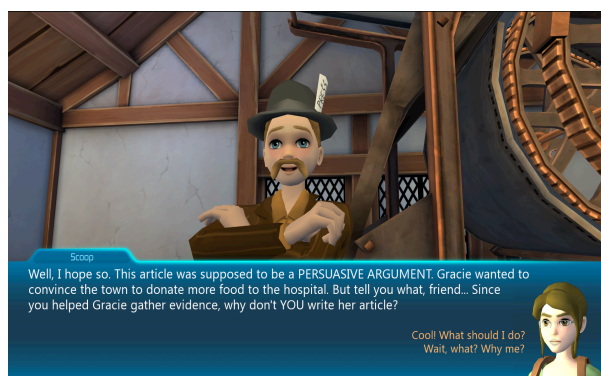


Figure 1. Screenshot from Doctors Cure.

Such play is transformational in that it changes the context in which play is occurring, at the same transforming the player and his or her potential to interact with the world. When creating learning environments to bring about transformational play, we use game-based technologies to position person, content, and context in the following ways:

- Person With Intentionality (positioning players as protagonists with the responsibility of making choices that advance the unfolding story line in the game)
- Content With Legitimacy (positioning the understanding and application of academic concepts as necessary if players are to resolve the game-world dilemmas successfully)
- Context With Consequentiality (positioning contexts as modifiable through player choices, thus illuminating the consequences and providing meaning to players' decisions)

In this way, gaming technologies now make Dewey's vision a reality, putting learners as active protagonists in their own learning, taking on authentic roles via avatars, and seeing the consequences of their actions played out in a 3D immersive world.

Atlantis Remixed: The Doctor's Cure

Informed by this theory, and to further investigate its underpinnings, we evolved a 3D immersive role-play game for learning literacy skills, called "The Doctor's Cure." While based on previous work (Barab, Pettyjohn, et al., 2012), in this study we have designed a new engine with powerful new learning affordances including pedagogical scaffolds, embedded assessment, and a powerful argumentation mechanic and tool. This game is embedded within the Atlantis Remixed (ARX) Project, an international learning and teaching project that uses 3D virtual environments to immerse children, ages 9-16, in educational tasks (the second generation of Quest Atlantis). Through interactions with in-game mentors (non-player characters, or NPCs) and by using in-game tools to engage the academic content, the students are given the scaffolds and affordances necessary to take on the role of an expert in an authentic task and make influential decisions, which they see played out in their virtual world.

"The Doctor's Cure" (TDC) is a literacy game based on Mary Shelley's novel *Frankenstein*, and set in a gothic world, where students take on the role of an investigative reporter via their avatar, and complete a series of missions to uncover a moral dilemma involving Dr. Frankenstein's work. As reporters, students actively collect evidence through interviews and investigations, build logical arguments to support their theses, submit these to an in-game logic machine for evaluation, and get feedback about the alignment between their evidence and reasoning. Players are intentionally positioned as agents of change whose purpose is to help the village of Ingolstadt decide if they should allow "Dr. Frank" to keep looking for a cure in spite of his questionable research methods. Players soon learn that persuasive writing is a necessary tool to resolve the

game's narrative conflict. As the game progresses, players experience how their choices and use of persuasive writing dramatically change Ingolstadt, its citizens, and even the students' own identity as a writer and leader.

In-game tools provide support in the interrogation of texts, as well as a model for testing the logic of their argument, and immediate feedback in the process. One of the pedagogical scaffolds in the game is the 'Lens of Lumination' tool (see Figure 2), which allows students to examine texts above their current reading level. The Lens of Illumination is designed to help the player engage in meaning making by "illuminating" the relevant claims in the documents, so players can decide if they want to collect that evidence. Once players have gathered what they think is a good collection of evidence to support their thesis, they visit Scoop Perry, and use his Persuasive Argumentation Tool (PAT), which displays all the evidence they have collected thus far, they then drag-and-drop their evidence to match their claims, and create a chain of logic to support their thesis (see

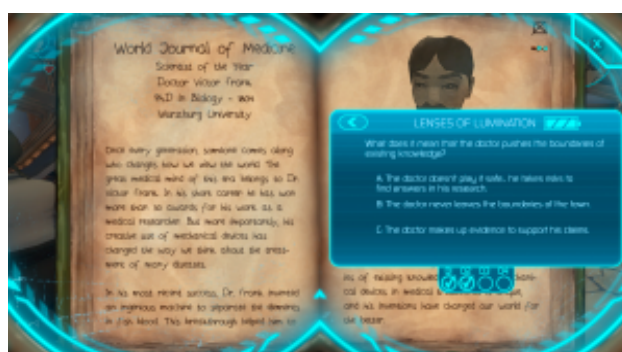


Figure 2. Screenshot of the Lens of Lumination.

Figure 3). The immediate feedback and flexibility of the tool allows players to move elements around easily, repeat and test various logic orders. Once sound, players proceed with the game and write their persuasive essay for the town newspaper, to convince the townspeople of their view. These tools are a small part of the much larger narrative.



Figure 3. Screenshot of the argument tool.

Consistent with our notion of games as a service, we provide a teacher toolkit and dashboard that allows teachers to easily manage their classes: registering students, assigning content and reviewing student work. This toolkit will build on the existing work done for the Atlantic Remix platform. This system provides ongoing data to teachers about the state of student gameplay, a review interface for evaluating and providing feedback on student work, as well as access to other relevant choices that can be used to support class discussions or Socratic dialogues with students.

Below is an image of the progress tracker, in which teachers can monitor what task and mission student is working on. Also below, is an example of the teacher feedback in which they can use a rich text editor to leave feedback and notes in the body of the student essay, or as a formal review on the side panel in the guise of an in-game character (e.g., Scoop). The fact that both the student work and teacher comments can be accessed and edited via basic html allows for continual feedback to support students in the writing revision process. At any point teachers can click on the blue button and see the supporting data that the students used to craft their essay. More than listing of facts the students have collected, teachers can see the actual embedded assessment and conceptual models that students generated while they were working through the game.

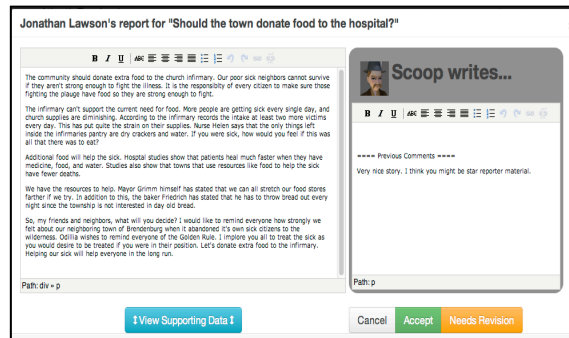


Figure 4. Teacher Feedback on Student's uploaded essay.

Progress Reports Summaries Evidence Choices Meters						
Student	Username	Intro	M1	M2	M3	M4
Babin, Erica	ebabin					
Bateson, Phillip	pbateson					
Curtice, Brian	bcurtice					
Elochs, Jordyn	jenochs					
Gulian, Kristoffer	kgulian					
Henry, Jeffrey	jholmes					
Henry, Matthew	huemas					
Kim, Younsu	yskim1					
Lawson, Jonathan	jrlawson					
McCoy, Lisa	lmccoy					

Figure 5. Class Mission Progress.

Methodology

Study Overview

To test the theory of Transformational Play, we implemented The Doctor's Cure game across a school district of 7th grade Literacy classes. This study compared a 3D gaming curriculum and context with that of a similarly novel curriculum based in the graphic novel *Frankenstein*. Interview measures showed the graphic novel to be incredibly popular, and equal to the game-condition in its engagement measures. The comparison study was a quasi-experimental design of intact classes, with 8 classes assigned the control and 8 assigned the experimental conditions—about 450 total students with just over 400 completing the pretest and the posttest.

Participants

This school district, located in the southwestern United States and bordering on Mexico, is a fairly large district with more than 18,000 students, many disadvantaged, and have a demographic breakdown that includes 95% minority groups (87% being Hispanic), 83% low socio-economic status, and 20% English Language Learners. Additionally, the community recently approved over \$25 million dollars in bonds for technology upgrades, bringing one-to-one laptops to all students. Participants were 12-13 years old.

Measures

Measures included traditional instruments, quantitative assessments, as well as ethnographic techniques. Learning gains were measured by a pretest and posttest, which were identical for both conditions, and counterbalanced within conditions. Questions varied from lower level multiple choice, to short answer, to a final essay writing question, where all students crafted their own persuasive essay with a given prompt. A rubric was created to analyze all open-ended responses and two raters scored a subset of tests with an interrater reliability of $\alpha = .81$. To quantitatively assess the engagement of the learner, a student self-report measure was gathered. This Engagement Questionnaire was based on that of Csikszentmihályi's (1990) study with 'flow', where he interrupted students involved in various activities to respond to their current state of engagement, motivation and challenge in the task at hand.

Results

Learning Data

A repeated measures ANOVA on the multiple choice/short answer test revealed a significant main effect for testing time, $F(1,402)=46.25$, $p=.000$, a non-significant main-effect for condition, $F(1,402)=1.72$, $p=.191$, and a non-significant interaction, $F(1,402)=.07$, $p=.793$. In other words, while both the control and experimental groups demonstrated statistically significant learning gains on the *multiple choice/short answer tasks* with small/medium effect sizes (Con ES=.33, Ex ES=.33), there were no differences for group gains. See Figure 1, left side. A repeated measures ANOVA on the *persuasive essay task* revealed a significant main effect for testing time, $F(1,353)=14.47$, $p=.000$, a non-significant main-effect for condition, $F(1,353)=.87$, $p=.350$, and a significant interaction, $F(1,353)=18.32$, $p=.000$. Follow-up analyses indicate that while the control, $t(167)=.36$, $p=.72$, did not have statistically significant learning gains, the experimental condition, $t(187)=5.73$, $p=.000$, improved significantly with a moderate effect size (ES=.51) from pre-post on the essay writing task (see Figure 1b). See Figure 1, right side.

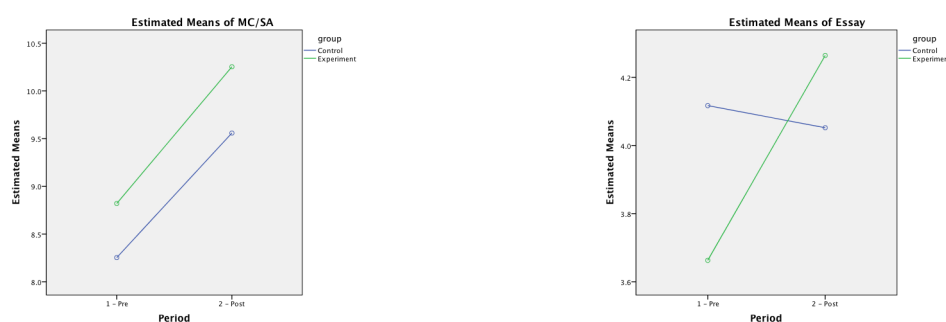


Figure 6. Learning Gains: Left side is comparison of groups over time on multiple choice/short answer. On the right side is comparison of groups over time on essay prompt.

These data show that both the comparison (graphic novel) group and the experimental (game) group had significant learning gains from pre to post test on low-level concepts related to literacy and persuasive argumentation. That is, both groups were able to identify and describe the basic elements of persuasive writing equally well. However, when looking at the final essay composition, requiring the highest levels of thinking, application, synthesis and evaluation, the students in the game condition students scored significantly higher in their compositions.

These findings lend support for Transformational Play, in that students who were actively engaged as first-person protagonists in the game narrative, with an authentic role and purpose for writing persuasively, did better in their writing than did those who had similar assignments based around a passive narrative in a graphic novel. Further, the game students had the contextual tools embedded in the virtual world to help them make sense of the evidence and create meaningful logic models. Students later demonstrated they were able to write persuasively even in the absence of these tools on the posttest, suggesting near transfer of these skills to a non-game context.

Engagement Data

We collected engagement data using a modified version of Csikszentmihalyi (1993) flow's instrument that consisted of 7 Likert-type questions (e.g., did you enjoy what you were doing, was the activity challenging, were you succeeding at what you were doing) and had .77 internal consistent using Cronbach's Alpha. Results showed that there were statistically significant differences between groups, $t(125)=5.41$, $p=.000$, with the experimental group ($M = 3.86$, $SD = .62$) outscoring the control group ($M = 3.29$, $SD = .58$) with a large effect size (ES=.95). Additionally, a Chi Square carried out on data in which students were asked: "What was your main reason for doing the task?" Results were significant, with the distribution being statistically different from chance, $X^2(121)=35.67$, $p < .001$. In fact, 74% (49/66) of students in the treatment condition attributed that choice to "because I'm interested in the task," as opposed to 22% (12/55) of the control. In contrast, 75% (43/55) of the control chose either "to get a grade" or "my teacher told me to," while only 23% (17/66) of the treatment condition chose this option. These findings are summarized in Table 1.

Table 1. Engagement Data. When students were asked “What is your main reason for doing this task?” a significant number of game-based students attributed it to being intrinsically interested in the task, rather than for an external grade or teacher direction.

Attribution	Treatment (Game)	Control (Graphic Novel)
‘I’m Interested in task’	49 (74%)	12 (22%)
‘Teacher told me to do it’	2 (3%)	14 (26%)
‘To get a good grade’	15 (23%)	28 (51%)
‘Other’	0 (0%)	1 (1%)

Qualitative Lessons

Interviews and written feedback were collected from the students throughout the implementation, and at its conclusion. When asked to describe their experiences in this unit, these are comments that are representative of the most commonly offered from students. First, there were several reactions that were similar for both the Game and the Graphic Novel students. Both groups described the experience as:

- “Challenging”, “Engaging”, “Fun”, “I am learning a lot”.
- “It will help improve my writing”,
- “This is something I’d like to continue on my own”.
- “When the story ended, I was shocked because I wanted to read more” (Graphic Novel)
- “Once I finished the game I was like, “Aw, man!” I wish there was another part. It was cool and fun to play.” (Game condition)

In fact, both conditions continued to access the games and novels after the study ended, and switched to participate in the activity they were not originally assigned.

Next, there were some reactions that were more isolated to only the Game Condition students. The following are representative reactions to the game that differed from the graphic novel condition, with categorical interpretations in parentheses:

- “I’m surprised that they let us play this game at school!” (engagement)
- “The game involved a lot of reading and writing.” (challenge)
- “I look forward to see what is happening next, and how it will turn out.” (consequentiality)
- “My opinion about who is right and who is wrong has changed over time.” (multiple perspectives)
- “I’m investigating the other side of the debate now.” (active participant, multiple perspectives)
- “The people agree and disagree with the Doctor. It will be interesting to see how they act after my article is published!” (embodied role, multiple perspectives)
- “I like it because it gives you some feedback and you can go back to fix it.” (engagement)
- “The game helps me be much more persuasive, so I can convince the people [in the game] about what I’ve found [as an investigative reporter].” (embodied role, engagement)
- “I’m looking forward to writing my essay so I can convince people that they should support the doctor!” (strong opinions, motivated to articulate)
- Some felt strongly that the doctor should not experiment on the creature because he has rights. Others felt that the ends justified the means, and the cost of the creature was worth saving the thousands of others from the plague. Others took middle ground, suggesting alternative ways to find a cure for the plague.

Game classes were incredibly on-task, and had little conversation going on that wasn’t related to the game. Many students were up and helping others, cooperating in game play, and celebrating when they were successful. The field observations and interview data are encouraging in that they further support the strengths of Transformational Play. While space limits prohibit extensive transcripts here, we will briefly overview three findings.

Importantly, teachers felt both conditions were equally engaging to students, with some teachers even arguing that the graphic novel might have been more exciting to students:

I think the kids found the game engaging, but they found the graphic novel equally engaging too. I think the difference is in the game **they also feel empowered**, even though I think that their world in the graphic game was in some way, I think, overwhelming. And it wasn’t so much the ‘3D-ness’ of it, as much as **they saw direct cause and effect**...They’re exploring the world at that age, of “**Can I make a difference?**” The first time they do something or write something that makes a difference, they think “**That was a real big deal!**”

However, what really seems to be powerful from this teachers' perspective was how the game provided player agency, consequentiality, and a sense of purpose that was relatively absent in the novel condition. More generally, one of the key observations was the sense of purpose or intentionality that students in the gaming condition as they often discussed the importance of their essay in accomplishing their goals. Reflecting across the field notes, the following three areas seemed to be especially important with respect to motivation learning and engagement.

Pedagogical Scaffolds

The pedagogical scaffolds and in-game tools were key for student success. Teachers and students reported how much they loved the scaffolding tools in the game, and the feedback they received on the quality of their logic model. Teachers felt that this was the largest contributor to student success in argumentation. The ability for students to drag and drop ideas alleviated much of the language barrier present in this population of students, and they could focus on just the logic model itself.

Consequential Feedback

One of the most engaging aspects of the experience for students was the consequential feedback they received around their choices. When players made a choice or performed a series of actions and the game responded, they would literally "hoot" out loud, or laugh and look around at other computers. We saw numerous examples of students discussing different outcomes with each other in the classroom, hallway, or in some cases we heard about students still discussing choices at home. As the unit progressed, a number of students' enthusiasm started to wane, however, and we need to figure out better ways of providing reinforcement and successes to keep students engaged for longer periods.

Professional Development

Game-based learning brings with it new pedagogies. Teachers need training in how to provide valuable feedback on student work, how to use embedded data in the system to support meaningful class discussions, and what is one's role and responsibilities for successfully implementing a game-infused curriculum. More generally, teachers needed ongoing support to engage the system successfully. Simply training them before implementation was inadequate, and we found that prompting teachers to review student work, to probe student understanding, and to provide just-in-time lectures was necessary.

Significance

The Doctor's Cure was created and grounded in the theory of Transformational Play, leveraging the three interconnected elements of *person*, *content*, and *context* in immersive gameplay. This study explored the nuances of learning and engagement, when those three elements are supported, demonstrating significant learning gains and deeper engagement at higher levels of authentic practice. Qualitative findings uncovered additional strengths, such as the in-game tools for providing practice and fluency in working through difficult logic models and complex texts. However, teachers need additional support to shift their pedagogy into a game-based approach. Importantly, this study took place in a challenging context with mostly struggling learners, and with a high percentage of at-risk learners who frequently opt out of school, in part because the lack of meaning with respect to academic content in terms of their personal lives. Through the game we were able to establish a rich context for learning that justified the use-value of the content, and gave these youth a sense of consequentiality to their learning and to potentially to their sense of self. We see this as an invaluable component of games.

Importantly, however, is that players are able to use this new found confidence and ability, to realize goals in their own lives outside of the game world or school. While constructing an individual bounded game has the potential to result in strong engagement and foster desired learning outcomes, central to our theory of change is to expand this vision of impact games to think of the medium as on-going services that support multiple game-infused experiences and real-world extensions where core lessons are brought outside the fictional gaming context. In this way, we are currently building in more features to scaffold rich conversations around the game play—multiuser experiences and collaborative smart tools to facilitate player interactions.

We are also adding in professional networks and the ability to extend the lessons learned but also the scaffolding tools into the real-world where players can build logical arguments and eventually support change in their own community. This involves, for example, having online profiles where one's professional network begins with "achievements" that are taken from the game world, and then as they use the argument tool on interest-driven agendas are able to expand towards accomplishing real-world goals in their community. A key challenge here is to balance player agency and emergent situations with productive constraint, and providing

designed experience where they can level up in ability, and build the necessary confidence and commitment to persist on complex real-world tasks.

References

- Arici, A. (2009). Meeting kids at their own game: A comparison of learning and engagement in traditional and 3D MUVE educational-gaming contexts. Unpublished doctoral dissertation, Indiana University, Bloomington.
- Barab, S., Gresalfi, M. and Arici, A. (2009) "Transformational play: Why educators should care about games". *Educational Leadership*, 67(1), 76–80.
- Barab, S., Gresalfi, M., and Ingram-Goble, A. (2010) "Transformational play: using games to position person, content, and context", *Educational Researcher*. v39 i7. 525-536.
- Barab, S.A., Pettyjohn, P., Gresalfi, M., Volk, C., & Solomou, M. (2012). Game-based curriculum and transformational play: Designing to meaningfully position person, content, and context. *Computers & Education* 58(3): 518–533.
- Csikszentmihályi, M. (1996) *Creativity: Flow and the Psychology of Discovery and Invention*, New York: Harper Perennial, ISBN 0-06-092820-4.
- Dewey, J. (1938). *Experience & Education*. New York, NY: Kappa Delta Pi. ISBN 0-684-83828-1.
- Gee, J. P. (2003) *What video games have to teach us about learning and literacy*. New York: Palgrave.
- Halverson, R. (2005) *What can K-12 school leaders learn from video games and gaming?* Innovate. 1(6).
- National Center for Education Statistics. (2011). *The Nation's Report Card: Reading 2011 (NCES 2012–457)*. Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.
- National Center for Education Statistics. (2012). *The Nation's Report Card: Writing 2011 (NCES 2012–470)*. Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.
- Shaffer, D. W. (2007) *How computer games help children learn*. New York: Palgrave Macmillan.
- Squire, K. (2006) From content to context: Videogames as designed experience. *Educational Researcher*, 35, 19–29.
- Steinkuehler, C. A. (2006) "Massively multiplayer online video gaming as participation in a discourse". *Mind, Culture, and Activity*, 13(1), 38–52.
- Thompson, S., Provasnik, S., Kastberg, D., Ferraro, D., Lemanski, N., Roey, S., & Jenkins, F. (2012). *Highlights From PIRLS 2011: Reading Achievement of U.S. Fourth-Grade Students in an International Context (NCES 20131010)*. Washington, D.C.: National Center for Education Statistics, IES, U.S. Department of Education.

Model-Based Reasoning: A Framework for Coordinating Authentic Scientific Practice with Science Learning

Julia S. Gouvea, Arash Jamshidi, Cynthia Passmore,
University of California, Davis, One Shields Avenue, Davis, CA, 95616, USA
Email: jmsvoboda@ucdavis.edu, ajamshidi@ucdavis.edu, cpassmore@ucdavis.edu

Abstract: In this paper we explore data from a two-year professional development program intended to help teachers adopt and enact a model-based reasoning view of science instruction. We use a resources framework to explore the teacher learning in this setting. By examining the experience of one teacher, we investigate how the ideas about learning and scientific models this teacher brought to the PD environment were aligned and coordinated over time to result in a change in her self-reported pedagogy by the end of the program. We explore implications of this view of teacher learning for PD in the current reform context.

Introduction

The assumption that science learning should be situated in authentic scientific practice underlies the vision of the recently released *Next Generation Science Standards* (NGSS) in the United States. The important step taken by the new standards framework is that it calls for supporting students to engage with and learn science ideas *through* engagement in authentic scientific practices. In this way the framework puts forward a vision of science learning that brings together constructivist theories of learning with an understanding of science as practice that helps scientists generate and refine knowledge about the world. For many teachers this will mean a radical shift in the character of their classroom instruction. In order to support teachers in enacting this new vision, teacher educators must consider what and how teachers learn (Wilson, 2013).

Much of the scholarship on teacher learning in science has focused on what teachers need to know, and how this knowledge can be brought into better alignment with education reform. In general, this literature tends to treat teacher knowledge and beliefs as deeply held core conceptions that influence how teachers plan and enact their instruction. If these core beliefs do not align with the vision for reform, then they need to be altered or changed through professional development experiences. In this paper we conceptualize teacher cognition in terms of finer-grained knowledge elements, or resources (Hammer et al., 2005a), and model teacher learning in terms of different patterns of activation and coordination of fine-grained knowledge elements. We present a case analysis of a teacher participant chosen because we observed evidence of substantial shifts in the ways in which she described her pedagogy over time in the context of a professional development program that emphasized model-based reasoning as a framework to guide science teaching and learning.

Theoretical Framework

Teacher Knowledge Is Complex and Dynamically Related to Teaching Practice

In the field of teacher learning there exist a number of complex constructs for describing teacher knowledge (attitudes, PCK, beliefs, orientations, core conceptions). What is common across these descriptions is the understanding that teachers bring together a range of different kinds of knowledge to their curricular and instructional decision-making. Researchers have attempted to describe different components and constellations of such knowledge in an effort to guide teacher learning and support professional development. A central aim of this body of work is to describe the kinds of knowledge held by exemplary teachers and how it differs from novice teachers (e.g. Shulman, 1986).

However, the idea that teachers have stable knowledge structures or belief systems that hold in all contexts has been challenged for both theoretical and methodological reasons (Friedrichsen et al., 2011). Mounting empirical evidence suggests that teachers have many different cognitive resources about teaching and learning, and they do not necessarily leverage that knowledge in consistent ways across contexts, or even across moments (Crawford, 2007; Kang and Wallace, 2005; Levin et al., 2009; Markauskaite & Goodyear, 2014). The field is moving towards a conception of teacher knowledge that is more dynamic and context dependent.

This shifting conception of teacher knowledge mirrors the theoretical movement in student learning to describe student knowledge as comprised of sets of interrelated “resources” – fine-grained cognitive elements that are differentially activated with context (diSessa, 2002; Hammer, 1996). Within a resources framework learning need not entail replacing or radically restructuring students’ knowledge frameworks. Instead, productive learning can often mean helping students build upon, connect or reorganize existing resources (Hammer, 1996; Hammer et al., 2005b) so that they access and leverage those resources in particular ways in various contexts. When viewed through this lens, it becomes important to understand the ways in which students *use* knowledge in different contexts, and insufficient to simply demonstrate that they *have* correct knowledge.

Applied to teacher learning and PD, a resources framework highlights the need to understand and help teachers refine the knowledge they already have. That is, we want to understand when, how and why particular resources are coordinated in service of pedagogical purpose and action and how to shift these patterns over time. In our case analysis we focus on one teacher, Quinn, who describes her pedagogy as transformed by our PD and ask, how can we make sense of this shift in terms of shifts in this teacher's cognitive ecology? We argue that Quinn developed *coherence* among knowledge resources and pedagogy as demonstrated in her descriptions of her classroom practice and her reflections on her actions. We posit that it is this coordination that is important for lasting change rather than the acquisition of any single particular idea.

Resources for Supporting Students' Engagement in Authentic Scientific Practice

In general, a "resource" can be defined as abstract basic knowledge unit that can be applied across a range of contexts. For example, Markauskaite & Goodyear (2014) propose a number of different knowledge categories from which such resources could be drawn (e.g. resources about *how people learn*, *disciplinary content*, and *student capabilities*). In this paper we focus specifically on two types of resources that are particularly relevant to the NGSS vision: resources related to (1) *student learning* (2) *the nature and purpose of scientific models*. Separating these two dimensions is important for understanding how teachers make instructional and curricular decisions. Often researchers collapse teachers' views of scientific practice and teaching and learning. For example, teachers' orientations towards "inquiry" tend to combine conceptions of the science as a practice of asking questions about the world and an orientation towards a student-centered classrooms in which students are the ones asking and answering questions. In this paper, we purposely tease apart resources so that we can explore how interactions among these resources changed during her participation in a professional development program that focused on model-based reasoning.

Methods

Study Context

The Innovations in Science Instruction through Modeling (ISIM) program was designed to engage science teachers in the conceptual work of problematizing current approaches to science instruction and restructuring their curriculum and instruction to include model-based reasoning (MBR) in their classrooms. From 2007 to 2011, two cohorts of 6-12th grade science teachers (N=57) participated in a two-year intensive professional development (PD) program focused on MBR. Here we report on one teacher from the second cohort (N=28).

MBR is an approach to science teaching that involves students in constructing and using models to explain natural phenomena (see Passmore & Stewart, 2002; Stewart, Cartier & Passmore, 2005; Windschitl, Thompson & Braaten, 2008 for more thorough treatment of this approach). Each year of the two-year PD program consisted of a three-week summer institute in which teachers experienced MBR as adult learners to deepen their own content knowledge and analyze their own learning process when introduced to the MBR approach. During the academic years of the program, teachers worked in content- and grade-alike teams to develop and enact curriculum appropriate for their own classroom integrating the MBR approach. Teams met regularly throughout the year to report and reflect on their experiences implementing MBR during afterschool meetings and a day-long retreat in the spring of each year.

Data Collection and Analysis

Over the course of the two-year program, we continually collected data in the form of artifacts from the PD program including approximately 100 written reflections per participant, field notes, video and audio of summer institutes and academic year events, and facilitator notes from lesson study team meetings. Each participant also participated in a small group and individual interviews during which participants discussed and reflected on prepared reflection prompts. All of the data were transcribed and entered into a FilemakerPro database. In this paper we report results from two types of analyses conducting using written prompts and interviews. The first is a pre-post comparison of a single written prompt teachers responded to both upon admission to and exit from the PD. The second uses the full set of prompts to track changes in teacher responses over time.

The majority of prompts were generated in an *ad hoc* manner throughout the program in order to provide opportunities for teachers to reflect on their participation in the PD as well as their evolving classroom practice. Because the nature of the prompts changed over time they cannot be used to measure absolute shifts in teachers' responses. However, because all teachers had the opportunity to respond to the same set of prompts, we can use relative differences across teachers as an indicator of the ideas that were most salient to them at different points throughout the program. We conducted a qualitative analysis of written prompts and interview transcripts in order to generate longitudinal descriptions of changing response patterns for the teachers in the second cohort. As a research team we iteratively developed a set of 14 categories to capture teacher statements by first reading through the full datasets for 10 teachers. We then separated statements into these different

categories (double-binning was permitted), which made it easier to parse the large dataset for further coding and analysis described in more detail below.

Pre-Post Analysis of Teachers' "Exemplary Lessons"

One of the written prompts asked teachers to "Describe one of your exemplary science lessons and explain your reasoning about the most important things that make it effective." Teachers responded to this prompt on their initial applications to the program (pre-lesson) and again two years later on exiting the program (post-lesson). To analyze specific features of these lessons we created a rubric that we used to score the lesson along two dimensions: attention to student ideas as an overall measure of student-centeredness, and authenticity to scientific practice as an indicator of how science is portrayed in the classroom. Each lesson was given a score ranging from 1 to 5 along each dimension. A score of 1 in the student-centered dimension described primarily lecture-based lessons, whereas a score of 5 signified lessons in which the teacher described giving students the opportunity to both share and reflect on their ideas with responsive feedback from the teacher. A score of 1 in the science dimension indicated lesson descriptions not involving students engaging in scientific thinking or practices, whereas a score of 5 represented lessons in which students combined science ideas and practices in order to make sense of a phenomenon. During scoring, all identifying markers were removed and pre- and post-lessons were randomized. Two researchers scored all lessons with initial reliability of 78%. Through discussion an inter-rater consensus of 100% was achieved.

Longitudinal Analysis of Shifts in Teachers' Instructional Choices and Pedagogical Strategies

One of the 14 categories captured statements about teachers' self-reported *instructional choices and pedagogical strategies*. We established qualitative codes for statements in this category by first reading through the binned dataset for each of the 28 teachers. As a research team we iteratively developed a set of codes drawing on both *a priori* expectations and emergent patterns in the data. The final set of codes were assigned to describe statements that fell into broad instructional strategies: (1) *traditional* approaches to science instruction including both lecture and activities focused on engagement (2) *student-centered* strategies such as group learning making student ideas visible (3) *procedure*-focused strategies including engaging students in laboratory skills (4) sense-making strategies focused on *explaining phenomena* and (5) *modeling* strategies that explicitly referenced involving students in model development or use.

Three researchers participated in the coding process by first coding a subset of the teachers and conducting extensive discussions to normalize code assignments. A single researcher conducted the remainder of the coding. Researchers were blinded to the identities of the teachers during the coding process.

In order to visualize changing patterns in code distribution over time, we divided teachers' responses into five time periods corresponding to the major time periods in the program (1st spring, 1st summer, 1st academic year, 2nd summer, 2nd academic year). We then counted up the number of responses corresponding to each code during each of the five time periods and plotted the percentage of statements made in each category for each teacher (see Fig 1). This provides a visual representation of the relative frequency of specific codes.

Qualitative Analysis of Quinn Focused on Alignment and Interactions Among Resources

We selected Quinn as a case for a more in depth analysis because of the dramatic change in the character of her exemplary lesson as well as the shifting patterns in her prompt responses related to her instructional practice. Because we were interested in the cognitive resources Quinn might be using to make and justify her instructional choices, we conducted a more in depth analysis of two categories of statements we found in the data: teachers' conceptions of *student learning* and views of *scientific models*.

We developed the Quinn case by reading through all statements in each category to first develop a qualitative narrative of how these statements changed over time. We then used the responses collected through the first summer institute to build inferences about the ideas that were most salient to Quinn during her early participation in the program and responses collected during the second summer and academic year to describe the ideas that were most salient to Quinn towards the end of the program. We then took a third pass through the data looking for responses coded for both views of student learning and scientific models to build up an account of how these different ideas were related for Quinn. Finally, we read through Quinn's descriptions of and reflections on the MBR framework to better understand what it meant to her and in what ways it might be influencing the way she thinks about learning and teaching science.

Results: The Case of Quinn

Quinn's Shifting Descriptions of Her Teaching Practice

One indicator of Quinn's perception of her changing teaching practice is given by the comparative analysis of her pre (before entering the PD) and post (upon exiting two years later) exemplary lessons (Table 1). Quinn's first lesson received a score of 2 in both rubric dimensions. In terms of engaging with student ideas, Quinn's pre lesson relied mainly on direct instruction with some time given to allow students to rotate through lab stations, while the scientific aspect of the lesson is dominated by procedural laboratory skills.

Quinn's post lesson received a score of 5 in both dimensions. In it Quinn describes students working in small groups while she played the role of questioning and facilitating whole class discussions. The science component of the lesson involved an exploration into a number of phenomena leveraging ideas about particle movement in response to temperature. Lessons that shifted from a score of less than 3 to a score of 4 or greater in both dimensions were considered to be large shifts. Quinn was one of 11 teachers (42%) whose lessons shifted in this way (N = 4, had smaller shifts, N=4 had high scoring lessons in both pre and post and N=7, had low scoring lessons in both pre and post).

Table 1: Comparison of excerpted sections from Quinn's description of an "exemplary lesson"

	Attending to Student Ideas	Authentic to Scientific Practice
Pre Lesson	Score: 2 "After providing some direct instruction about elements, the Periodic Table, and chemical bonding, the students partake in a lab activity..."	Score: 2 "...the students partake in a lab activity where they rotate through nine different stations. ...the students perform a task of mixing the chemicals together, heating the chemicals, or some other action."
Post Lesson	Score: 5 "I then had students respond to a challenge statement ... Afterwards, students practiced group norms as they shared in small and large groups." "I facilitated and questioned student thinking."	Score: 5 "We started with discussion about particles of liquids, gases, and solids and how temperature and heat affect it or are affected..." "we examined phenomena and attempted to explain them. Phenomena included: 1. Blue ice melting in warm tap water..."

A second indication of a shift came from examining changes in how Quinn described her teaching over the course of her participation in the PD program as captured by five different types of coded statements (Figure 1). Overall, 110 written or transcribed statements from Quinn were categorized as *instructional choices and pedagogical strategies*. What drew our attention to Quinn was the relative increase in statements about *attending to student thinking* and engaging students in *explaining phenomena* coupled with a decrease in statements coded as *traditional instruction* and *laboratory or procedure-based instruction*. Compare for example the relative frequency of statements made in these categories upon entering the program (1st spring) and near the end of the program (2nd academic year) as depicted in Figure 1.

The final strand of evidence we used to triangulate our account of Quinn is her own descriptions of her teaching. During the first summer institute Quinn described her current teaching practice as "a set of lecture notes of facts that the students should know for the test, lab experiments that only serve to prove the facts already presented, and reinforcement activities of the same rote facts" (135, 1st sum.). Quinn also described attempting to engage her students with "hands-on" activities, but admitted that she often found herself "giving the answer" (132, 1st sum.). Two years later, Quinn recounted the changes she made in her classroom as she discussed her experience in the PD program with her peers in a group interview:

Labs used to be very cookbook, and, "here, do this, do that, don't do this, don't do that." And now I've given them materials, I say here's some driving questions you might want to think about, what do you think is happening or how do you think it's working or could you design an experiment to test your ideas? (240, 2nd acad. yr.)

Quinn also described her shifting role as an instructor in the classroom saying she "eliminated a lot of the direct instruction and lecturing" and was instead "question[ing] student thinking" (234, 2nd acad. yr.). For example, Quinn described beginning each class by asking students to think about "a phenomenon they may have experienced in their everyday lives." (240, 2nd acad. yr.). Quinn also drew attention to a in the way she organized her classroom wall space. The information she has displayed on her walls was "no longer posters of this fact or

that fact and the periodic table. It's more of what the students' ideas are and how they've changed over the course of the term" (240, 2nd acad. yr.). Her approach to assessment also shifted: Quinn described focusing on whether the "reasoning makes sense with the model" when assigning points to her students' work (240, 2nd acad. yr.).

Overall, the comparison of Quinn's descriptions of exemplary lessons and the quantitative and qualitative analysis of Quinn's descriptions of her instructional approach suggest a shift in emphasis from traditional lecture-based instruction and hands-on laboratory activities to classroom discussions in which students are encouraged to develop their own explanations for phenomena.

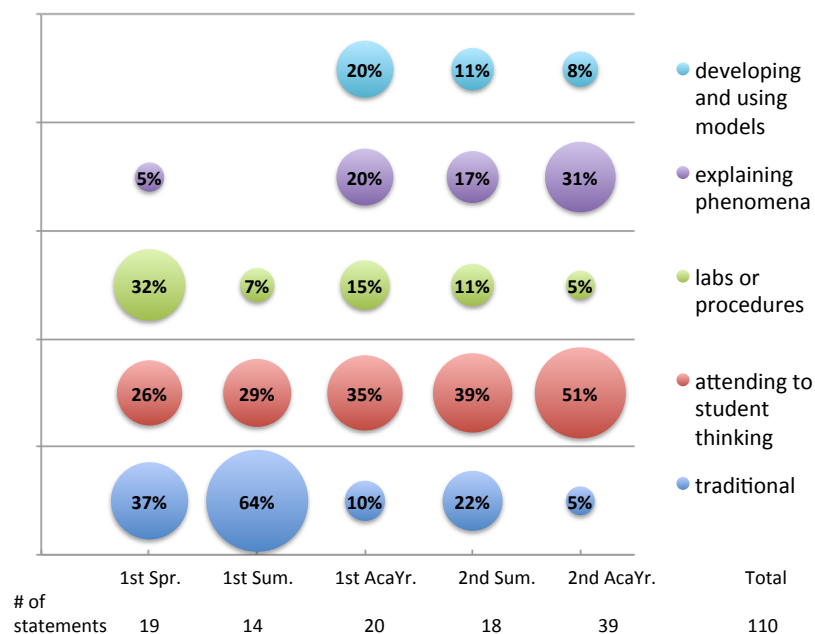


Figure 1. Quinn's self-described *instructional choices and pedagogical strategies* over time. Size of bubbles and number labels indicate percentage of statements assigned to each of the five codes within each time period.

Quinn's Conceptions of Student Learning and Science Teaching

Early in the PD, Quinn's statements about the learning process involve a number of different references to students' prior knowledge. She describes students as having "many misconceptions about science, specifically about chemistry" and that students must be given opportunities to "question what they think that they already know." For example, Quinn described how many students "assume there is only a chemical reaction if the result is an explosion" so she provided them with the chance to observe other kinds of chemical reactions in a series of "hands on" lab stations (180, 1st spr.).

She also described the importance of making connections between her instruction and students' prior knowledge. In general, she stated, "throughout lecture/direct instruction, I try to ask questions of students to access their own personal experiences and prior learning" (155, 1st sum.). Specifically, Quinn described how "teaching for understanding" in her classroom might happen when "I present the information using analogies to concepts that the students have in their prior knowledge or life experience (like when I compare the molecular movement and energy of solids, liquids, and gases to students in class (solid), at lunch (liquid), and after 3pm (gas)) (132, 1st spr.). However, Quinn also expressed dissatisfaction in the way her current practice aligns with her ideas about learning. "My teaching has evolved into something that does not reflect what I do know about prior knowledge and preconceptions," she wrote, stressing the mismatch between her current teaching and her ideas about learning. "Rarely do I find the time to examine students' prior knowledge," (135, 1st sum.).

Towards the end of the program, Quinn still expressed a concern over misconceptions. Her goal was to help students "eliminate all the previous theories," but described this process in a way that emphasized giving students time to think and talk to one another. For example, Quinn indicated that markers of learning for her now included being able to say that during a lesson, "at least they [students] were thinking, at least they were challenging what was being said by their peers, or what they thought themselves.... The kids learn more as opposed to just filling their heads with the facts and not actually committing it into their long-term memory." (238, 2nd acad. yr.).

Quinn's Conceptions of Scientific Models and the Role of Models in Science Teaching

Early in the program Quinn's responses reflect a number of different conceptions of the nature and utility of scientific models. Like many teachers Quinn described a model as "a plastic form representing an important structure of process in science i.e. a model of the water cycle or a cell model" (122, 1st spr.). She suggested that the role they play in teaching is as, "visual representations of structures that are not readily accessible to students-whether it be due to the lack of equipment necessary (like you can't see a cell without a microscope) or the lack of tangible attributes of a topic (for example, students who live in the suburbs may have difficulty understanding the rock cycle)." (123, 1st spr.) At the same time, Quinn stated that, "Models, like theories, are possible explanations for phenomena." (125, 1st spr.). In this role she described how models could be used in the classroom to, "Engage students and explore curiosity through understanding the cause-effect of phenomena" as well as to allow, "them to really try to formulate hypotheses and test them!" (129, 1st spr.).

In a science classroom, we use models everyday to teach students. After all, the models are what make up the ideas that we try to get our students to understand. However, the students and the teachers that often present the curriculum look at the models in the form of facts that are memorized and regurgitated through assessments and projects. This does not demonstrate a thorough UNDERSTANDING of the model, but instead, a memorization of the facts and definitions at hand. (emphasis original, 133, 1st spr.)

Later Quinn consistently referred to models as things that "allow us to develop explanations for natural phenomena (196, 1st acad. yr). A model, she says, is not "just a description," but rather something that is used for "explaining/known mechanisms" (221, 2nd sum.). What makes models useful in science classrooms is that they allow students "to create their own driving questions make observations, collect data, and develop the model," and, "more importantly, I would like them to reason with their model" (197, 1st acad. yr.).

Quinn's "MBR" framework: Coordinating student thinking and explanatory models

Early in the PD, Quinn described MBR as a process that "has participants come to their own conclusions to explain a phenomenon" (129, 1st sum), combining a constructivist ideas about learning with a focus on making sense of phenomena. We see this integration echoed throughout Quinn's responses. For example, at the end of the two-year PD, the teachers were asked to reflect on what they had learned from their participation. Quinn's response once again brings together ideas about learning and the role of models in science.

I don't need to be the source of information. I learned that I don't have to direct them to a source of information, that a lot of things that students gain through this process are in looking at or experiencing phenomena and reasoning out for themselves that they are the source of knowledge, and that they don't have to go somewhere else to find it, and that they can work together to come to the same answers they would have found in a textbook, and in the process eliminate all the previous theories that used to be presented in coming up with those scientific ideas. (238, 2nd acad. yr.)

We posit that it was the refinement and coordination of these ideas that constitute the learning underlying Quinn's described pedagogical shift. We see evidence of this in the way Quinn linked MBR to specific instructional strategies. In describing what it looks like to enact MBR in a science classroom, Quinn described "assessing constantly, to see oh, there's a new misconception, oh, this group is totally on the right path, and then you have another group that's totally on the wrong path, and it's okay, because eventually they'll get there." Quinn described how in her experience as a learner she often had moments of profound confusion, but that she saw these as an integral part of learning within this framework. Quinn stated that in her own classroom, "with my students I almost feel a sense of accomplishment when I leave their group and they seem just as confused as I did." Quinn also described how MBR means focusing her lessons around the following question: what do I want my students to be able to do and explain? (223, 2nd sum.)

In summarizing what she saw as the "essential elements" of the MBR approach, Quinn names three instructional heuristics: "1. Don't steal 'ah-hahs' 2. Make thinking visible and making thinking happen, 3. importance of a driving question" (249, 2nd acad. yr.). The first refers to the teaching strategy of providing space and time for students to come to their own understanding rather than telling. The second refers to a set of strategies, including asking students to draw and reflect, while prompting them with challenging questions and the third refers to organizing lessons around a central driving question about some phenomenon. Taken together these essential elements combine strategies meant to support student thinking as well as keep that thinking rooted to a scientific question.

Finally, we see evidence of Quinn reflectively applying the MBR framework to her past instruction. At the end of the program Quinn critiques her original lesson (Table 1) noting that while it did "address a preconception," it did not challenge students to "reason, take risks, or thinking deeper" and for this reason she

no longer believed it would be effective in changing student thinking. She also noted that the lesson, “asked for answers without explanation” (253, 2nd acad. yr.). In contrast, she described her new lesson as exemplary because it incorporated strategies such as using pictures and words to make students thinking visible and engaging students in explaining a phenomenon. She further described her role in the lesson as “facilitating,” not “giving answers” (252, 2nd acad. yr.).

Discussion

Using a Resources Framework to Make Sense of Quinn’s Shift

This case study relies solely on self-report data as evidence of a shift in how a teacher thinks about her pedagogy. While research has documented that descriptions of practice do not always reflect classroom enactment, we argue that the quality and consistency of the shift we see in Quinn’s responses represents an outcome worth exploring. We acknowledge however that more work is needed to see how these self-reports play out in the classroom.

Continuing from the assumption that we have captured a consequential change in how Quinn approaches her teaching, how can we make sense of Quinn’s shift over time in the program? One interpretation could have been to evaluate whether she attained the target knowledge about MBR and then attribute shifts to that knowledge acquisition. However, when we actually look at the data they tell a much more complicated story. If we model stable beliefs as mediating instructional practice, then two trajectories are possible: Either Quinn entered the PD with a set of sophisticated beliefs that allowed her to take up the messages of the PD program and change her practice, *or* the PD program was able to change Quinn’s core knowledge and beliefs about science teaching and learning. Our analysis of Quinn’s statements about student learning and scientific models do not support the interpretation that she had a stable set of core beliefs coming into the program.

A resources framework provides an alternative way to think about teacher learning in terms of building coherence around sets of productive resources. Looking across our analyses we can begin to make sense of the shift we saw in Quinn’s descriptions of her teaching. Quinn’s evolving understanding of MBR seems to have provided her with a way to coordinate, refine and enact ideas that she brought with her to the program. Quinn’s initial inclination to connect to students’ prior knowledge was refined into a strategy that allowed her students to make their thinking visible by generating public representations of their ideas. Rather than demonstrating or telling them the correct ideas, as she described doing early in the PD, Quinn described how to her, MBR means giving students more time to think, and allowing for, even valuing, frustration as part of the learning process.

Crucially, MBR also gave her a way to structure this thinking around an authentic scientific aim – to use models to explain natural phenomena. That is, it gave her students something to think *about*. Rather than demonstrating or directly teaching the scientific models to her students, the model ideas became the reasoning tools that her students would use to think about and explain phenomena. For Quinn, ideas about what it means to learn deeply and what it means to engage in authentic scientific reasoning appear to have come together under the MBR framework. That is, she came in with the inclination to attend to student ideas and an intention to engage students in science, but she admitted that she didn’t actually do much of that in her day-to-day practice. By the end of the program Quinn demonstrated that she was using MBR as a coherent framework and as a lens through which to view her pedagogy. In reflecting on her past lessons she was critical of both the quality of the science and the nature of her attention to student ideas. In describing her current aspirations, she described the importance of both choosing a phenomenon to investigate and ensuring that her students were the ones doing the thinking, while she played the role of facilitator. In her own analysis of her pedagogy, Quinn seemed to productively bring together and coordinate a number of ideas around student learning and scientific practice.

Implications for Professional Development

This PD program was not explicitly designed with a theoretical commitment to a resources view. Given the success of Quinn, what can we say about PD more broadly? The ISIM PD included many design features of good PD including: engaging teachers as learners in the MBR process, long duration, community building and coordination of theory with professional experimentation (Wilson, 2013). This work does not challenge these design decisions, but rather provides us with a way to reframe the way with think about teacher learning. For example, one of the objectives of the PD was to help teachers take up the MBR framework. We now have a better understanding of the mechanisms that may have led one teacher to take up many of our intended messages and a more robust understanding of what uptake even means. For Quinn it was not about merely acquiring or replacing existing conceptions of teaching and learning. What seems to have mattered to her was that the PD generally and the MBR framework specifically allowed her to coordinate and refine some of her existing ideas about teaching and learning. This work echoes the understanding that PD is situated in a much longer trajectory of teacher’s experiences. The PD aim should not be to completely overturn and replace teachers’ systems of beliefs. Rather than labeling teachers’ knowledge as correct, unproductive or transitional,

we can focus more specifically on the kinds of resources teachers have that can be built upon, coordinated and activated in particular contexts while helping them to reflect on their pedagogy and refine their practice.

As we go further into our data analysis and consider the implications for PD more broadly, we will need to also attend to the varied reasons why teachers may not have taken up these ideas. Increasingly, research on teacher learning suggests that merely targeting particular beliefs may not result in changes in classroom practice (Blanchard et al. 2009). Instead, what Franke et al. (1998) call “self-sustaining, generative change” seems to depend on the degree to which teachers have a coherent framework that can be used *flexibly* and thoughtfully to guide action. It is a more thorough understanding of the cognition of teachers in context that will be useful moving forward. More work is needed to understand how to systematically support teachers as they apply wide-ranging sets of cognitive and epistemological resources to the array of contexts in which they find themselves. In particular, the implementation of the Next Generation Science Standards in the U.S. creates the potential for a major perturbation of the status quo in science classrooms. Understanding how teachers develop, activate and cohere their new and existing cognitive resources around learning, teaching and scientific practice in service of aligning classroom practice to the new reform context will be critically important.

References

- Blanchard, M. R., Southerland, S. A., & Granger, E. M. (2009). No silver bullet for inquiry: Making sense of teacher change following an inquiry-based research experience for teachers. *Science Education*, 93(2), 322–360.
- Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, 44(4), 613–642.
- diSessa, A. A. (2002). Why “conceptual ecology” is a good idea. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: issues in theory and practice* (pp. 29–60). Boston: Kluwer Academic Publishers.
- Franke, M., 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.
- Friedrichsen, P., Van Driel, J., & Abell, S. K. (2011). Taking a Closer Look at Science Teaching Orientations. *Science Education*, 95, 358–376.
- Hammer, D. (1996). Misconceptions or P-Prims: How May Alternative Perspectives of Cognitive Structure Influence Instructional Perceptions and Intentions. *Journal of the Learning Sciences*, 5(2), 97–127.
- Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005a). Resources, framing, and transfer. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89–120). Information Age Pub Inc.
- Hammer, D., Russ, R., Mikeska, J., & Scherr, R. (2005b). Identifying inquiry and conceptualizing students’ abilities. In R. Duschl & R. Grandy (Eds.), *Teaching scientific inquiry* (pp. 138–156). Rotterdam, The Netherlands: SensePublishers.
- Kang, N.-H., & Wallace, C. S. (2005). Secondary science teachers’ use of laboratory activities: Linking epistemological beliefs, goals, and practices. *Science Education*, 89(1), 140–165.
- Markauskaite, L., & Goodyear, P. (2014). Tapping into the mental resources of teachers’ working knowledge: Insights into the generative power of intuitive pedagogy. *Learning, Culture and Social Interaction*.
- 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.
- Passmore, C., & Svoboda, J. (2011). Exploring Opportunities for Argumentation in Modelling Classrooms. *International Journal of Science Education*, (October), 1-20. doi:10.1080/09500693.2011.577842
- Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Stewart, J., Cartier, J., & Passmore, C. (2005). Developing Understanding Through Model-based Inquiry. In M. S. Donovan & J. Bransford (Eds.), *How Students Learn: History, Mathematics, and Science in the Classroom* (pp. 515–565). Washington D.C.: National Research Council.
- Wilson, S. (2013). Professional Development for Science Teachers. *Science*, 340, 310-313.
- 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*, 1–27.

Acknowledgments

This work was supported by NSF award #0554652. Opinions and findings are our own and not necessarily those of the NSF. We are grateful to the teachers in our program for opening their practice and thinking to us over the course of the two-year program. We also thank Arthur Beauchamp, Rich Hedman, Connie Hvidsten, Lin Xiang and Wendell Potter for their assistance in data analysis and for contributing to the conceptualization of this work more broadly.

The Impact of Text Genre on Science Interest in an Authentic Science Learning Environment

Steven McGee, The Learning Partnership, Westerns Spring, IL, mcgee@lponline.net

Amanda M. Durik, Northern Illinois University, DeKalb, IL, adurik@niu.edu

Dena Ann Pastor, James Madison University, Harrisonburg, VA, pastorda@jmu.edu

Abstract: A gap exists between research on learning and research on interest. Cognitive researchers rarely consider motivational processes, and interest researchers rarely consider cognitive process. However, it is essential to consider both since achievement and interest are in fact intertwined. In this paper we (1) discuss a theoretical model that intertwines cognitive and interest development, (2) describe how that model informed the development of educational materials, and (3) report on the results of the motivational components of a randomized research study examining the impact of text genre on interest. We hypothesized that students with low levels of interest would receive greater benefit from narrative text formats, whereas students with high levels of science interest would benefit more from expository text formats. The results of this research showed the opposite effect. Students with high levels of interest perceived the narrative texts as more interesting and useful for learning.

Significance of the Project

Research on the low levels of science proficiency and motivation, overall, as well as the findings on the gaps between White and African American/Hispanic students provide an important backdrop to the myriad of national reports making clarion calls for increasing the number of students who pursue careers in science (Robelen, 2010). The recommendations stemming from these reports tend to focus either on increasing student achievement in science or increasing student interest in pursuing a career in science. However, these reports tend to ignore research indicating that achievement and interest are in fact intertwined (Hidi, 1990). This body of research has shown that as student achievement in science increases, students tend to value science more. In parallel, the development of interest in science increases the selection of opportunities to further develop skills, such that individuals who value math and science are more likely to seek out additional math and science courses. For example, high school students who aspire to major in science tend to take more advanced science courses in high school (Simpkins, Davis-Kean, & Eccles, 2006), thus performing better on college achievement tests (Easton, Ponisciak, & Lupescu, 2008).

These findings suggest that both motivational and cognitive processes need to be examined together. In this paper we (1) discuss a theoretical model of interest development that intertwines learning and interest, (2) describe how that model informed the development of educational materials about the life history of a variety of consumer species in the El Yunque rainforest in Puerto Rico, and (3) report on the preliminary results of the motivational components of a randomized research study examining the impact of those materials on students' interest in science. Future reports will incorporate analyses related to the impact on learning outcomes.

Four-Phase Model of Interest

Hidi and Renninger (2006) articulated a phase-based model of interest development describing how instructional elements interact with both interest and cognitive development. They make a distinction between *situational* interest, which is supported primarily by the environment, and *individual* interest, which is driven by internal characteristics. In addition to describing the phases of the model, they specify which instructional supports are likely to facilitate interest development at each phase. A key element of the model is that the educational supports for situational and individual interests are not the same. Therefore, instructional enhancements will only promote interest if they are customized to learners' existing levels of individual interest.

In the case of situational interest, it is primarily the characteristics of instructional materials that trigger interest. Situations that are rich in novelty, surprise, and incongruity as well as group work, hands-on activities, fantasy contexts, and narrative texts can trigger situational interest, which is the first phase of situational interest (see Schraw & Lehman, 2001 for a review). Triggered situational interest is a short-term experience of focused attention and involvement that is supported almost entirely by features of the situation (Hidi & Renninger, 2006). If individuals become sufficiently involved in the material and find the content meaningful, they might experience maintained situational interest, the second phase of situational interest, and continue to engage in the material or choose to enter very similar situations in the future.

On the other hand, the phases of individual interest are supported more by factors inside the person. Individuals with individual interest have some stored knowledge of the content area and value that knowledge. They have also developed some fluidity in the way they work with domain content. However, individual interest still requires some external support. Specifically, the presence of factors such as expert models, challenging

goals, and opportunities for knowledge building can promote the development of individual interest.

Given this framework and the variability between instructional supports that enhance situational versus individual interest, a critical question is how to design instructional materials to support the development of interest for all learners. An assumption underlying most instructional materials and most research on the factors that can promote interest is that “one size fits all.” In most of the prior experimental research testing the effects of situational factors on interest, pre-existing individual interest, if measured, is not tested as a moderator of the effects of situational factors (e.g., Harp & Mayer, 1997). In other words, there is an assumption that learners will respond to situational enhancements with greater interest regardless of their level of individual interest.

Consistent with the four-phase model, evidence is beginning to accumulate that the effects of situational enhancements work differently depending on individuals’ pre-existing orientations towards the task (Hulleman & Harackiewicz, 2009). For example, learners with low individual interest in math, showed higher situational interest after learning the material with colorful versus plain instructional materials, whereas learners with high individual interest in math showed the opposite pattern. In addition, learners with high individual interest in math showed higher task interest after learning how a new math technique could be personally useful to their lives (Durik & Harackiewicz, 2007). In other words, the effect of situational enhancements designed to raise interest in educational contexts seem to work best if the type of enhancement is paired with learners’ existing level of individual interest.

Implication of Four-Phase Model of Interest on the Design of Instructional Texts

The four-phase model of interest development suggests that students’ interest in any given learning task is a function of the task characteristics and students’ prior individual interest in the subject of study. The genre of a reading passage is a task characteristic that has been shown to influence interest. Research in which both expository and narrative texts are included often shows that readers find narrative texts more interesting than expository texts (Dai & Wang, 2007). Several characteristics of narrative texts have been identified to increase text interest (see Schraw & Lehman, 2001), such as, rich detail about concepts and ideas (Schraw, 1997), as well as suspense and resolution (Iran-Nejad, 1987). Coherence is also a critical variable related to text interest because texts that readers perceive to be coherent are more interesting (Sadovski, Goetz, & Rodriguez, 2000).

On the other hand, narrative texts can be more challenging than expository texts in supporting reading to learn (Lee & Spratley, 2010). Expository texts tend to be organized around the structure of the subject matter to be learned, whereas narrative texts can lead readers astray from the main ideas conveyed in a text (Kintsch, 1998). Consistent with this, Harp and Mayer (1997) measured two kinds of self-reported situational interest following a reading activity in order to separate the affective and cognitive value of texts. They found that seductive details, which are common in narrative texts, increased affective interest (enjoyment) but decreased cognitive interest (utility). In this study, we also examine the effects of text characteristics (i.e., narrative or expository genre) on cognitive and affective situational interest. We also expand on prior research by testing whether the effects of genre differ depending on initial levels of individual interest.

Design of Instructional Texts

The four-phase model of interest development and the empirical research on the impact of text characteristics on learning and interest provide a framework for the design of instructional materials that can be used to enhance interest and learning at different phases of interest. For students who are low in science interest, it is important to maximize the instructional enhancements that will promote interest and engagement. For students who are high in individual science interest, it is important to provide students with challenging material.

In this paper, we describe an effort to modify the background materials in a middle school, inquiry-based curriculum unit called *Journey to El Yunque* (<http://elyunque.net>). We have developed two parallel forms (narrative and expository) of the background readings. The narrative form describes the life history of each species by presenting anthropomorphic characters that show students key limiting factors through vivid imagery and rich descriptions. The narrative version contains the text-based features that prior research has shown to trigger situational interest in students with low individual interest. In contrast, the expository versions present the same key limiting factors as in the narrative form, but the information is conveyed as a general description. The expository version minimizes the text-based features that prior research has shown to be distracting for students with high individual interest. We will now describe the context of the *Journey to El Yunque* program and outline the creation of parallel narrative and expository versions of the background readings.

The *Journey to El Yunque* program aims to improve middle school students’ understanding of the dynamic interrelationships among organisms and the environment. The four-week, web-based unit exposes students to authentic research practices at the Luquillo Long-Term Ecological Research station in Puerto Rico, commonly known as El Yunque. It is the only tropical rainforest among 25 sites within the overall Long-Term Ecological Research network. El Yunque is the site of some of the earliest ecosystem-level studies in ecology (Odum, 1970). More recently, El Yunque has been struck by two severe hurricanes (Hurricane Hugo in 1989 and Hurricane Georges in 1998). Researchers in El Yunque have provided comprehensive studies of the

resilience of ecosystems to natural disturbances like hurricanes. The research surrounding these hurricane recovery periods provides a rich example of basic ecosystem processes at work.

Journey to El Yunque engages students in the same problems that researchers in El Yunque are investigating. The program consists of four modules that support the overarching goal of investigating what will happen to the rainforest if severe hurricanes end up striking the rainforest more frequently, as suggested by climate models. The introductory module introduces students to the historical patterns of hurricane frequency and damage in El Yunque. In the second module, students investigate what happens to the producers in El Yunque after a hurricane. In the third module, students investigate what happens to the consumers in El Yunque after a hurricane. In the fourth module, students explore the impact of a hurricane on an entire food chain.

The background readings in this study come from the third module on consumers. Prior to engaging in modeling activities, students are presented with background readings about the life history of each species. The program introduces students to six consumer species that are representative of the types of hurricane responses researchers have found in El Yunque. There are two decomposers, two primary consumers, and two secondary consumers. As part of the investigation cycle, students read about the life history of their assigned consumer, make a prediction about the population dynamics after hurricane Hugo, and explore a model of population dynamics to explain what happened to their species after hurricane Hugo.

When investigating population dynamics after a disturbance, scientists in El Yunque typically consider changes to five primary limiting factors for a given species: access to prey, avoidance of predators, direct mortality from the hurricane, and suitability to changes in environmental conditions, in particular drought conditions and the influx of forest debris from the hurricane. Each species has five background readings focused on providing the life history of each limiting factor. During the modeling activities, students manipulate parameters related to these five limiting factors and investigate the impact on population dynamics.

Parallel Readings

In order to address the need to have instructional texts that can accommodate students with either low or high individual interest, we developed parallel background readings (narrative or expository) for each of the five limiting factors for each of the six species. There are three dimensions on which we ensured that the readings were parallel: reading complexity, word length, and idea units. We developed the readings to be at a 5th grade reading level to ensure that the readings would be within the reading range of most middle school students. We used the Lexile text score to measure reading complexity (Smith, Stenner, Horabin, & Smith, 1989). The Lexile bases text difficulty on the average sentence length and the complexity of the text vocabulary. In classroom settings, teachers can assess students' reading abilities on the same Lexile scale. Therefore, students' reading abilities can be matched to text difficulty. The interquartile Lexile range for 5th grade students is 565L – 910L(1). Books that are more than 50L above students' Lexile scores are considered too complex for students to read (2). Therefore, each of the background readings was developed to have a Lexile score between 565L and 910L and the Lexile scores for the parallel versions were within 50L points of each other to be considered equivalent.

At the end of 5th grade, the average student's reading fluency for a 5th grade text is 139 words correct per minute (Hasbrouck & Tindal, 2006, Table 1, p. 69). Based on that result, we assumed that most middle school students would be able to read a 5th grade text at 125 words per minute. Therefore, five 500-word essays would take students about 20 minutes to complete. We limited the word length differences between the parallel versions to be within 12 words for each reading, which would correspond to a 30-second differential for the average student when reading about all five limiting factors for one species.

In order to ensure that the readings had parallel content, we first outlined the key concepts for each species. These content outlines were provided to a professional fiction author. He created narrative, anthropomorphic characters and situations to convey the life history of the species. Next, the narrative stories and the content outline were provided to a science textbook writer, who developed a parallel expository version that was within the word length and reading complexity parameters of the narrative version. Lastly, the narrative and expository versions were reviewed by an ecologist who studies El Yunque to ensure that the content was scientifically accurate and the content was parallel across both versions. Through this process we developed five narrative and five expository readings for each species that are parallel in text complexity, word length and idea units. Therefore, the primary difference between the parallel versions is the genre — narrative or expository.

Method

The study was conducted during the 2012-13 school year in 20 seventh and eighth grade classrooms from 7 different middle schools in the Midwest. Five of the seven schools were urban elementary schools with greater than 90% of the students eligible for free or reduced lunch. One school was a Catholic elementary school situated in an urban neighborhood with a 20% poverty rate. The remaining school was a suburban junior high school with 0% of the students eligible for free or reduced lunch.

The intervention took place over five class periods. During the first class period, students completed pre-surveys and on the fifth class period, students completed post-surveys. During the middle three days of the intervention, students studied three of the six species—one species on each day (randomly assigned at the classroom level to ensure generalizability across species). Students within classrooms were randomly assigned to read either narrative or expository versions of the readings. In order to ensure that levels of prior individual interest in ecology were evenly distributed across conditions, we did blocked random assignment on students' prior individual interest (see the Measures section below for a description of the prior interest survey).

The delivery of the educational materials was conducted through a web-based system. Once a student was assigned to a treatment condition, an account was created for the student and the account was assigned to readings associated with the assigned experimental condition. The experimenter provided students with their login information at the start of class period 2. Class periods 2-4 took place in a computer lab setting in which participants in the same class were reading about the same species, but presented in different genres. All of the experimental instructions were delivered via text. The experimenter monitored the students as they participated in the activities and took note of any behavior that fell outside of the study protocol, such as students talking with each other, obvious distractions, and technical problems.

For class period 2, students logged in to the system and went through an initial 5-minute introduction and training on the system. Then, the system presented students with the activities for the first of the three species they had been assigned. The timing and structure of the activities was the same regardless of which genre students were presented based on their experimental condition.

The students were instructed that they had 20 minutes to read the articles associated with the species with the goal of learning about the factors that affect the size of the population and then applying that information to make a prediction about what happened to that species after Hurricane Hugo. At the end of 20 minutes, the students were automatically moved to the next step. If students finished early they could move on to the next step. We controlled for time on task by obtaining the actual time spent on the reading task, which could be used as a control variable in the analyses. Students were able to take notes on a paper-based worksheet that contained a column for the limiting factors included in the reading and a column for taking notes about that limiting factor. The paper notes could be used during subsequent activities related to the species.

Once the students completed the reading and note-taking task, they completed a comprehension task by summarizing the main factors that affect the survival of the species. After completing the comprehension task, we gauged students' situational interest by first asking them to describe anything they were curious about. Next, students responded to situational interest survey questions (see measures below). After completing the situational interest survey, the system presented students with the prediction activity, which is an application of their reading comprehension. Students were instructed to use what they learned from the background readings to draw a prediction of what the graph of the population looked like after Hurricane Hugo. The students used the interactive graphing tool in *Journey to El Yunque* to draw a graph predicting what the populations levels would be each quarter over a 60-month period. After the students completed their drawing, they developed and submitted a justification for their prediction. Once they completed the graph and submitted the justification, the program overlaid a graph of the actual population on their graph of the predicted population. The students then generated a hypothesis that used what they had learned about the species from the background reading to explain how changes in the limiting factors affected the population. The students repeated the same protocol on class periods 3 and 4 for the second and third species they been assigned respectively. In this paper, we focus on the interaction of prior individual interest and genre on students' situational interest.

Measures

The pre- and post-surveys contain measures of individual interest and general ecology knowledge. The pre-survey also measures students' reading fluency and collects demographic information. Since this study focuses on individual and situational interest, the individual and situational interest measures are described below.

Self-report measure of individual interest. Individual interest in ecology was measured on the pre-survey with 6 items that were adapted from prior research (Harackiewicz, Durik, Barron, Linnenbrink-Garcia, & Tauer, 2008). Half of the items focused on feelings associated with the domain (e.g., "How much are you interested in ecology?") whereas the other half focused on meaningfulness of the domain (e.g., "How important is ecology as a subject?"). The response options ranged from 1 (*Not at all* ___) to 7 (*Very* ___) and the anchor labels were tailored to each item (e.g., *Not at all interested* to *Very interested*). Prior to answering the questions, students were given a basic definition of ecology. The reliability for the scale for the current sample, as measured by Cronbach's alpha, was 0.80.

Self-report of cognitive and affective situational interest. Two types of situational interest were measured on each day after students completed the reading tasks (Harp & Mayer, 1997). Cognitive situational interest refers to the extent to which individuals believe that the learning materials helped them learn the content (e.g., "How much did the readings help you understand this topic?"). Affective situational interest reflects students' beliefs that the material was interesting and entertaining (e.g., "How interesting was the material in the

readings?”). Participants rated each item on the same 7-point scale as used for individual interest, and each construct was tapped with two items. The reliability estimates for cognitive and affective interest for each day were high, with Cronbach’s alphas exceeding 0.84 for each scale on each day.

Population

There were 488 students who participated in the intervention. There were 321 students included in the analyses. Students were dropped from the analyses if they did not provide informed consent or were absent on the day of the pre-survey. Of the participating students a little over half were girls (52%). The largest racial group was Hispanic (57%), followed by Caucasian (33%), and African American (18%). Around 6% of the students indicated membership in other racial groups. The percentages add up to more than 100% since students were able to indicate more than one racial group. Students who were designated as special education or English Language Learners (6%) were also included in the study. In addition to blocking on prior interest, students were also blocked during the random assignment process based on their special population status.

Results

In our analyses, we explored the impact of prior individual interest and genre on students' cognitive and affective situational interest related to the background readings about the species. The overall averages for both cognitive and affective situational interest was above the scale midpoint (4.0) on each day, indicating in general that students found the readings enjoyable and useful for learning. Prior to examining the interaction effect, we explored the main effect of three factors: group of readings, genre, and prior individual interest.

The first factor we examined was whether there were any differences in cognitive or affective situational interest depending on which species the students read about. The six species were organized into two groups of readings (A or B). If there are minimal differences between the groups then we can combine the two groups for subsequent analyses. We conducted separate ANOVAs for cognitive and affective situational interest for each day using group as the independent variable (Tables 1 and 2 report the means and standard deviations). For affective situational interest, there were no statistical differences between the two groups on any of the three days of the intervention (see Table 1). For cognitive situational interest, there were no statistical differences between the two groups on the first two days of the intervention (see Table 2). However, on the third day, students found the snails readings slightly more beneficial for learning than the mushroom readings, with a small effect size, $F(1,300) = 3.94, p < .05, d = 0.2$. Given that there was only a small difference on one dimension of situational interest for one of the species, we combined the two groups for all subsequent analyses.

Table 1: The effect of group, genre, and prior individual interest on affective situational interest

Day	Group		Genre		Prior Interest Slope (R^2)
	Group A	Group B	Narrative	Expository	
1	4.3 (1.5)	4.5 (1.4)	4.3 (1.4)	4.5 (1.5)	0.09* (20%)
2	4.7 (1.5)	4.6 (1.5)	4.4 (1.6)	4.8 (1.4)	0.13* (36%)
3	4.4 (1.6)	4.5 (1.5)	4.5 (1.6)	4.4 (1.5)	0.13* (35%)

Table 2: The effect of group, genre, and prior individual interest on cognitive situational interest

Day	Group		Genre		Prior Interest Slope (R^2)
	Group A	Group B	Narrative	Expository	
1	4.8 (1.3)	5.0 (1.1)	4.8 (1.2)	5.0 (1.2)	0.07* (14%)
2	5.0 (1.3)	5.0 (1.2)	4.8 (1.3)	5.2* (1.2)	0.08* (22%)
3	4.7 (1.4)	5.0* (1.2)	4.8 (1.4)	4.9 (1.3)	0.08* (21%)

The second factor we examined was whether there was a main effect of genre on cognitive or affective situational interest related to the background readings. We conducted separate ANOVAs for cognitive and affective situational interest for each day using genre as the independent variable (Tables 1 and 2 report the means and standard deviations). For affective situational interest, there were no statistical differences between the narrative and expository readings on any of the three days of the intervention (see Table 1). For cognitive situational interest, there were no statistical differences between the narrative and expository readings on the first and third days of the intervention (see Table 2). However, on the second day, students found the expository readings slightly more beneficial for learning than the narrative readings, with a small effect size, $F(1,285) = 7.8, p < .01, d = 0.3$. Given that there was only a small difference on one dimension of situational interest for one of the days, we concluded that there was no main effect of genre on situational interest. Given our hypothesis that there is an interaction effect between prior individual interest and genre, it is possible that the effects of genre are masked by the interaction.

The third factor that we examined was whether there was a main effect of prior individual interest on cognitive or affective situational interest related to the background readings. We conducted separate regressions

on cognitive and affective situational interest for each day using prior individual interest as the independent variable. Tables 1 and 2 show the beta coefficients and R^2 values for each regression. Overall, prior individual interest in ecology was a strong predictor of both cognitive and affective situational interest related to the background readings on each of the three days of the intervention. Students with higher levels of individual interest found the readings to be more enjoyable and useful for learning than students with lower levels of individual interest.

In order to analyze the interaction between genre and prior individual interest on situational interest, there are two primary issues that negate the appropriateness of conducting separate analyses for each day using linear modeling. First is that days are not independent of each other. Days are embedded within students. Students' experiences on one day may impact their experiences on subsequent days. Second, due to absences from one day to the next, there is variation in the number of observations on each day. A dataset that only includes students who were in attendance on all three days would be significantly reduced. Hierarchical linear modeling (HLM) addresses both of these issues. HLM analysis constructs a growth model for each student, which includes an intercept and a slope for each student. These individual growth models can be constructed even if students were not in attendance on all three days. These intercepts and slopes for each student can then be analyzed at the student level for interactions between prior individual interest and genre. Tables 3 and 4 show the results of the HLM analyses for affective and cognitive situational interest. For ease of presentation, we computed the expected values of the intercepts and slopes for students whose prior individual interest was at the average of the population (medium), one standard deviation above the average (high) and one standard deviation below the average (low). The growth models were centered at the first day of the intervention, so that the intercepts represent the expected values on the first day of the intervention.

Table 3: The interaction of prior individual interest and genre on the intercepts and slopes related to affective situational interest

Prior Individual Interest	Narrative		Expository	
	Intercept	Slope	Intercept	Slope
Low	3.7	-0.12	4.0	-0.12
Medium	4.4	0.09	4.7	-0.07
High	5.1	0.30	5.4	-0.01

Table 4: The interaction of prior individual interest and genre on the intercepts and slopes related to cognitive situational interest

Prior Individual Interest	Narrative		Expository	
	Intercept	Slope	Intercept	Slope
Low	4.3	-0.15	4.5	-0.07
Medium	4.8	-0.002	4.9	-0.08
High	5.4	0.15	5.2	-0.09

There is a statistically significant interaction effect between prior individual interest in ecology and genre of the background readings. Students with higher levels of prior individual interest tended to rate the narrative readings higher by the end of the intervention than the expository readings and students with lower levels of prior individual interest tended to rate the expository readings higher by the end of the intervention than the narrative readings. Figures 1 and 2 provide a graphical representation of these results. The intercepts for affective situational interest show a positive main effect for expository over narrative and a positive main effect for prior individual interest, but no interaction effect. The intercepts for cognitive situational interest show a main effect for prior individual interest in ecology and an interaction effect with genre. Students with lower levels of prior individual interest rated the expository readings more helpful for learning whereas the students with higher levels of prior individual interest rated the narrative readings as being more helpful for learning.

The slopes for both affective and cognitive situational interest show an interaction effect. In the expository condition, students' ratings of both cognitive and affective situational interest generally decreased over the course of the three-day intervention. However, in the narrative condition, the change in situational interest over time depended on students' prior individual interest in ecology. Students in the narrative condition with high prior individual interest showed positive growth in both cognitive and affective situational interest to the point that they rated their narrative readings higher than the high interest students in the expository condition. In contrast, students with low prior interest in both the narrative and expository condition showed negative growth in both affective and cognitive situational interest over the three days.

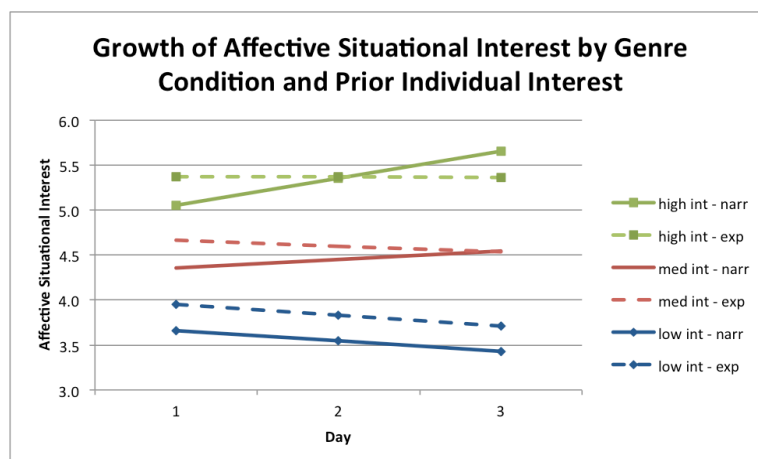


Figure 1. Interaction effect of prior individual interest and genre on affective situational interest

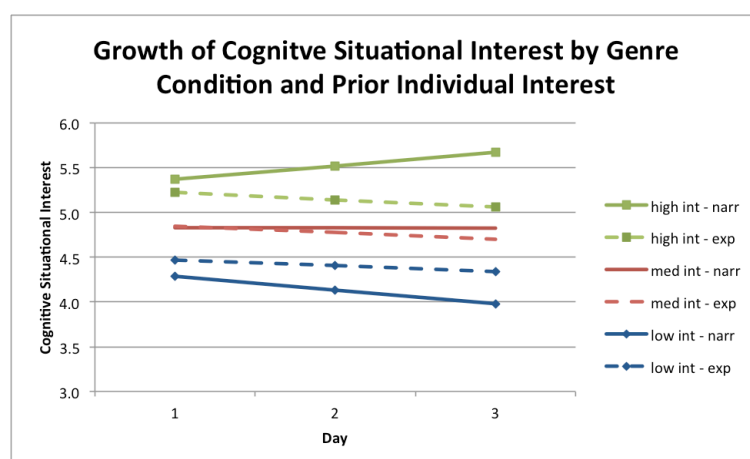


Figure 2. Interaction effect of prior individual interest and genre on cognitive situational interest

Conclusion

The four-phase model of interest development (Hidi & Renninger, 2006) suggests that students' interest in any given learning task is a function of the task characteristics and students' prior individual interest in the subject of study. Task interest for students with low individual interest in a subject area is increased through task enhancements that are outside of the subject of study. Prior research on text characteristics has found that narrative texts with rich descriptions are generally more interesting than expository texts (Dai & Wang, 2007), and could be used as enhancements to trigger situational interest in a science text among students with low individual interest. On the other hand, for students with high individual interest in the subject area, creating challenging learning tasks enhances task interest. Narrative texts can be more challenging in a scientific context than expository texts because they can lead readers astray from the main ideas conveyed through a narrative text (Kintsch, 1998).

The results of this study support the conclusion that students' task interest in reading about the life history of species in El Yunque is dependent on their individual interest in ecology and the characteristics of the text. We attempted to vary text characteristics so as to optimally support the task interest of students who were low or high in individual interest in ecology. For students who were low in individual interest in ecology, both the narrative and expository readings triggered at least some level of interest at first (3.5-4.0 out of 7 on affective interest and 4.0-4.5 out of 7 on cognitive interest). However, neither genre of reading was able to maintain the triggered interest of students who were low in individual interest in ecology. The ratings of situational interest for both narrative and expository decreased as the intervention progressed. These results are not consistent with prior research indicating that narrative texts tend to be more interesting than expository texts. Further work needs to be done to uncover the characteristics of science tasks that will both trigger and maintain situational interest for students who are low in individual interest in science.

On the other hand, for students who were high in individual interest in ecology, both the narrative and expository readings triggered high levels of interest at first (5.0-5.5 out of 7 on both affective and cognitive interest). As was the case for students low in individual interest in ecology, the expository readings were not

able to maintain the triggered interest of students who were high in individual interest. The ratings of situational interest for the expository readings decreased as the intervention progressed. In contrast, the narrative readings did maintain and even increase the triggered situational interest of students who were high in individual interest. When viewing these results through the lens of prior research that characterizes the challenge of extracting important ideas from narrative texts, these results seem to be consistent with the four-phased model of interest development. Hidi and Renninger (2006) suggest that challenging tasks heighten the level of task interest for those students who are already high in individual interest in a subject area. These results raise interesting possibilities about the use of narrative texts in science as a means to challenge students with high individual interest in science.

Endnotes

- (1) <http://lexile.com/about-lexile/grade-equivalent/grade-equivalent-chart/>
- (2) <http://lexile.com/using-lexile/lexile-at-home/reading-outside-of-your-lexile-range/>

References

- Dai, D. Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology*, 32(3), 332-347.
- Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology*, 99(3), 597-610.
- Easton, J., Ponisciak, S., & Luppescu, S. (2008). *From High School to the Future: The Pathway to 20*. Chicago: Consortium on Chicago School Research at the University of Chicago. Retrieved March 23, 2011 from http://ccsr.uchicago.edu/content/publications.php?pub_id=126.
- Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology*, 100(1), 105-122.
- Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, 89, 92-102.
- Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. [Article]. *Reading Teacher*, 59(7), 636-644.
- Hidi, S. (1990). Interest and Its Contribution as a Mental Resource for Learning. *Review of Educational Research*, 60(4), 549-571.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111-127.
- Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, 326, 1410-1412.
- Iran-Nejad, A. (1987). Cognitive and Affective Causes of Interest and Liking. *Journal of Educational Psychology*, 79(2), 120-130.
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. New York: Cambridge University Press.
- Lee, C. D., & Spratley, A. (2010). *Reading in the Disciplines: The Challenges of Adolescent Literacy*. New York: Carnegie Corporation. Retrieved March 18, 2011 from http://carnegie.org/fileadmin/Media/Publications/PDF/tta_Lee.pdf.
- Robelen, E. (2010, September 14). Advisory Panel to Offer Obama Ideas for Advancing STEM Education. http://blogs.edweek.org/edweek/curriculum/2010/09/stem_ed_conference.html
- Sadoski, M., Goetz, E. T., & Rodriguez, M. (2000). Engaging tests: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology*, 92, 85-95.
- Schraw, G. (1997). Situational interest in literary text. *Contemporary Educational Psychology*, 22, 436-456.
- Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, 13, 23-52.
- Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. L. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology*, 42, 70-83.
- Smith, D. R., Stenner, A. J., Horabin, I., & Smith, M. (1989). *The Lexile Scale in Theory and Practice. Final Report*. Washington, DC: MetaMetrics.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A110810 to The Learning Partnership. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

When Experts Disagree: Sourcing Practices While Reading Conflicting Online Information Sources

Sarit Barzilai, University of Haifa, 199 Abba Hushi Blvd., Haifa, Israel
Eynav Tzadok and Yoram Eshet-Alkalai, The Open University of Israel, 108 Ravutsky St., Ra'anana, Israel
Email: sarit.barzilai@edu.haifa.ac.il, eynav150@hotmail.com, yorames@openu.ac.il

Abstract: Laypeople who use the Internet to learn about issues of personal or social relevance often encounter online information sources that present conflicting expert accounts. The aim of the current study was to provide a close observation of spontaneous sourcing practices while reading conflicting online information sources, to examine the relation between sourcing while reading and subsequent argument construction, and to assess the role of epistemic perspectives, topic interest, and topic knowledge in sourcing. 61 university students thought aloud while reading four blog-posts that provided conflicting accounts of a socio-scientific controversy. The findings revealed a wide range of sourcing practices. High sourcing participants made more sourcing activities, paid more attention to source characteristics, and made source-source comparisons. Higher levels of sourcing were found to be related to subsequent argument complexity. Epistemic perspectives and gender played a significant role in sourcing practices and highlighted their socio-cultural nature.

Introduction

The ubiquity of Internet and Web 2.0 technologies has created a contemporary learning challenge: On the one hand, information technologies have dramatically increased the abundance and variety of information sources that are available to laypeople who wish to learn about current issues. On the other hand, the diverse and conflicting nature of Internet sources often makes it difficult to make sense of their claims and to reach informed decisions. Effectively dealing with online information necessitates, therefore, acquiring digital literacy strategies, such as the ability to critically evaluate and integrate multiple information sources (Eshet-Alkalai, 2004; Goldman & Scardamalia, 2013). A crucial aspect of evaluation and integration strategies is the ability to form connections between sources and their contents: Knowledge construction from multiple information sources entails attributing information to specific sources and paying attention to the ways in which source characteristics, such as identity, expertise, perspectives, motivation, and credibility, impact the information presented (Rouet, 2006; Wineburg, 1991). However, multiple studies indicate that learners often ignore or downplay source characteristics and tend to focus on medium and content characteristics (Britt & Aglinskas, 2002; Gasser, Cortesi, Malik, & Lee, 2012).

Sourcing involves attending to and evaluating source information and using source information to interpret the content (Bråten, Britt, Strømsø, & Rouet, 2011; Britt & Aglinskas, 2002). Wineburg (1991) noted that sourcing provides anticipatory frameworks for the subsequent encoding of text. Although sourcing has been often studied in the context of reading historical documents, recent studies have focused on sourcing processes with varied types of documents, including online information sources (e.g., Goldman, Braasch, Wiley, Graesser, & Brodowska, 2012; Strømsø, Bråten, Britt, & Ferguson, 2013). Sourcing studies often present learners with information sources of varying levels of expertise and authority, thereby highlighting evaluation of author trustworthiness. In the current study we were interested in examining what happens when learners encounter conflicting information sources written by authors of comparable expertise and authority. Such examination may help improve the understanding of how online readers process and evaluate source information as they learn about socio-scientific controversies. Additionally, the study examined whether or not sourcing matters in terms of argument construction and how sourcing might be related to learners' views of knowledge and knowing, to their topic interest, and to their topic knowledge.

Multiple Document Comprehension

The challenge of comprehending multiple information sources stems from the fact that these information sources often do not provide a single coherent account of the issue or situation at hand, but rather may provide partially or wholly inconsistent accounts (Rouet, 2006; Wineburg, 1991). For example, online searching about socio-scientific controversies, such as cell phone use or global warming, can yield numerous search results in which experts provide conflicting accounts and explanations. Therefore, source expertise and authoritativeness are necessary but insufficient criteria for evaluating and integrating online information sources about such controversies. Deeper understandings of author backgrounds and perspectives and of the supports they offer for their claims are needed. The *Documents Model* (Bråten et al., 2011; Britt & Rouet, 2012; Rouet, 2006) describes the structures and mechanisms of multiple document comprehension. In brief, the Documents Model proposes that readers construct an *Integrated Mental Model*, which integrates information across documents, as well as an

Intertext Model, which represents the documents and their relations. Constructing the Intertext Model involves forming a representation of each document (i.e. author and document characteristics, rhetorical goals, and key claims or position), creating source-content links (i.e. connecting source information to content and using that information to interpret the content), and creating source-source links between documents (i.e. comparing documents and noting similarities and differences). Although the Documents Model has been used as a framework in multiple studies, there have not been, to the best of our knowledge, many attempts to examine the model *in situ* using think-aloud techniques (cf. Strømsø et al., 2013). In this study we focused on examining how learners construct Intertext Models by analyzing spontaneous references to source representations, source-content links, and source-source links while reading.

Contributors to Multiple Document Comprehension

Several contextual and individual factors may contribute to multiple document comprehension. Source and content characteristics have been found to impact attention to and memory for sources. Specifically, conflicts between information sources have been found to increase source salience, source comprehension, and memory for sources (Braasch, Rouet, Vibert, & Britt, 2012). Interest in the topic may lead people to pay closer attention to messages, to evaluate their arguments and sources more carefully and critically, and thus improve document comprehension and memory for sources (Strømsø, Bråten, & Britt, 2010). Disciplinary expertise and topic knowledge may also affect learners' capacity to evaluate sources critically (Wineburg, 1991). Finally, learners' personal epistemology, that is, their thinking about knowledge and about how people know, has also been found to have an impact on source evaluation and multiple document integration (Barzilai & Zohar, 2012; Bråten et al., 2011).

In the current study, personal epistemology was conceptualized using a model proposed by Kuhn and her colleagues (Kuhn & Weinstock, 2002). In brief, this model distinguishes between three main epistemic perspectives: (1) absolutism— knowledge is perceived as objective, absolute, and certain; (2) multiplism — knowledge is perceived as subjective, relative, and uncertain; and (3) evaluativism— knowledge is perceived as combining subjective and objective aspects, and although it is uncertain can be evaluated and improved. Learners' epistemic thinking develops with age and schooling and is shaped by disciplinary standards as well as by socio-cultural values (Tabak & Weinstock, 2008). Learner's epistemic thinking includes both metacognitive knowledge, skills, and experiences regarding the nature of knowledge and knowing, as well as cognitive strategies and processes for reasoning about the epistemic characteristics of specific information, knowledge claims, and their sources (Barzilai & Zohar, 2014). Thus, reasoning about the reliability of specific information sources and about the justification and coherence of their claims are cognitive-level epistemic processes (epistemic cognition). Whereas learners' views regarding the certainty, sources, structure, and justification of knowledge and knowing are a meta-level epistemic knowledge (epistemic metacognition). Learners' epistemic metacognitive knowledge is proposed to interact with their cognitive-level epistemic processes (Barzilai & Zohar, 2012, 2014).

The Present Study

The goal of the present study was to observe and analyze learners' spontaneous sourcing practices as they read online information sources that present conflicting expert accounts regarding a socio-scientific controversy. In addition, the study examined some of the predictors and products of sourcing.

The study questions were:

1. To what extent do learners spontaneously attend to source information and process that information while reading conflicting online information sources? Specifically, do learners spontaneously form source representations, source-content links, and source-source links, while reading?
2. What are the characteristics of learners' sourcing practices?
3. Is sourcing while reading related to subsequent argument construction?
4. Is sourcing while reading related to learners' epistemic perspectives, topic interest, and topic knowledge?

Method

Participants

Participants were 61 Hebrew-speaking students from a distance learning university (64% female; 84% BA students; 16% MA students; mean age = 30.50 years, *SD* = 8.44).

Materials

The topic chosen for the study was the socio-scientific controversy regarding whether or not Israel should develop an extensive seawater desalination system to supply growing water demands. We designed several short blog-posts that were all based on authentic online sources. The blog-posts were of similar length, writing style,

and author expertise. All authors were presented as owning a doctoral degree and working as consultants to government departments and agencies. However, the authors presented conflicting arguments for and against desalination from different disciplinary perspectives (economic and environmental). Thus, in contrast to many previous sourcing studies, the sources were of similar expertise and authority, thereby focusing the attention on differences between source perspectives and positions.

Measures

Scenario-Based Epistemic Thinking Assessment: Learners' epistemic perspectives were measured using a scenario-based assessment that related to the desalination context (Barzilai & Weinstock, 2014). The assessment is a topic-specific measure of learners' epistemic metacognitive knowledge. Participants were asked to rate their agreement with typical absolutist, multiplist, and evaluativist views, on a scale from one to ten (e.g., an absolutist statement was "eventually there will be one right answer"). Internal consistency reliabilities of the scales were: absolutism, 10 items, $\alpha = .87$; multiplism, 8 items, $\alpha = .85$; and evaluativism, 10 items; $\alpha = .83$. A score for each position was calculated based on the item mean. Mean scores were: absolutism, $M = 7.49$, $SD = 1.54$; multiplism, $M = 3.39$, $SD = 1.66$; and evaluativism, $M = 6.76$, $SD = 1.62$.

Topic Interest: Topic interest was assessed using a 10-item questionnaire by Mason, Gava, and Boldrin (2008) that was translated to Hebrew and adapted to the task topic. Items were scored on a six-point scale. Internal consistency reliability was $\alpha = .92$. The topic interest score was based on the item mean, $M = 3.76$, $SD = 1.03$.

Topic Knowledge: We developed a multiple-choice test composed of 12 items, which related to the claims made in the blog-posts. Internal consistency reliability was $\alpha = .45$ which is lower than desirable and may have resulted from the fact that the assessment tapped multiple areas of knowledge regarding desalination and Israel's water sources. The topic knowledge score was based on the sum of the correct responses, $M = 4.75$, $SD = 2.09$.

Argument task: Following reading, the participants were asked to write arguments presenting their position regarding the desalination controversy. The writing instructions were: "Please write an argument that addresses the question: Should the State of Israel continue to encourage the construction of seawater desalination plants? Present your position on this issue and justify it". The arguments were coded for complexity and the number of sources that the argument was based on.

- Argument complexity - Coding was based on criteria employed by Schwarz et al. (2003):
 - No sound argument (0 points) - A claim is not provided or no justifications are provided for a claim.
 - One-sided argument (1 point) - A single claim is supported by a reason or a series of reasons.
 - Two-sided argument with partial justification (2 points) - A claim and a counter-claim or a two-sided claim is presented but reasons are provided for only one of the sides.
 - Two-sided argument with full justification (3 points) - A claim and a counter-claim or a two-sided claim is presented and reasons are provided for both sides. The first two authors independently coded 40 arguments. Interrater reliability was Cohen's kappa = .92.
- Sources in argument - The number of blog-posts referenced in the participants' arguments was counted. Interrater reliability was Cohen's kappa = .74.

Procedure

All questionnaires, blog-posts and tasks were displayed on a computer screen via an Internet browser. Data collection was conducted individually. We used the think-aloud method (Ericsson & Simon, 1993) to tap learners' sourcing activity. Participants were instructed and trained to say everything they think or do as they read. If participants were silent for more than 15 seconds they were reminded to think-aloud, using prompts such as "please tell me what you are thinking". No other instructions or feedback were provided.

Coding Scheme

The unit of analysis was defined as a comment or a set of comments that relate to specific source information (Strømsø et al., 2013). The coding scheme was developed by adapting and expanding a coding scheme by Strømsø et al. (2013). Two main additions were made to the Strømsø et al. scheme: (1) Because readers made references to themselves as a source of knowledge, the reader was added as one of the possible sources; (2) Source-source link activities were added to the coding scheme. Interrater reliability ranged from Cohen's kappa .87 to 1.00. The coding scheme is provided in Table 1.

Table 1: Sourcing coding scheme

Sourcing process	Category	Sub-category	Description
Source representation (SR)	Type of source	Present blog	Reference to the source of the current blog.
		Other blog	Reference to the source of one of the three other blogs.
		Other source	Reference to a source that is not part of the document set (e.g., a newspaper article previously read).
		Reader as source	Explicit and reflective reference to the reader as a source of knowledge (e.g., “ <i>I remember that</i> after the last winter, the water level in the Sea of Galilee actually rose”).
	Type of reference	Explicit	Clear and precise expression of source characteristics such as source name, profession, or affiliation (e.g., “Dr. Rabinovich, hydrologist”).
		Implicit	Statements that indicate sourcing without precise verbalization of source characteristics (e.g., “ <i>She says that...</i> ”).
	Source characteristics	Expertise	Explicit reference to source expertise, qualifications, professionalism, prior experience, and knowledge (e.g., “They are all <i>doctors...</i> ”).
		Position	Explicit reference to source’s stance regarding the desalination controversy (e.g., “It seems as if she is <i>for desalination</i> ”).
		Disciplinary perspective	Explicit reference to the disciplinary perspective or point of view of the source (e.g., “She has an <i>ecological</i> approach and not just an <i>economic</i> one”).
		Motivation	Explicit reference to financial, professional, or social motivations and interests (e.g., “He may have <i>financial interests...</i> ”).
		Currency	Explicit reference to the time in which the blog was written (e.g., “I wonder if these blogs are <i>up-to-date</i> ”).
		Other	Any other source characteristic (e.g., writing style or familiarity).
	Sourcing activity	Attention	Mention of the above source information without any further consideration of source reliability.
		Evaluating source reliability	Explicit evaluation of the trustworthiness of the source (e.g., “I am thinking if <i>he is right or not</i> ”).
Source-content (SC) links	Sourcing activity	Predicting	Use of source information to anticipate information to appear in the blog (e.g., “ <i>This is going to be different</i> because he is a consultant...”).
		Connecting	Relating source information to a specific knowledge claim made in the blog (e.g., “ <i>She writes that</i> the population is expected to grow”).
		Interpreting	Using source information to interpret a specific knowledge claim.
		Evaluating content reliability	Using source information to evaluate the reliability of the blog’s content (e.g., “I am not familiar with the data but I assume that <i>he is familiar</i> ”).
Source-source (SS) links	Sourcing activity	Comparing source claims	Comparing and contrasting specific knowledge claims made by the sources (e.g., “ <i>In the previous blog they said</i> that the water wells were becoming salty <i>but here they say</i> that there are other water wells”).
		Comparing source perspectives or positions	Comparing and contrasting source point of views, opinions, or positions (e.g., “She says that <i>desalination shouldn’t be done</i> and the previous one said that it <i>should</i> ”).
		Comparing source characteristics	Comparing and contrasting source characteristics such as expertise, affiliation, or currency, not including source positions or perspectives (e.g., “Before there was someone who was employed in the <i>Water Authority</i> and he is an <i>economist</i> ”).
		Comparing source reliability	Comparing and contrasting the trustworthiness of the sources (e.g., “It seems as if <i>everyone is exaggerating a bit</i> ”).

Results

Sourcing Profiles

The participants made a total of 219 sourcing comments that related to at least one document (i.e., present blog, other blog, or other source) while reading (3.6 comments per participant on average). However, examination of sourcing activities revealed substantially different patterns of sourcing. Using the Documents Model, four distinct sourcing profiles were identified:

Profile A - No sourcing – 12 participants (19.7%) made no sourcing comments at all.

Profile B - Minimal sourcing: Source representations only – 2 participants (3.3%) mentioned SR only without making SC or SS links. This profile was quite rare because SR were usually connected to blog content.

Profile C - Low sourcing: Source representations and source-content links – 20 participants (32.8%) mentioned SR and made SC links only. Participants in this group referred to the authors of the blogs, to their claims, and sometimes to their characteristics, but did not engage in explicit comparisons. For example:

It is not clear to me if when *she* speaks of alternative sources *she* is speaking about desalination or something else... OK, so this is actually pro desalination... [Goes on to read another blog] I agree with what is written here... I'm not sure this is sufficient, such a solution.... OK... What *he* says is interesting... [P40; emphasizes by the authors]

Profile D - High sourcing: Source representations, source-content links, and source-source links – Lastly, 27 participants (44.2%) mentioned SR and made both SC and SS links. For example:

It seems as if *she* has an interest in this because *she* is a consultant to the Water Authority... Compared to the previous one, they didn't mention the [pollution of the] water wells... It is clear that *she* supports desalination... [Goes on to read another blog] ... This doesn't make sense, in the first blog they mentioned different data and the numbers don't work out... Here they don't talk about desalination and they do talk about other solutions... [P7; emphasizes by the authors]

Participants in the high sourcing profile also made considerably more sourcing activities. See Figure 1.

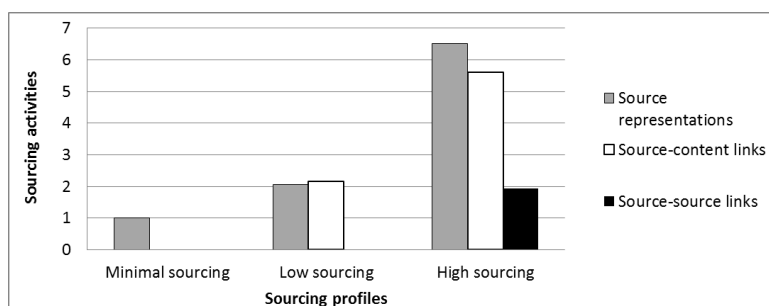


Figure 1. Sourcing profiles: Mean number of sourcing activities by sourcing process

Source Representations

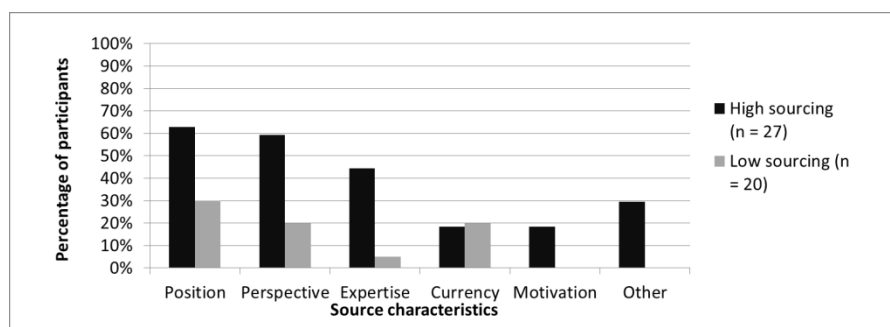


Figure 2. References to source characteristic by sourcing profile

The sourcing comments were overwhelmingly implicit (79.5% of the comments). However, high sourcing participants were more likely to make explicit source references compared to low sourcing participants, $\chi^2 (1, N$

= 47) = 15.09, $p < .001$. High sourcing participants were also significantly more likely to mention other blogs in their sourcing comments than low sourcing participants, $\chi^2 (1, N = 47) = 35.66, p < .001$. Lastly, high sourcing participants referred to a wider range of source characteristics. See Figure 2.

Sourcing Activities

Among SR activities, attention to the source (84.9% of the comments) was much more frequent than explicit evaluation of source reliability (15.1% of the comments). Figure 3 provides the percentage of participants who engaged in each sourcing activity. Low sourcing and high sourcing participants were equally likely to engage in simple attention activities. However, high sourcing participants were marginally more likely to reflect on source reliability, $\chi^2 (1, N = 47) = 3.64, p = .056$. The predominant SC activity was creating connections between sources and claims, i.e., noting who says what (75.8% of the comments). Use of source information to evaluate content reliability was a much less frequent activity (12.3% of the comments). With a single exception, participants did not predict or interpret the content using source information. High sourcing participants made more SC activities than the low sourcing participants but there were no significant differences between groups in the distribution of SC activities. The most frequent SS activity was comparing source claims (13.7% of the comments), followed by comparing source perspectives or positions (6.8%), comparing source characteristics (2.7%), and comparing source reliability (0.5%).

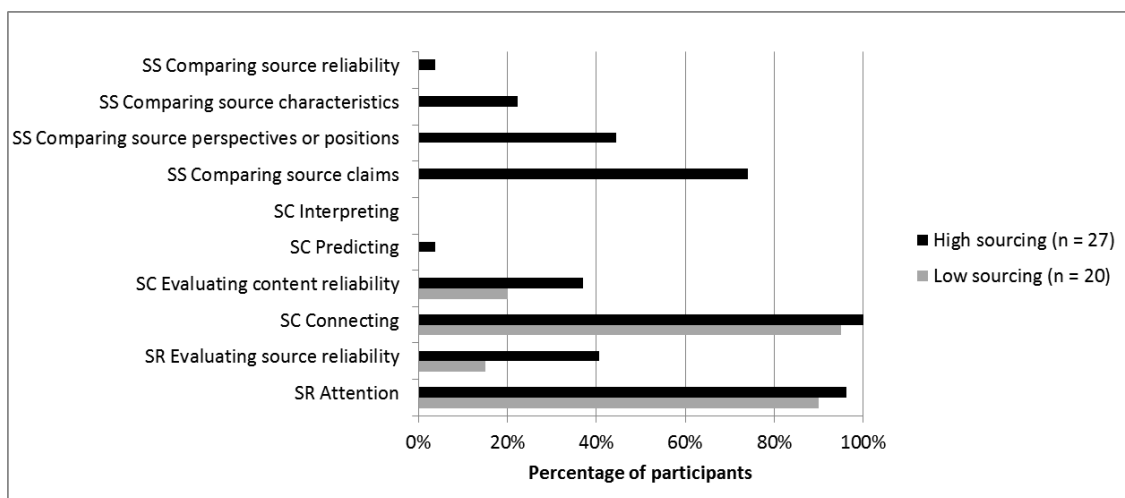


Figure 3. Sourcing activities by sourcing profile

Relation between Sourcing and Argumentation

In order to examine the relation between sourcing and argument construction we conducted ANOVAs with sourcing profile as a between-subjects variable (no sourcing, $n = 12$, low sourcing, $n = 20$, and high sourcing, $n = 27$) and argument complexity and argument sources as dependent variables. No effect of sourcing profile on the numbers of sources mentioned in the arguments was found. However, the effect of sourcing profile on argument complexity was statistically significant, $F(2,56) = 5.68, p = .006, \eta_p^2 = .17$. Post-hoc Tukey HSD tests indicated that high sourcing participants produced significantly more complex arguments, $M = 1.89, SD = .80$, than low sourcing participants, $M = 1.20, SD = .89, p = .013$, and no sourcing participants, $M = 1.17, SD = .58, p = .030$. No differences were found between no sourcing and low sourcing participants, $p = .993$. A possible explanation of the relation between sourcing and argumentation emerged from the participants' comments. Several high sourcing participants described in their arguments how source-source comparisons introduce epistemic doubt and promote an awareness of the complexity of the situation:

The articles presented contrary opinions that made me doubt the situation. ... I am afraid that if financial aspects will dominate the ecological aspects we will destroy the ecological balance and create irreversible damages and therefore I still have some reservations... [P15; emphasizes by the authors]

My opinion is conflicted. On the one hand, desalination has a positive aspect, and on the other hand, negative. My opinion is that we need to find the best solution, and when I say best I mean a solution that will cause minimal damage to nature while balancing the water sources. I found it hard to base my opinion on what was written [in the blogs] because many positions and opinions were presented. [P28; emphasizes by the authors]

Predictors of Sourcing

A multinomial logistic regression indicated that epistemic perspectives, topic interest, and topic knowledge were not significant predictors of sourcing profile (no sourcing, low sourcing, and high sourcing). We further examined the correlations between epistemic perspectives, topic interest, and topic knowledge, and specific sourcing activities among all participants who made document-related sourcing comments. Multiplism was found to be negatively correlated to SR activities, $r_s = -.31, p = .031$, and to SC link activities, $r_s = -.31, p = .028$. Accordingly, multiplism was also related to low attention to source characteristics, $r_s = -.38, p = .007$, and specifically to low attention to the positions and perspectives of the authors, $r_s = -.31, p = .031$, and to information source currency, $r_s = -.33, p = .020$. In contrast, multiplism was positively related to reference to the reader as a source of knowledge, $r_s = .36, p = .012$. Evaluativism was also positively correlated with references to the reader as source, $r_s = .29, p = .040$. However, evaluativism, in contrast to multiplism, was not significantly associated with document-related sourcing activities. Absolutism, topic interest, and topic knowledge were not correlated with sourcing activities.

Unexpectedly, gender was found to be related to sourcing profiles, $\chi^2(2, N = 59) = 6.80, p = .033$. Examination of sourcing profiles revealed that 29.7% of the women and only single man were included in the no sourcing profile, 35.1% of the women and 31.8% of the men were included in the low sourcing profile, and 35.1% of the women and 63.6% of the men were included in the high sourcing profile.

Conclusions and Discussion

A close examination of spontaneous sourcing while reading conflicting online information sources, revealed a wide range of sourcing practices. Some learners did not explicitly engage in sourcing as they read, whereas others actively engaged in the full span of Intertext Model building activities. High sourcing participants made more sourcing activities, paid more attention to source characteristics, and made source-source comparisons. However, even though high sourcing participants were aware of source characteristics and relations, they did not often employ source information in order to explicitly evaluate source and content reliability and to predict and interpret the content of the blog-posts. That is, even when learners did create Intertext Models, these were often loosely connected to an understanding of the text at hand and were not always used in order to construct an integrated understanding of the documents. Nonetheless, higher levels of sourcing while reading were found to be related to argument complexity, suggesting that sourcing while reading is an important practice when constructing knowledge from diverse information sources. Source-source comparisons may highlight differences between source positions and perspectives and thus promote knowledge construction. Additionally, such comparisons may introduce epistemic doubt (Bendixen & Rule, 2004) which can lead learners to consider and weigh alternate accounts.

A possible interpretation of our findings is that the authors' expertise might have conferred high reliability on the information sources and therefore may have negatively impacted learners' tendencies to evaluate source and content reliability. Further studies will be needed in order to examine how different types of contrasts between online information sources and different task conditions might affect sourcing practices.

The finding that multiplism is negatively related to attention to sources of the documents and positively related to attention to the reader as source of knowledge is in line with previous findings that views of knowledge as subjective and of justification as personal are related to lower attention to the sources of knowledge (Bråten, Ferguson, Strømsø, & Anmarkrud, in press). Evaluativism was also related to attention to the reader as a source of knowledge. However, evaluativist perspectives were not related to lower attention to the sources of the documents, suggesting that, in this case, awareness of the reader as source was balanced by attentiveness to additional sources of knowledge. Topic interest and topic knowledge did not emerge as significant precursors of sourcing, possibly because topic interest was sufficiently high in our sample and because the texts were self-explanatory and did not require extensive background knowledge.

The study revealed that not all learners spontaneously engage in sourcing. In our study, women were much more likely than men to adopt a position of silence (Belenky, Clinchy, Goldberger, & Tarule, 1986) vis-à-vis the authors/authority of the information sources they were reading. The finding that gender may have a role in sourcing was unexpected and requires further examination. However, this finding suggests that sourcing is a socio-cultural practice that is deeply related to how learners' view themselves as knowers in relation to expert authorities and to how they perceive their role in the knowledge society.

Design Implications

Our findings suggest that all learners might benefit from instruction aimed at fostering sourcing while learning with information sources. Promoting closer attention to source characteristics, consideration of their implications, and creation of source-source links appear to be particularly important aims for sourcing instruction. We propose that the design of such instruction needs to take into account not only the nature of online information sources but also learners' existing sourcing practices. These practices develop in the context of everyday use of the Internet for information seeking and reflect learners' views regarding what it means to

know and regarding their roles and goals as knowers in the knowledge society. Therefore, learning sourcing is not just acquiring a set of skills; it is also learning how to become a knower and a producer of knowledge in increasingly networked and information-rich communities. This reflection leads us to conjecture that sourcing instruction that addresses learners' epistemic perspectives regarding the nature of knowledge and knowing and regarding their roles as knowers in society might promote more active and critical sourcing.

References

- Barzilai, S., & Weinstock, M. (2014). *Development and validation of a scenario-based assessment of epistemic thinking*. Paper presented at the Annual Meeting of the American Educational Research Association (AERA), Philadelphia, Pennsylvania.
- Barzilai, S., & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources. *Cognition and Instruction, 30*(1), 39-85.
- Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist, 49*(1), 13-35.
- Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice and mind*. New York, NY: Basic Books.
- Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist, 39*(1), 69-80.
- Braasch, J. L. G., Rouet, J.-F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition, 40*(3), 450-465.
- Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46*(1), 48-70.
- Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (in press). Students working with multiple conflicting documents on a scientific issue: Relations between epistemic cognition while reading and sourcing and argumentation in essays. *British Journal of Educational Psychology*.
- Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction, 20*(4), 485-522.
- Britt, M. A., & Rouet, J.-F. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (pp. 276-314). New York, NY: Cambridge University Press.
- Ericsson, A. K., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (rev. Ed.). Cambridge, MA: MIT Press.
- Eshet-Alkalai, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia, 13*(1), 93-106.
- Gasser, U., Cortesi, S., Malik, M., & Lee, A. (2012). Youth and digital media: From credibility to information quality: Berkman Center for Internet & Society.
- Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehending and learning from internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly, 47*(4), 356-381.
- Goldman, S. R., & Scardamalia, M. (2013). Managing, understanding, applying, and creating knowledge in the information age: Next-generation challenges and opportunities. *Cognition and Instruction, 31*(2), 255-269.
- Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking and why does it matter? In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. (pp. 121-144). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mason, L., Gava, M., & Boldrin, A. (2008). On warm conceptual change: The interplay of text, epistemological beliefs, and topic interest. *Journal of Educational Psychology, 100*(2), 291-309.
- Rouet, J.-F. (2006). *The skills of document use: From text comprehension to web-based learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schwarz, B. B., Neuman, Y., Gil, J., & Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity. *Journal of the Learning Sciences, 12*(2), 219-256.
- Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction, 20*(3), 192-204.
- Strømsø, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing among students reading multiple documents. *Cognition and Instruction, 31*(2), 176-203.
- Tabak, I., & Weinstock, M. (2008). A sociocultural exploration of epistemological beliefs. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs: Epistemological studies across diverse cultures* (pp. 177-195). New York, NY: Springer.
- Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83*(1), 73-87.

Learning and Becoming in an After School Program: The Relationship as a Tool for Equity within the Practices of *Making and Tinkering*

Daniela K. DiGiacomo and Kris D. Gutiérrez, CU Boulder, 249 UCB, Boulder, CO 80309-0249
daniela.digiacomo@colorado.edu, kris.gutierrez@colorado.edu

Abstract: El Pueblo Mágico is a social design experiment (Gutierrez & Vossoughi, 2010) in which youth and adults are developing deep and meaningful relationships which facilitate learning that is inclusive, participatory, and robust. This paper focuses on ‘Making and Tinkering’ practices to examine the relationship that develops as both adults and children engage in the joint activity of making, re-making, and designing artifacts such as ‘Squishy Circuits’. Grounded in sociocultural theory, and situated within Nasir’s (2012) work, this study draws on a corpus of data that includes weekly observations and interviews over one year, to examine how the M & T activities are socially organized to increase room for feedback, one of the four important aspects of expansive learning contexts (Nasir, 2012). Specifically, I analyze interactions to understand how relationships facilitated by the social organization of the practice of tinkering can be tools in the design of inclusive, equitable learning spaces.

Introduction

Re-Designing for Equity: The Potential of Making and Tinkering

Making & Tinkering (M & T) practices have shown to be a powerful means for engaging and exciting children around Science, Technology, Engineering, and Mathematics (STEM) learning (Resnick and Rosenbaum, 2013; New York Hall of Science, 2010). Because of the playful, imaginative nature of many M & T activities, the traditional notion that ‘science is for scientists’ begins to dismantle, as children discover that they too can engage in scientific pursuits. M & T environments have recently been lauded not only for their ability to engage children in STEM learning, such as figuring out what materials conduct electricity or how to create a circuit, but also for their reimagining of what learning can look like. M & T practices can open up new spaces for students to develop a sustained engagement with learning processes (Washor and Mojkowski, 2010; Resnick and Rosenbaum, 2013). However, though the idea of ‘diverse learners’ is invoked in recent M & T publications, the research around what potential M & T practices can have on environments and relationships aimed at creating equity remains lacking. Toward this end, I explore what role M & T can have in one learning context designed for a racially and economically diverse student body and their predominantly White, middle class teacher-figures.

The relationship between teachers and students in classrooms has historically been one of teacher-as-knower and student-as-receiver (Rogoff, 1990). Complicating this hierarchical relationship is the fact that most teachers in the United States are White, middle class women, many of whom have had few, if any, meaningful relationships with students of color before they enter the classroom (Matias, 2013). Alongside these realities is often a ‘managerial-like’ approach to classroom organization, where teachers see their teaching identities as being heavily intertwined with their ability to control the classroom and transfer their own content knowledge to their students. Considering the grave inequities that continue to pervade the US educational system in what many term as the ‘opportunity gap’, there is a continued imperative to reconsider this traditional teacher-student relationship and the participation structures of learning environments, especially in light of the changing demographics and the increasing disparity between them (Carter and Welner, 2013).

To address the need for more robust and equitable learning environments, it seems fruitful to re-imagine how classrooms could be organized differently and what design principles could alter the traditional teacher-student power dynamic. The ways in which a learning space is organized is known to have a relationship with one’s ability to access and be a participant in that space (Rogoff, 1990). As such, a social practice designed to re-organize the space of learning in ways that encourage everyone to be a valued and active participant seems smart. M & T is an emergent framework for designing learning activity systems that attempts to do just that. Its design principles are governed by the idea that children should be the designers, rather than consumers, of the artifacts with which they interact. M & T practices encourage students to engage in an interest-driven collaborative process of (re)design, (re)production, reflection, and remixing. In the particular activity system for this study, undergraduate students and youth are brought together to engage in new media and tinkering practices that promote learning. The focus on children as designers, complimented by the immediate feedback nature of the activity of Squishy Circuits, is helping the undergrads to foster productive and

more symmetrical partnerships with the youth, often undercutting their initial ideas about how the relationship of adults and children in educational spaces is supposed to look.

Relationships as Tools for More Inclusive Practices of Teaching and Learning

While we know that relationships matter in overcoming normative stereotypes and discrimination (Nasir, 2012), I was curious as to if the ‘partner-like’ approach of M & T could have the ability to breakdown the power and authority which characterizes traditional teacher-student relations. I argue that the instantiation of M & T practices in educational environments has the potential to engender relationships and identities that are characterized less by the traditional teacher/student power dynamic to which we have become accustomed, and more akin to the valued partnerships known to be crucial to equity oriented social practice (Gutierrez and Vossoughi, 2010). Accordingly, I explore how the social organization of the M & T practice affords four tenets of learning (yet focus specifically on the first): consistent feedback, room for personal contribution, a sense of social belonging, and the availability of multiple roles; all of which combine to produce room for the development of meaningful, alternative relationships.

Theoretical and Methodological Approaches

Conceptualizing Learning and Relationship Processes through Analysis of the Social Organization of *Making and Tinkering Practices*

Recent interest in notions of connected learning (Ito, et al., 2013) have highlighted the importance of engaging youth in practices that are both interest-driven and expand youths’ repertoires of practice across settings and institutions, particularly for youth from non-dominant communities. M & T activities provide a context for connecting youths’ everyday interest and practices with new forms of activity and participation. Given the significant disconnect between youths’ out-of-school and in-school practices, the design of such activities becomes more salient. Within this framework, I analyze how the social organization of M & T affords new forms of engagement and participation between youth and adults working in joint activity.

In M& T, students take ownership over their own learning processes, and the design of the activity facilitates immediate feedback, open exploration, and fluid experimentation (Resnick and Rosenbaum, 2013). M & T activities encourage people to embody the ‘maker’ spirit and identity during learning-centered activities and to develop a sense of ownership and sustained participation. Furthermore, the “maker movement nurtures communities of practice that bring adults and young people together around common interests” (Washor and Mojkowski, 2010: 1), invoking Nasir’s (2012) notion that common interests aid in the development of meaningful joint activity and positive relationships between students and teachers.

Because M & T practices have been conceptualized as tools for fostering increased and sustained engagement among diverse types of youth (Washor and Majkowski, 2010), it seems fruitful to examine how its practices could engender and support greater equity in environments, such as schools, that are producing unequal outcomes. Because we know that teaching and learning are relational, and learning is a social practice that is organized in its context of development (Rogoff, 1990), it is important to investigate the relation between the organization of a learning environment and its effect on relationship building central to learning. To do this, I draw on insights and methods from Nasir (2012) as well as Gutierrez & Vossoughi (2010) to better understand how we can design for equity in learning environments.

Nasir (2012) analyzed how the social organization of the practice in alternative learning spaces such as dominoes and track and field afforded *four* key tenets that she identified as crucial to the creation of equity oriented learning environments: consistent feedback, availability of multiple roles, a sense of social belonging, and room for personal contribution. She found these four tenets of a practice to be especially salient for children from non-dominant backgrounds in spaces dominated by Whiteness, such as in schools. For the purposes of this paper, I emphasize and explore the tenet of ‘consistent feedback’, as it reflects, albeit slightly differently, the M & T notion of how ‘immediate feedback’ from process and product in activity facilitates increased levels of inquiry and engagement in practice.

Taken together, Nasir argued that when spaces are organized in ways that engender these elements, students can develop social and academic identities that build on rather than oppose each other. In looking at the social organization of an activity, she found feedback to be a critical aspect of the practice. Consistent feedback speaks to the availability of critical and supportive feedback from adults in moment-to-moment activity, guidance/assistance when needed, opportunities for observation and modeling, and multiple chances to try again.

The activity of Squishy Circuits, in particular, merits further exploration as the central activity of analysis in this discussion. “Squishy Circuits, developed by AnnMarie Thomas, consist of two kinds of play-doh; one is conductive, the other not. By layering conductive and non-conductive play-doh in different configurations, simple, tangible, ‘squishy’ circuits can be made and hooked into simple electronics” (LeDuc-

Mills, et al., 2013:618). There are multiple elements of this specific medium of play-doh that lend itself to playful exploration and playful learning without needing certain levels of prerequisite knowledge, both important elements of learning from a sociocultural perspective. Johnson and Thomas note that “these compounds have extremely low entry barriers; anyone can learn from, and enjoy them. The procedures for implementing basic circuits are very simple as well...one can almost immediately start building circuits...This learning tool was especially effective [for improving knowledge about circuits and electricity] among students that, judging from the preliminary test, had almost no pre-existing knowledge of these subjects” (2010:4103). These qualities of the Squish Circuit allow both youth and adults to “jump into the practice”, a central design principle of M & T at large and one especially helpful for facilitating co-learning by doing among intergenerational ensembles such as those at EPM.

Johnson and Thomas, in their discussion of Squishy Circuits and Science Education, also point to benefits of Squishy Circuits that are in line with Nasir’s social conditions for learning, namely ‘personal contribution’ and ‘social belonging’. “Playful learning and tangible mediums have been shown provide motivation to learn. Students are most involved in learning a topic when it intrigues their own *personal interests*. When students care about their work, they develop a profound understanding of their subject matter...Building circuits with the conductive and non-conductive dough, as well as various electronics components, gives students a *personal experience*, because they are designing their own implementations” (Johnson and Thomas, 2010: 4102-4103, *emphasis mine*). These localized learning experiences with the circuits should moreover be considered as tools to be taken up within the larger issue of lack of representation of diversity in STEM fields. “Research has shown a disconnect, between scientific direction presented in classrooms and students’ pursuit of science on their own. By late elementary school many students do not see their efforts outside of the classroom as “science” at all. Playful learning through tangible mediums [e.g., Squishy Circuits] bridges this gap by combining what students learn, and what they do for fun...Using squishy circuits has the potential to bring playful learning methodologies to electronics education” (Johnson and Thomas, 2010: 4102). Johnson and Thomas illustrate the potential power that this particular M & T activity has for engaging students in a practice (and ultimately field) to which they may otherwise have not readily entered, through designing for play and leveraging student interests.

So how do Squishy Circuits serve to help dismantle inequities between traditionally hierarchical relationships in learning activity? My argument is that the social organization of the practice of M & T facilitates room for the transformation of participation structures and the building of meaningful intergenerational relationships in educational settings. From this perspective, instead of receiving knowledge from adults, the children jump right into the practice of creating and re-making, and in doing so, there is more collaboration and less stratification of power among participants (Washor and Mojkowski, 2011; Resnick and Rosenbaum, 2013). Because the child may find a way to light up his circuit before the adult, traditional roles of expert and novice may shift. These elements of the practice produce an alternative type of relationship between the adult and child, one in which all participants become doers and learners, alongside each other. This alternative relationship, I argue, may have the potential to re-mediate traditional teacher-student relations in practice, as well as re-mediate some of the deficit ideologies that teachers from dominant communities may bring to their classrooms of color (Gutierrez, et al., 2009).

Methods

There are many ways to organize learning environments that invite increased participation and positive relationships among teachers and students. However, in my research in this particular historically indexed, equity-oriented social design experiment, I noted particularly interesting practices and processes of relationship building that emerged from the Making and Tinkering activities. I began this research with an emergent interest in relationships and was oriented to the following types of research questions:

1. What are the affordances of the social organization of the activities of *Making and Tinkering* at El Pueblo Magico?
2. In what ways is relationship development between youth and adults shaped by the social organization of the *Making and Tinkering* activities?

These questions are aligned with my guiding curiosities about what power this practice could have on designing for equity. In dismantling the deficit frameworks that novice teachers often bring into the classroom, there seemed to be the potential to facilitate a more partnership-like relationship between adult and child.

Because my research questions are focused on the social practices around learning and becoming within an activity system, I situate my central constructs of learning and becoming within a sociocultural theory of learning that views learning and becoming processes as socially constituted and relationally based. I understand learning and becoming in practice as illustrated by instances of shifts in participation, often signaled

by discourse that reflects a sense of value and identity in the practice. For instance, when the children said things like “Look, I got it to light up, I’m a scientist!” I understood them to be engaging in some form of learning and becoming in practice. In order to get at how relationships are engendered in the M & T activities and environment, my unit of analysis was the social organization of the practice.

The research context for my project was a designed based intervention—an afterschool program *El Pueblo Mágico (EPM)*, located in a predominantly Latino suburb of Denver, Colorado. EPM was conceptualized and directed by Professor Kris Gutiérrez, who designed EPM based on the principles of the Los Angeles 5th Dimension antecedent, *Las Redes*. *Las Redes*, like EPM, involves participants from both the elementary school and the University, to participate in a ‘change laboratory’ where transformative learning for both the student and the undergraduate is privileged (Gutiérrez and Vossoughi, 2010). On each day, there were about thirty children from low income and Latino backgrounds who attended EPM. The CU undergraduates, as part of their Educational Psychology requirement, worked within ensembles of two to six children. The semester’s target population of undergrads was primarily female and Caucasian.

I conducted a qualitative study drawing on principles of ethnography, to capture the social practices, specifically within the spaces of Making and Tinkering. I focused my eight field notes on the interactions and discourse among groups who demonstrated engagement over time with two primary M & T activities of ‘Squishy Circuits’, and in fewer cases, ‘Scribbling Machines’. Scribbling Machines activity involves creating a moving drawing machine out of batteries and art materials, employing similar guiding tenets for design as those in Squishy Circuits. These activities served as ideal practices in which to study the affordances of “alternative learning spaces” (Nasir, 2012) supporting relationship development that involve processes of learning and becoming through doing, albeit disguised more informally as collaborative art/game design and play.

In order to investigate the affordances of the social organization of the practice, I focused my note taking on capturing the four tenets of Nasir’s framework (social belonging, personal contribution, availability of roles, consistent feedback), as well as the discourse and social interactions among the participants and materials. To capture discourse around identity and relationship development *in situ*, I worked side by side in creating play dough within groups doing Squishy Circuits. After I analyzed and categorized the instances of affordances of social practices of M & T into the four categories, I began to understand the ways in which the M & T environment was lending itself to certain patterns of identity and relationship development. I added another layer of coding to my data based on the M & T design principles framework of Resnick and Rosenbaum (2013), such as ‘easy to connect,’ ‘open exploration,’ and ‘fluid experimentation,’ to examine their relation with Nasir’s affordances of alternative learning spaces. I looked for disconfirming evidence as well, of which there were a few instances when the M & T practices engendered the traditional ‘teacher as all knower’ pedagogical practices and relationships. However, to a greater degree, the learning affordances between how Nasir (2012) conceptualized the practice of dominoes mirrored the learning affordances in Resnick and Rosenbaum’s (2013) ‘Scratch’ Tinkering practice, overlapping in such a way as to confirm the reasoning behind deductively coding the second pass in this way. These analyses suggested that my constructs were appropriate and viable in addressing my topic of interest.

Major Findings, Conclusions, and Implications

Presentation of Findings

In my role as participant observer and interviewer, I documented the ways undergraduate participants discursively framed students. I noted a regularity in the use of phrases like “But this child is a problem, I don’t know how else to describe it” or “It’s hard for me to imagine what it must be like not to have support at home”, or “Things went well today, all the children listened and did what they were supposed to.” This seemingly deficit view of the children at *El Pueblo Mágico* was part of the everyday language I noted in this educational setting. As aspiring teachers, many of the undergrads initially talked about children, most of whom are predominantly low income children of color, in ways that positioned themselves as the sole source of knowledge, without an understanding of how children’s repertoires of practices were essential to learning (Gutiérrez & Rogoff, 2003). However, the organization of designed Squishy Circuits activities of which they were a part afforded new pedagogical and social relations in which undergraduates assumed ‘partner-like’ approach to teaching and mentorship.

The participation structures of the ecology of El Pueblo and M & T activity generated a range of ways for the undergraduates to mediate children’s learning and participation. Specifically, I documented three primary forms of feedback: 1) verbal feedback from the undergraduates in activity; 2) the feedback from the construction and design of the SC; and 3) feedback arising from the joint activity and collaboration with peers. For example, when the children’s attempt to make a circuit failed, the undergrads were able to provide feedback to the children in ways that were markedly different from their discourse outside the activity. Their feedback was supplemented by the consistent and immediate feedback nature of the actual Squishy Circuit. Ultimately, the forms of feedback helped to create an environment where the undergraduates could reframe their notions, in

ways that challenged pre-existing views of children from nondominant communities. As a result, new social relations between the undergraduates and the children developed into more collaborative and positive partnerships.

Drawing on Nasir's analytical framework of the four social conditions for learning, I coded the data deductively, seeking instances and examples of feedback, availability of multiple roles, social belonging, and personal contribution (Nasir, 2012). The results are shown below in Table 1. These analyses allowed me to see how the social organization of the practice of M & T was creating an environment in which the youth and adults, as co-learners, co-constructed relationships with the process, products, and participants that led to increased and sustained participation in the learning activities.

Table 1. Instances of the 4 social conditions for learning (Nasir, 2012) in m & t practices across 8 observations and 6 interviews

Feedback	Availability of Multiple Roles	Social Belonging	Personal Contribution
-22 instances of 'consistent feedback' - 13 instances of 'see the result' - 19 instances of 'see the process'	-35 instances	-33 instances	- 22 instances

However, as the table reveals, the emergence of consistent and immediate feedback (invoking both Nasir and Resnick and Rosenbaum's notion of the potential of feedback) seemed to be especially important to the particular practices of Squishy Circuits (and Scribbling Machines, despite their less prevalent use). To that end, I focused my analysis for this paper on the role that feedback played in helping to transform participation structures in this social design context.

I use the following representative example to illustrate the shift from discourse to practice. Undergraduate Miley described Ericka, a Latina 2nd grader as a really "sweet, obedient girl" (Interview 3, 3.20.13). When asked about what role race and socioeconomic status played in education, she said "they dictate the type of relationship you have with education...because like if your parents didn't go to college you may not be as inclined to go." At one level, Miley's words were reflective of a deficit framework commonly used to talk about low income non-dominant students. Her discourse also suggests that she may not have viewed Ericka as a likely candidate for college, on the basis of her parents' education. Interestingly, I had specifically sought out Miley for an interview towards the end of the semester because her relationship and practice with Ericka during M & T activities stood out as an exemplary of the affordances of the practice, and highlighted the relationship between activity and mediation. When working together to make Squishy Circuits, as we see below, Miley continually asked Ericka (the child) for guidance in creating the perfect conducting dough, as well as questioned her about how to say things in Spanish.

Field Note 5, 3.5.13: I positioned myself close to Miley, a blonde undergraduate, and her two Latino youth: Jennifer and Ericka, who were seated on the floor together making insulating play dough. Miley asked: "What do you want to be when you grow up?" to which Ericka says "a princess." Miley tells her that she had also wanted to be a princess. They continue to talk about the princess life, and Miley asks Ericka for help in adding the right amount of water to her dough as hers is too dry. The group next to them is talking about Spanish words for animals. I then hear Miley say, "What Spanish words can you teach me", and Ericka says "P-e-r-r-o"...

It was particularly noteworthy that Miley's discourse shifted significantly in the M & T practices. The collaborative activity of making a Squishy Circuit requires all participants to sit together on the floor, coordinating their activity in ways that positioned them physically at the same level with each other. These practices also invite participants to work in close proximity, distributing their expertise and shared ideas in ways that were more intimate and reflective of a partnership. Moreover, Squishy Circuits is an activity that appeared to foster conversation-in-process in which the participants were able to move from task-based conversation to personal conversation in a fluid and integrated way, allowing for increased potential for relationship building. Lastly, the use of play dough afforded different ways of designing the artifact, providing opportunities to continually repair both discourse and practices. I argue that these elements of the activity helped to create an

environment in which Miley could be expert and novice—a space in which she felt comfortable to ask questions and seek assistance from her youth.

Specifically, the consistent and immediate feedback provided throughout the development of the artifact allow for the learners to engage in “fluid experimentation” (Resnick and Rosenbaum, 2013). This fluid experimentation has been identified as important for youth learning and becoming in practice. Even in the case of a student who had struggled with behavior at EPM (Jose), we see how participation in the M & T activity engendered immediate feedback practices that helped transform one’s identity in the practice, and thus one’s identity as learner. The below excerpt is illustrative of how children’s ability to ‘succeed’, even briefly, is important for their sense of becoming in the practice.

Field Note 1, 1.29.13: Jose, a Latino boy who I had worked with last semester, stands close to Lisa as she explains Making and Tinkering activities. He seems uninterested and reticent, as evidenced by the fact that he says “I’d rather be playing video games” while Lisa was giving the two minute introduction speech. After about thirty minutes playing with the Squishy Circuits, however, he says “I’m a genius” when he was able to conduct electricity through a new manipulation of his pepperoni pizza play dough. Many others in his group seem excited and engaged based on their continual touching of the materials and talking with each other and the undergraduates, and their laughing, smiling, and saying things like “See, I’m a genius!”

This example demonstrates a shifting participation in practice (Rogoff, 1990). From a sociocultural theoretical perspective, one’s connection and identity in a learning environment are central to one’s ability to engage in expansive learning. Jose, I argue, shifts his understanding and engagement in the activity after recognizing his ability to be a valued participant in the practice; aided by the immediate, iterative, and localized feedback nature of the Squishy Circuit. Jose, similar to Jacqueline in the example below, seemed to display a sense of ownership in the activity, thereby increasing his/her proclivity to continue to engage in the learning process (Nasir, 2012).

From field note 8, 4.5.13: I sit close to Jacqueline, who is creating a butterfly out of the dough. She says that she hopes to light up its eyes. When I get up to see what she is doing and she can’t make it light up, she tries various ways to plug in the wires and the batteries. She then says “I’ll call you back when it’s ready.” Jacqueline’s undergraduate says to her: “How do you think you made it work the first time?” Paige and Devon, two other kids who are working to light their designs up next to Jacqueline, are also struggling initially either because of battery, LED, or wire issues. A few minutes pass of them talking about previous successful or failed attempts, and then Paige (youth) says “Devon it’s so cool-it’s so bright because I’m touching it!” Jacqueline then chimes in: “Wait I got it- It’s working! It’s working, and it’s so bright!”

Instead of asking for help, Jacqueline chose to continue to re-make her circuit through repeated instances of trial of error. This room for repair is a necessary social condition for learning (Nasir, 2012), and in this case not only does Jacqueline have the time and space to repair, but she seemed to value this feature of the activity, as she wanted to take charge of this process. When she is finally able to make it light up, she took pride in her success so much so that she called the attention of others to her learning accomplishment.

The opportunity for consistent feedback from peers and mentors, as well as the availability of multiple roles, increased participants’ connections to the learning activity. The interaction with Rosdy below highlights how both the actual activity as well as the environment in which it was embedded allowed Rosdy room to negotiate her understanding of the science reasoning behind the functioning of the ‘Scribbling Machine’.

Field Note 4, 2.27.13: When I asked Rosdy what she was thinking about when she was designing her scribbling machine, she said “I was thinking about what was going to work.” When she tries out her machine on paper, the propeller spins vigorously but the cup doesn’t move on the paper like it is supposed to. She and Jordan talk about why this could be the case, such as because the propeller is too heavy or the pen is bad. I signal to Andrew, a boy seated nearby at the computers with more experience in making Scribbling Machines, to come over and help. He says that we need to change pens before going back to his station. I offer to go get thicker pens and bring them back to Rosdy. The scribbling machine now scribbles around the page in a circular fashion. When I ask Rosdy if she knew that the machine would go in circles when she was making it, she said no. A few moments later, she says, “Maybe it’s because it is a circular cup.” Another pause... “Or because the propeller spins in a circle”.

Without being asked to do, Rosdy proffers a number of possible explanations to why her machine was working better or worse during our exchange. The social organization of the practice was such that she was able

not only have the ready assistance and ‘repair’ opportunities (Stone and Gutierrez, 2007) of her undergraduate, Jordan, but she was also able to talk through her thinking with Andrew and me, both more ‘experienced Tinkerers’. This immediate and localized feedback and room for repair is important for the development her own identity as a learner, in that she had the time and space to learn from others through modeling and to situate herself as competent within the practice. If the experiment had ended with Rosdy’s first attempt at making her Scribbling Machine draw, her understanding of motion and battery propelled energy (or simply her enthusiasm for exploration) may have been left wanting.

Because different designs produce different results in the activity of Squishy Circuits, it is difficult for there to be an expert in the practice. There are often unpredictable results, again dismantling the more traditional, recipe-like instruction so common in classrooms. The activity operated in a way that allowed for participants to play different roles at different times. This meant that undergraduates and youth became co-learners in this activity, as they both needed to try out different designs in order to create the brightest or biggest circuits. As the below excerpt from my interview with Miley demonstrates, she herself acknowledges how she did not know the science behind the functioning of the activity.

Interview with Miley, 3.20.13:

Me: What do you believe to be the affordances/constraints or benefits/obstacles within Tinkering?

Miley: Um, to get them to actually do the projects instead of just making the dough and playing with the dough, cause I think that the appeal of it is that it feels cool on their hands and stuff like that and they like making it because it's like cooking with their families. Ericka has started to like doing that lighting it up even though it was hard, but like when she figured it out and stuff, she liked it.

Me: And what would be like the benefits within that exercise, or the activity?

Miley: Um, if you can get them to give you responses to why the light bulb is lighting up, and understanding the insulating dough and everything like that. I mean remember you asked me yesterday why our circuit was lighting up a certain way, and I was like I don't even know the answer to that!

Miley’s responses in regard to the affordances of M & T reflected Nasir’s tenets of learning: “to get them to actually do the project” invoked Nasir’s ‘room for personal contribution’, “they like making it because it’s like cooking with their families” mirrored Nasir’s ‘sense of social belong’, and “like doing that lighting it up...when she figured it out” demonstrated its ability to provide consistent feedback. These aspects of the Squishy Circuit practice facilitated an environment which had the necessary social conditions for learning (Nasir, 2012). Moreover, the fact that Miley herself is actually unable to position herself as an expert, despite her status as an adult and mentor in the learning activity, highlights the potential power this M & T practice has in transforming participation and engagement structures in classroom-like settings.

Conclusions and Implications

Through Nasir’s conceptualization of how the social organization of alternative learning spaces helped to provide necessary social conditions for learning, I found the Squishy Circuits activity to similarly afford four crucial elements for inclusive, participatory learning: consistent feedback and room for repair, the availability of multiple roles, the valuing of personal contribution, and a sense of social belonging among the students. Students’ connection to the M & T practice was further facilitated by the expanded opportunities for learning and becoming in practice, and this affected their sustained engagement in the learning activities. Specifically, I argue that the availability of consistent and immediate feedback engendered space for the development of alternative and meaningful connections between the dominant undergraduates and their non-dominant students of color.

The undergraduate’s relations with students during M & T activities more closely resembled partnerships, rather than the typical power dynamics often seen in educational environments. In their discourse, the undergraduates demonstrated a managerial, one-sided approach to teaching. Yet during M & T activities, their practice looked very different: they often positioned themselves as novices, engaged in question-asking teaching practices, and easily enacted a mentor/partner-like role with the children. I argue that these different types of relationships were facilitated by the social made available by the instantiation of M & T.

These findings of new learning practices merit further investigation. I argue that alternative relationships are important not only for practices of good teaching and learning, but for designing for learning spaces that embody an equity-oriented lens and the abandonment of racialized identities and deficit pedagogies. Historically marginalized communities may benefit even more from the instantiation of M & T practices, as students from these communities are typically not given access equal amounts of access to higher order material, relational, and ideational resources (Nasir, 2012). The need for more opportunity to engage in meaningful practices should direct us to continue to focus on how to design robust and equitable spaces. The present study

suggests that M & T practices have the potential to foster alternative and positive relationships in educational spaces, with the opportunity to engender new types of participation structures where students and teachers can learn and become in practice, together.

References

- Brown, J., Collins, A. and Duguid, P. (2012). Situated cognition and the culture of learning. *Educational Researcher*, 18 (1), 32-42.
- Carter, P.L. and Welner, K. G. (Eds.) (2013). *Closing the opportunity gap: What America must do to give every child an even chance*. Oxford.: Oxford University Press.
- Cole, M. and Engestrom, Y. (1993). A cultural-historical approach to situated cognition. In G. Salomon (Ed.), *Distributed Cognitions: Ecological and Educational Considerations* (pp. 1-46). New York: Cambridge University Press.
- Cooper, C. (2011). *Bridging multiple worlds: Cultures, identities, and pathways to college*. New York: Oxford University Press,
- Engestrom, Y. (2004). From design experiments to formal interventions. University of Helsinki, Finland, 1-29.
- Gutierrez, K. and Vossoughi, S. (2010). Lifting off the ground to return anew: Mediated praxis, transformative learning, and social design experiments.” *Journal on Teacher Education*, 61 (102), 110-117.
- Gutierrez, K. Morales, P. and Martinez, D. (2009). Re-mediating literacy: Culture, difference, and learning for students from nondominant communities. *Review of Research in Education*, 33 (1), 212-245.
- Gutierrez, K. and Rogoff, Barbara (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher*, 32 (5), 19-25.
- Ito, M. et al. *Hanging Out, Messing Around, and Geeking Out*. Cambridge: The MIT Press, 2010.
- Johnson, S. and Thomas, A. Squishy circuits: A tangible medium for electronics education. CHI 2010, Work-in-progress. Atlanta, GA, 4099-4104.
- LeDuc-Mills, B., Dec, J., and Schimmel, J. (2013). Evaluating accessibility in fabrication tools for children. Proceedings of the 12th International Conference on Interaction Design and Children - IDC '13, New York.
- Matias, Cheryl. Keynote Speech at Teachers of Colors and Allies Summit, CU Boulder, October 22, 2013.
- Nasir, N. (2012). *Racialized identities: Race and achievement among African American youth*. Stanford: Stanford University Press,
- New York Hall of Science. (2010). Proceedings from the “Innovation, Education, and the Maker Movement” Workshop. Accessed through <http://www.nysci.org/media/file/MakerFaireReportFinal122310.pdf>
- Resnick, M. and Rosenbaum, E. (2013). Designing for tinkability. In Honey, M., & Kanter, D. (eds.), *Design, Make, Play: Growing the Next Generation of STEM Innovators*, pp. 163-181. Routledge.
- Rogoff, B. (1990). *Apprenticeship in Thinking: Cognitive Development in Social Context*. Oxford: Oxford University Press, 1990.
- Stone, L. and Gutierrez, K. (2007). Problem articulation and the processes of assistance: An activity theoretic view of mediation in game play. *International Journal of Educational Research*, 46 (1-2), 43-56.
- Washor, E. and Mojkowski, C. Making their way: Creating a new generation of “Thinkers”. Accessed through <http://www.handoc.com/documents/makersymposium.pdf>

“It’s Intentional”: Co-Construction of Transformational Processes and Pathways within and across Hubs of Interdependence in an Urban Community

Joanne Larson, Courtney Hanny, Joyce Duckles, Joel Gallegos Greenwich, Eric Meyer
University of Rochester, Box 270425 LeChase Hall, Rochester, New York 14627
jl Larson@warner.rochester.edu, channy@u.rochester.edu, joyce.duckles@warner.rochester.edu,
joel.greenwich@warner.rochester.edu, Eric.Meyer@warner.rochester.edu

Robert Moses, George Moses, Kimberly Jones, Jeremy Smith
Northeast Area Development, 360 Webster Avenue, Rochester, NY 14609
mosesrobert94@yahoo.com, gmoses@neadrochester.org, kbjn09@gmail.com, smith.jeremy950@gmail.com

Abstract: Beginning with data from a long-term participatory ethnography of an urban corner store, this paper discusses transformational processes and pathways intentionally being co-constructed across an urban community. Within the store and across multiple hubs in this neighborhood, we found six central processes that facilitated change: building relationships, building community, being family, communicating, belonging, and becoming. Additionally, we found that transformational pathways situated in the social, cultural, educational, economical, political, and research realms drove and connected these hubs, creating momentum towards sustainable change. We use the concept of a rhizome to explore how various hubs were connected through these transformational pathways and how redundancy across hubs created new developmental trajectories for community members. Looking across hubs, we found that transformational processes were generated when multiple and divergent discourses were intentionally allowed to remain in tension, rather than being compelled into consensus (creating what we call spaces of dissensus).

Introduction

The long-term participatory ethnography that informs this paper began with one community’s work to address the issues of “food deserts” (Pothukuchi, 2005) or “food swamps” (Rose, et al. 2009) in their neighborhood and the implications of food access for the health and well being of the community. In the absence of full service groceries or fresh produce markets, corner stores that sell preservative-laden, high priced foods, alcohol, and tobacco products often populate food desert neighborhoods. By selling these items, typical corner stores may exacerbate the social, economic, and health concerns within local communities (McClintock, 2011; Pothukuchi, 2005). Community activists in the Buttonwood neighborhood purchased a local corner store with the intention of turning it into a site of transformation toward healthier lifestyles in a healthier community rather than just a space of economic transaction. Early in the process of acquiring the store, a partnership was established between the community activists and researchers at a local university. The result was a community-university partnership that has spanned two and a half years so far. As a blended team of community based and university based researchers, we first assumed that we would be studying food and health practices. However, we soon discovered that we were studying the development of transformational pathways that fostered new ways of being through the ongoing co-construction of an organic curriculum with new kinds of relationships and new practices in the store.

Our findings indicate that relationship building is a foundational precursor for any change in practice. Additionally, five central processes emerged as well: building community, being family, communicating, belonging, and becoming. We identified these processes first in the corner store, but realized that within the various initiatives of this community-university team, there were other hubs in which these processes were intentionally and spontaneously co-constructed (See figure 1). We found that transformational pathways situated in the social, cultural, educational, economical, political, and research realms drove and connected these hubs, creating momentum for sustainable change.

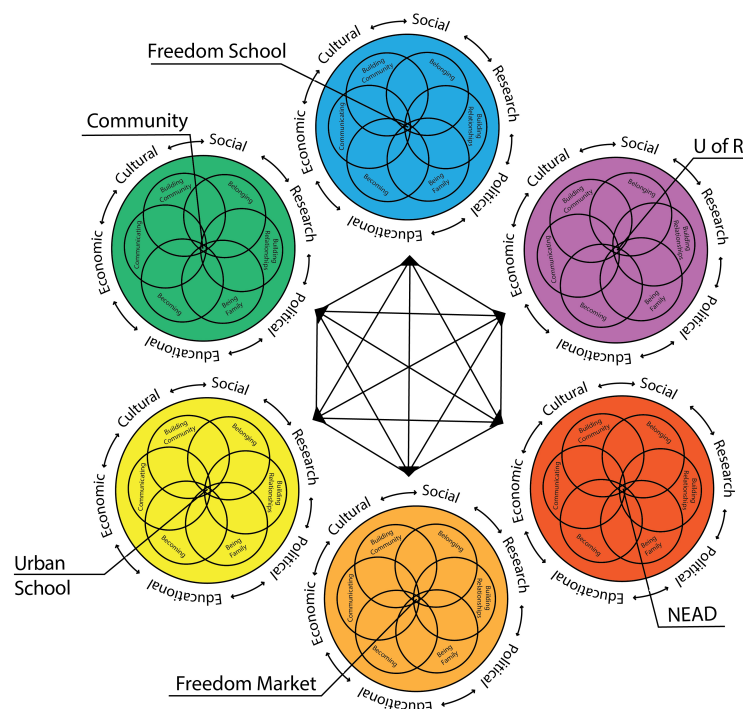


Figure 1. The interdependence model

For this paper, we focused on one hub, the Freedom Market (an economic space Soja, 2010), as it was situated within a complex web of interdependent hubs and practices (see figure 2). Accessibility to our six central processes across spaces, including those outside the store, proved essential. We adopted the metaphor of a rhizome (Deleuze & Guattari, 1987) to develop a model of interdependence through which we explore how various hubs were connected through these transformational pathways and how redundancy in processes across hubs created new developmental trajectories for community members. This interdependence model challenges the notion that learning and change occurs in a one-way trajectory or at a single site. Further, it challenges the idea that consensus-based models are adequate for understanding holistic and sustainable change practices. Looking across hubs, we found that transformation processes were generated when multiple and divergent discourses were intentionally allowed to remain in tension, in what we call spaces of dissensus (Ziarek, 2001), rather than being compelled into consensus, which is a direct challenge to the politically correct neatness that has caused well meaning stakeholders not to address the issues directly.

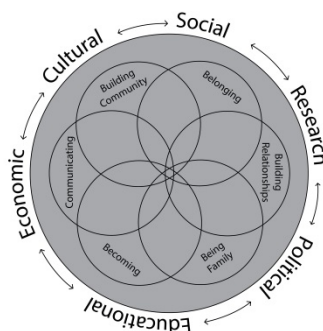


Figure 2. Hub and Transformational pathways at Freedom Market

Our findings provide insights into a new and dynamic way of conceptualizing the community as a multi-nodal site for change, where new ways of being manifest in unique but connected ways. Our goal in this paper is to recognize the value of these emergent processes and pathways and to make the intentionality of these change efforts explicit.

Theoretical Framework

In order to adequately account for the complexity and dynamism that became important in our data, we drew from postmodern conceptualizations of knowledge production and distribution, identity, and change, as intersubjective and networked practices. Specifically, we employed the concepts of dialogical learning (Bakhtin, 1981; Marková, 2003; Rommetveit, 1991), a rhizomatic model of interdependence and generativity (Deleuze & Guattari, 1987; Leander & Boldt, 2013; Leander & Rowe, 2006), and dissensus as an ethical political model (Ziarek, 2001). These perspectives allowed us to move between interactional, spatial, and sociocritical frames fluidly, recognizing that they were interdependent and mutually constitutive.

Methodology

Building on the extensive relationships that the community organizers had already fostered, we collaborated with community members to research their own neighborhood, which in turn informed the nature and structure of this project. We followed a participatory action research (PAR) (McIntyre, 2008) cycle of data collection-analysis-implementation-collection in iterative cycles to build a local evidence base documenting the effectiveness of this project, using terms defined by residents themselves. Residents were trained in research techniques and educated as co-researchers, co-authors and, along with university personnel, co-implementers so that they could better gather, analyze, and understand data on health and education issues in their community and participate in developing solutions to community-identified issues.

Our data corpus included two and a half years' worth of participant observation field notes, audiotaped and transcribed interviews with residents and customers, store personnel and families, audio/video taped and transcribed research team meetings, photographs, and surveys. We used a constructivist grounded theory approach for coding and analysis (Charmaz, 2006). As a research team, we each individually open coded data sets, then met as a team to develop categories. We worked together to saturate those categories and to build themes. From the themes, we developed the theoretical models of interdependent hubs and transformational pathways discussed below.

Findings

We found six central processes emerging within and across each hub: building relationships; building community; being family; communicating; belonging; and becoming. Our data suggest that, in each space, these processes were intentionally and spontaneously co-constructed by community members. Further, we found that these processes were becoming recognizable and evident through intentional (albeit informal) "teaching" and through organic processes of distributed cognition and dialogical knowledge construction (Hutchins, 1995; Marková, 2003). Through our findings, we show how the store was becoming a *transformational* space (manifesting new ways of knowing and being in relation) rather than a *transactional* space (transferring goods and currencies in a typical, impersonal retail exchange) and how, in that new space of transformation, learning moved across persons, spaces, and practices.

Building Relationships

The first and most foundational process we identified was *building relationships*. Our analysis led us to define building relationships as a process beginning with meeting people "where they are." It required sincerity, compassion, and a willingness to work *with* people as opposed to working *for* people. Relationships were found to develop through what people said and what they did – embodied and communicated directly or indirectly in everyday interactions at the store. Relationships involved a combination of building trust and expecting responsibility—a combination that generated tensions and challenges that were addressed collaboratively. This process proved salient for store personnel, customers, and researchers.

Store personnel at the Market were known collectively as the "Food Corps." The Food Corps was comprised of neighborhood residents hired to work inside the store as "nutrition interventionists." Although they had all the typical duties of store workers, such as stocking shelves and running the cash register, their primary duty was to engage customers and the community in making healthier choices of food consumption by suggesting healthy options for purchase, and even offering recipes and doing food demonstrations. These practices of situated and distributed learning exemplify teaching and learning woven into the everyday interactions of the community. This is not to suggest that there was not push back from store patrons who were accustomed to established ways of being in this space. Indeed, some community residents (mostly adult males) were resistant to these new interactions at what had been a space in which to "hang out" with (adult male) friends (field notes, 2/2013). Food Corps workers and the store leadership allowed for this resistance and engaged in respectful dialogue with residents who felt this way. Eventually, this play of difference instituted a new practice between "worthy adversaries" (Mouffe, 1993). Although differences of opinion remained, the differences became a source for almost ritualized joking and banter about a store that had become a research site (field notes, 2/2013-4/2013). Our analysis of data over time suggests that by allowing the space and time for

different ways of being in this space, Market workers allowed for a space of dissensus in which differences did not equate to disrespect—hence relationships could be built despite (in some ways on the basis of) opposing perspectives.

This openness to difference also manifested in forms of assistance to customers whose needs went beyond items in a store. The following interaction was recorded in one researcher's notes:

Ted is at the front of the store with Walter (a community elder who works at the store) and Noah (a store employee also from the neighborhood). They have moved away from the activity of the register are talking in quiet voices with heads down. I can overhear talk about "this time" and Ted seems to be agreeing. They talk over 10 minutes in this way. I find out after Ted leaves that he was trying to borrow money – enough to get him by for a bit. (Field notes, 3/2013)

Rommetveit (1991) notes that shared understanding is an interactional achievement that requires a strong degree of "attunement" to the awareness of the other. At the store, no one taught this; it was not available in a manual or how-to book. Rather, participants learned from each other in processes of relationship building. Building relationships involved trust, but trust was something that was earned. It was negotiated, starting with honest exchanges about everyday life. Instances of tacit "contracts" (Rommetveit, 1991) such as the one involving Ted, in which participants temporarily and partially shared perspectives to make meaning together appeared in these small ways where mutual respect was subtly (sometimes delicately) offered and tested. Furthermore, university folks had to be real with themselves and acknowledge many of the perceptions and narratives they come in the door with, which could have been construed as "hidden agendas."

Building Community

The second process we identified was *building community*. As the store manager noted, generations ago, barbershops and beauty salons were traditionally the "hubs" in which the broader community related and communicated—establishing informal communities of practice (Lave & Wenger, 1991). With heightened economic disparities in recent years, this sort of consistent, frequent interaction at such community establishments has decreased. Our analyses suggest that Freedom Market, with its model of transformational relationships, has filled this vacancy to become a new community hub. Numerous respondents in interviews and surveys indicated that the store was recognized as a central site for support, reaching across the community. One customer noted that the store is more like a "mom and pop's community store—most of the customers know the staff and they know everyone in here." This perspective reflects the old model of corner stores being owned and operated by folks who lived in the neighborhood, not seen as outsiders. Field notes indicated that it was common for people to stop in several times per day to ask for advice, or seek help with children or elders. One researcher's notes included the following:

A woman walks in with a folder full of forms – she's having trouble with them. She can't reach anyone for help and isn't sure who to go to. It seems to be some kind of credit assistance program. Walter asks if I can Google it while they go through the forms together. (Field notes, 2/2013).

The comfort and lack of second-guessing about accessing expertise within the store suggested a developing network of shared expertise, or an informal distributed cognitive system (Hutchins, 1995; Lave & Wegner, 1991). No one was assigned the role of "comforter," "internet information gatherer," etc. Rather, a mutually understood distribution of expertise evolved over time and through interactions across spaces. Similar instances in which expertise (far outside the "boundaries" of retail business expertise) frequent the corpus of field data, including the following:

- A teenager called from school when he was being sent home so that someone could notify his mom
- A six year-old came by to share his report card with the adults at the store
- A mother called the store to ask Walter to speak to her teenaged daughter, whom the mother fears is getting involved with a dangerous crowd (field notes, 3/2012-4/2013)

This sense of a new kind of community, emerging spontaneously but with intentionality, was observed to be evolving within the store as both workers and community members came to recognize and value the distributed expertise of Food Corps workers, management, customers, and even university researchers.

Being Family

The third process we identified was *being family*. A regular customer described the Market this way: “this is my family from another mother.” It was also called a “safe haven, my home away from home.” This notion of a new, homelike space was observed as children and adults visited the store up to ten times per day, taking items from the store’s coolers as one would a refrigerator in their own home. We began to uncover how the community came to see and co-construct the store as a “homeplace” (hooks, 1990). Roger, the store manager, noted “people have come to trust us with their lives.” He recounted how one day:

A woman who was about to faint due to her diabetes came into the store. She just sat on the stoop, and then proceeded to lie down. She told him she had been diagnosed two days earlier with diabetes. Roger notes, “having recently been diagnosed with diabetes myself, I knew to give her orange juice.” When she was able to stand, she expressed how thankful she was for the first aid provided.

Our data suggested that the Market became a safe haven, a place for nurturing in times of need, and a space in which family-like experiences such as doing homework and chatting about medical concerns were everyday occurrences. [H]ooks’s (1990) notion of “homeplace” is helpful for understanding this dynamic as not only comforting, but also powerfully and politically transformative. That is, “homeplace” complicates assumptions about marginality or outsideness by pointing out the powerful covert capacities for resistance within spaces from the “margins,” as margins are the very ideal spaces for “others” to “meet.” [H]ooks makes the case that these spheres of cultural practice and sites of resistance have always been there, operating under the radar without recognition or validation from a dominant vantage point. Seen in this light, the store as a site that fosters new ways of being family is a politically and socially potent site for change. This is contrary to the local trend that sees corner stores as places of loitering and illegal behavior, which caused the City of Rochester in 2013 to establish new more restrictive zoning for corner stores to address neighborhood complaints.

Communicating

The fourth process we identified was *communicating*. Communicating builds on the idea of the store as a hub, a resource and a gathering space. Data indicated that the store was seen as a place to get information and to share stories and cultural values, particularly regarding what was going on in the neighborhood. This was found to occur in at least two ways: spontaneously and intentionally.

Sometimes communicative practices at the store seemed primarily spontaneous and organic. For example, on one occasion, a customer stopped in and asked when a community member’s wake was scheduled. On another occasion, a customer asked about applications for local housing. Here, there were no constraints on what counted as a “valid” topic of conversation. Rather, dialogicality was emergent and situation-specific: conversations evolved organically, and they were often personal, sometimes “gossipy”, and very often highly political. One researcher noted the following:

When Walter came in I asked him, “Why were you carrying those newspaper articles in your pocket that day?” He said that he often carries articles with him. He said that, at first, he keeps articles for his own information. He shared that, eventually, he uses the articles to initiate conversations with community members and to challenge what is going on. He referred again to the fact that the CSD is the last in the country for Black and Latino males. “It makes you question what’s going on,” he added. Walter concluded that it is important for him to encourage his community to take a more active role in the education of its youth because “it takes a village.” (Field note, 3/2013)

Walter’s comment points to the other way that communications were seen to occur at the store—as the result of intentionality, a part of the new space/way of being that is being constructed. Another example of this was identified in the way that Food Corps workers distributed Family Nutrition Records to community members. The official purpose of these forms was to track the amount of fruit, vegetables, and sugary drinks families consumed on a weekly basis. The numerical data was being gathered for funders. At the same time, however, Food Corps workers used these forms to initiate dialogues about eating habits and food purchasing practices in the interest of fostering positive measurable results. Through everyday interactions such as these, Market staff would consistently “check in” with customers. As a result, they knew what was going on with families in the neighborhood.

Whether intentional or spontaneous, communicating in this space was dialogic and synergistic (Bakhtin, 1981; Community member, 2012). It was observable as a practice-in-development, filled with tensions, the easing of tension, and the generation of new tensions (Marková, 2003). Engaged participants worked through these tensions as respectful participants in authentic dialogue. To responsibly consider learning,

conversation, mediation, and change practices within the store and among the research team, we needed to recognize the varied histories of the individuals, the community, and the broader sociopolitical sphere that impacted store interactions. We found that concepts from radical democracy, which see difference and struggle as generative, productive, and necessary for ethical political engagement (Laclau & Mouffe, 1985; Mouffe, 1993), were far more effective for understanding the practices our research examined than were political models that foreground normative consistency or consensus building.

Belonging

The fifth process we identified was *belonging*. People who came to the store noted feeling acknowledged for who they were, in a non-judgmental way. A customer told a research team member, “Before, they didn’t do that.... But the store here, I feel you come in here and it doesn’t matter who you are, people are always kind and curious and respectful to you.” As noted above, even individuals with irreconcilable points of view seemed to find ways to be authentic—knowing their views were respected, even if not condoned—in this community space.

Some participants noted that churches and schools no longer seemed necessarily to offer sites for belonging, but that as a community hub, the Market was emerging as one place where they felt this sense of belonging. Further, many acknowledged that this sense of belonging was new to them. In one team meeting, a researcher identified this feeling in himself:

They came, they’re like they’ll just, I want to do something. I like how this makes me feel. I feel at home, some it’s like church...so they want to contribute. So they say how can I contribute to the way I feel. I feel good and I’m matching this feeling. I belong. I want to help. (Team meeting transcript, 9/2013)

The researcher’s reflection on his own learning pathway was telling. It suggested that he was not only observing, but also starting to feel like a part of the community and processes being observed. The boundaries between roles and places were in flux or tension, but the desire to belong and to be part of the change processes came through. This sense of nascent processes is a key factor in our sixth process: becoming.

Becoming

The last process we identified was *becoming*. Becoming implies the sense of change going on in the store and in the community (what we have identified as a shift from the transactional to the transformational). We found that this process was taking place at multiple levels—within individuals, among customers at the store, among the community members who work in and run the store, and among university research partners who found our roles evolving in unexpected ways. An employee of a local community development organization who started working in the store explained:

I found myself at the store more and more. I no longer wanted to just come to the store, I wanted to be part of the store. As I visited more, I found what was supposed to work for our community members was working on me as well. My role transitioned from being served to now serving. So how does one go from being the educator to being educated? No longer do I have an extended hand, but now a connected one. (Memo, 3/2013)

As community members themselves, the community activists who purchased the store and the Food Corps members placed high value on the human capital of this community, and held high expectations for the community’s capacity to come to know and understand what needed to be done for lasting, substantive change. For them, the store was a restorative space that could help the community “get itself back in order” (because perceptions from insiders and outsiders suggested that it had been out of order for too long). Although they drew from local traditions to enact these processes of change, store leaders made this space intentionally creative and new. They explored opportunities for fresh ways of keeping people engaged through creative expression and ways of combining marketability with ethics—being very “cutting edge” and yet very “old school” (field notes, 9/2012; 2/2013). Store staff were observed to consistently engage the community in creative ways to celebrate successes, and in ways to mediate problems through restorative justice practices so that consequences were fairly meted out for failures to be responsible to the newly developing, shared expectations of the emerging vision, or for contradicting the central values that were being solidified.

This double sentiment, of both old school practical wisdom and almost metaphysical attention to care of the other, might seem at first to be contradictory. Our analysis indicates, however, that such simultaneity of difference is a regular component of the agonistic, dialogical practices that took place here, and were part of what fueled its transformational capabilities and hence cleared new pathways.

Interdependent Nodes and Transformational Pathways

Within and across hubs, we identified transformational pathways. We found that the store was a social space; that it built on cultural practices of food as well as practices of restorative justice (“good food for good people”); that it created economies through opportunities for the dollars earned in the community to stay in the community; that politics and policies framed the everyday business of the store and many of the conversations, from school policies to drug laws; that it was an educational space where homework was done, forms were completed, and job training took place; and that it was a research site (as noted on the flyer posted on the door and on the side of the cooler) where community practices and processes of transformation were documented. We argue that nurturing and making these pathways explicit enervates the hubs by creating momentum toward change. One researcher observed at a meeting:

I just want to point out that it’s not just spontaneously transformational, right, it’s intentional. Because initially coming in they were standoffish, back, but now they come in with their problems, whether it’s cultural, social, economics...coming to understand the value of being within your neighborhood (team meeting transcript, 2013)

The transformational pathways we identified were not static, but rather dynamically interconnected to, and affected by, the different hubs within the whole rhizomatic network.

We came to our conceptual frameworks as a result of ongoing and collaborative attempts to make sense of our data. We often found ourselves struggling to find terms for the complexities and overlapping processes and ways of meaning making we were seeing. Dialogicality served to make sense of how meaning was constructed between and among persons, spaces, and times. We found that Deleuze and Guattari’s (1987) concept of the rhizome, based on the underground system of nodes, roots, and shoots (what one community members refer to as grass roots), helped to make sense of the ways in which the store, local schools, and other organizations (see figure 1) appeared to be distinct and separate, but actually had complex subterranean connections. The rhizome gave us a metaphor with which to explain how new ideas “took off” in new and unexpected directions, developing into new hubs or nodes that functioned alone, but remained connected to the store and the community. We propose that these “hubs of interdependence” stand in contrast to the ideologies of independence that drive many policies, social structures, and notions of family in the United States. For example, we found that economies in the store were connected with the community development organization across the street and the lives of community members. Likewise, the cultural framework of the Freedom School, with the foundation of the nine principles of *nguzo saba*, was present in the store through service, unity, and community. Further, community members and parents at the school have become co-researchers along with university faculty and students, who use community-defined evidence to inform practices, to present together at conferences, and to write grants.

Conclusion

We have presented the story of just one interdependent hub – the Freedom Market – to illustrate how new ways of being a customer, a community member, and a researcher evolved through the interactions at a corner store that was far from typical. Our research began at this store, but crossed into new spaces, as six common processes were nurtured across other hubs in the community. Our data demonstrate how, in each space, these processes were intentionally and spontaneously co-constructed by community members and university researchers. Furthermore, we suggest that movement across these hubs creates “transformational pathways” toward a larger activist network. We found that these “travels” constituted new developmental trajectories through which community members constructed new and changing ways of being, in community, and suggest that by making these practices explicit, other community researchers can develop transformational pathways relevant to the communities they serve. We must also balance quantitative data with qualitative data to give greater insight into the “why” at the same time balance research with implementation strategies that are informed by real time data.

References:

- Bakhtin, M. (1981). *The dialogic imagination* (M. Holquist, Ed.; M. Holquist & C. Emerson, Trans.). Austin, TX: University of Texas Press.
- Charmaz, K. (2006) *Constructing grounded theory: A practical guide through qualitative analysis*. London, UK: Sage
- Deleuze, G. and Guattari, F. (1987). *A thousand plateaus: Capitalism and schizophrenia*. Minneapolis, MN: University of Minnesota Press.
- Hooks, B. (1990). *Yearning*. New York, NY: South End.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: The MIT Press.

- Laclau, E. and Mouffe, C. (1985). *Hegemony as socialist strategy: Toward a radical democratic politics*. London, UK: Verso.
- Lave, J. and Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: University of Cambridge Press.
- Leander, K. and Boldt, G. (2013). Rereading “A pedagogy of multiliteracies”: Bodies, texts, and emergence. *Journal of Literacy Research* 45(1), 22-46.
- Leander, K. and Rowe, D. (2006). Mapping literacy spaces in motion: A rhizomatic analysis of a classroom literacy performance. *Reading Research Quarterly* 41(4), 428-460.
- Marková, I. (2003). *Dialogicality and social representations: The dynamics of mind*. Cambridge, UK: Cambridge University Press.
- McIntyre, A. (2008). *Participatory action research*. London, UK: Sage.
- Pothukuchi, K. (2005). Attracting supermarkets to inner-city neighborhoods: Economic development outside the box. *Economic Development Quarterly*, 19(3), 232-244.
- Rommetveit, R. (1991). On axiomatic features of a dialogic approach to language and mind. In I. Marková, and K. Foppa, (Eds) *The Dynamics of dialogue*, New York, NY: Springer-Verlag.
- Rose, D., Bodor, J.N., Swalm, C.M, Rice, J.C., Farley, T.A., Hutchinson, P.L. (2009). “Deserts in New Orleans” Illustrations of urban food access and implications for policy. National Poverty Center Working Paper. <http://www.npc.umich.edu/news/events/food-access/index.php>
- Soja, E. (2010). *Seeking spatial justice*. Minneapolis: University of Minnesota Press.
- Ziarek, E.P. (2001). *An ethics of dissensus: Postmodernity, feminism, and the politics of radical democracy*. Stanford, CA: Stanford University Press.

Middle School Learners' Ontological 'Trying-on' of Dimensions: A Phenomenological Investigation

Keri Duncan Valentine, Theodore J. Kopcha, The University of Georgia
221 Rivers Crossing, Athens, GA 30602
Email: frogstuff@gmail.com, tjkopcha@uga.edu

Abstract: This paper shares findings from a post-intentional phenomenological study aimed at understanding learners' experience investigating space and dimension concepts in a fifth and sixth grade mathematics class. Findings indicate experiences in this study manifest as an ontological 'trying-on' of geometric dimensions (e.g., through sight, perception, and motion), leading learners to conjecture about dimensional relationships. The paper discusses implications informing the second iteration of the design-based research project.

Introduction

Becoming is an ontological notion explicated in the phenomenological philosophical tradition of Heidegger and Merleau-Ponty. Heidegger questions the epistemological focus (i.e., consciousness of something) of phenomenological investigations, recognizing that experience is already situated in the world of being. Throughout his major work, *Being and Time* (1927/2008), he uses the phrase 'Being-in-the-world' to indicate the ontological phenomenological starting point. According to Heidegger, phenomenology's task is to question what it is to be in the everyday world. Merleau-Ponty (1945/2002) agrees with Heidegger's notion of 'Being-in-the-world' and makes embodiment a central theme in his work – a subject embedded in a certain position and time in space. This is exemplified in his work with the primacy of perception.

This paper shares middle school learners' experiences conjecturing about the relationships between dimensions following the film *Flatland* (Travis & Johnson, 2007). The film offers viewers an opportunity to contemplate analogously within and between dimensions. In addition to the film and subsequent discourse, learners constructed models to demonstrate 3D/4D relationships. For learners, the investigation was taken up ontologically as they 'tried on' the lenses, motion, and perceptions of imagined characters.

Issue Being Addressed

Rarely do learners reflect on the ontological experience of *becoming* in mathematics class. Yet, children form conceptions of space early in life simply by *being* in the world. They engage in everyday activities like "looking, walking, drawing, building, and manipulating objects" (Lehrer, Jenkins, & Osana, 1998, p. 169). They develop intuitions about spatial structure before formalized words are attached to their concepts (Freudenthal, 1983; Gravemeijer, 1998; Piaget & Inhelder, 1948/1967; van Hiele, 1986). It is possible that their pre-formalized concepts formed by *being* in the world could come in conflict with Euclidean geometric concepts formally taught in school.

These conflicts are shown to relate to the materials and processes used in schools (Lehrer et al., 1998). For example, most textbooks and teachers present images of triangles to students with a horizontal base, although this is not a defining characteristic. This results in students primarily recognizing prototypical examples of triangles to the exclusion of non-prototypical examples. These limited conceptions can inhibit students from reaching the intended goals of instruction and will most likely resurface in later geometry learning. Spiro, Coulson, Feltovich, and Anderson (1988) describe reciprocal misconception compounding as a problem resulting from simplifying complex concepts. Many misconceptions have been identified in geometry and their origins might lie in this simplification process as well as instruction that fail to incorporate learners' prior knowledge (e.g., Carroll, 1998; Monaghan, 2000).

The current educational climate rarely provides opportunities for learners to problematize space in geometry, rather it appears taken-for-granted that everyone shares the same concept – a Euclidean mathematical space represented by x, y, and z coordinates. Even more problematic is that school explorations of space start in an abstract manner, building on idealizations of the point (0 space), the line (1 space), and the plane (2 space) – spaces that are actually impossible to *find* in the world. In this study, we created a learning environment where learners could build on prior 3D experiences and the experiences of characters from the film *Flatland*.

Context

Forty-three 5th and 6th grade mathematics students were shown *Flatland* as a way to follow up on a recent unit investigating properties of polygons. After watching the movie, students questioned what it would be like to live in a four-dimensional world – to eat, to move, to *see*. There were no simplistic answers to their questions, so we conjectured as a class and eventually created three-dimensional shapes out of straws and pipe cleaners. These were dunked in bubbles, which allowed students to make physical models of hyper-shapes (i.e., representations

of 4D shapes). A year later, we conducted a post-intentional phenomenological study with students to explore their lived experiences. In particular, we explored if and how children take on the perspective of another dimension not their own (0D, 1D, 2D, 4D).

Theoretical Perspective

Our perspective acknowledges that humans' experience of phenomena in the world is complex. Phenomena emerge over time, are context dependent, and exhibit complex dynamics such as emergence, ambiguity, and adaptation (e.g., Davis & Sumara, 2006; Jacobson & Wilensky, 2006). Not only is the concept of space and dimension taken up as complex and ill-structured, but designing learning environments is viewed as a complex, iterative process informing and developing theory as well as practice (Barab & Squire, 2004; Gravemeijer & Cobb, 2006). At the early stages in design, our aim is to explicate phenomenological understandings concerning the ways shifts in perspective manifest in complex, emergent environments integrating innovative practices (e.g., modeling, mathematical discourse) and emerging forms of technology (e.g., animations, video cases). In particular, we used conceptual change strategies from science education including a discrepant event and bridging analogies as described by Clement (1993, 2008). These manifestations of shifts in perspective were then used to inform refinements in the design and theoretical conjectures.

There is little insight concerning middle school learners investigating the fourth dimension or even how this may play out using multimedia animations. Banchoff (1990) is one who passes along geometric insights to teachers of young children. He writes "[t]he invitation to examine coordinates from a dimensional standpoint is available at all times: We only have to make students aware of what they are seeing" (p. 33). We agree that the invitation is available but that it also encompasses more than sight. Freudenthal (1973), attributed with developing Realistic Mathematics Education, describes geometry as embodied:

since it is about the education of children, [geometry] is grasping that space in which the child lives, breathes and moves. The space the child must learn to know, explore, conquer, in order to live, breath and move better in it. (p. 403)

Clements (1998) describes children's investigation of shape through various actions (e.g., touching, drawing, discussing, moving); yet, one of his descriptions stands out – "development of perspective taking" (p. 3). He describes the "three mountains" task, where a doll is placed in several positions around a mountain scene. Children are asked to describe the scene from the perspective of the doll as it is moved, but always describe their own viewpoint instead. He writes, "it is not just familiarity or experience, but *connecting different* viewpoints, that develops perspective-taking ability" (p. 3). How does one connect different viewpoints? Taking on an external perspective is nuanced because children would also have to make a transition from two to three-space. It may help to think about the context of maps to understand viewpoint. According to Presson (1987 as cited in Clements, 1998) children:

must grow in their ability to treat the spatial relations as separate from their immediate environment. These secondary meanings require people to take the perspective of an abstract frame of reference ("as if you were there") that conflicts with the primary meaning. (p. 13)

Methodological Approach

Vagle's (2010) post-intentional phenomenological (PIP) research method was used in this study to capture the lived-experiences of middle school students investigating space and dimension. PIP brings post-structural thinking in conversation with phenomenology philosophers, such as Husserl (1901/1970) Heidegger (1927/2008), Merleau-Ponty (1945/2002), Gadamer (1975/1994), and methodologists such as van Manen (1990), and Dahlberg, Dahlberg, and Nyström (2008). Phenomena, and humans experience living with them, are seen as *tentative manifestations* – dynamic and continuously changing. Experiences are not only interpreted, but are lived in tentative, partial, and fleeting ways. Vagle (2010) has designed a five-component process for conducting PIP research: (1) Identify a phenomenon in its multiple, partial, and varied contexts, (2) devise a clear, yet flexible process for collecting data appropriate for the phenomenon under investigation, (3) make a bridling plan, (4) read and write your way through your data in a systematic, responsive manner, and (5) craft a text that captures tentative manifestations of the phenomenon in its multiple, partial, and varied contexts.

Phenomenology is the study of experience, or as Dahlberg et al. (2008) write, "the science of the world and its inhabitants, the "things of experience" understood as the world of experience" (p. 33). One cannot investigate phenomena without due attention to the central tenet of intentionality. Intentionality is best understood as an embodied relationship to things and beings in the world. Although similar to the word "intention," or action (i.e., I intend to go to bed early), intentionality is different. We choose to draw on Vagle's (2009) notion of intentionality as related to PIP research, thus situating intentionality within this study. Vagle writes (2010) "intentionality is shifting and forever partial and thus can be read through post-structural frames"

(p. 2). This presents a point of departure from other forms of phenomenology, such as transcendental phenomenology, which seeks to describe essences. Rather, we ascribe to Vagle's approach that identifies phenomena as *tentative manifestations*.

Phenomenological philosophers and methodologists have discussed in length and with great detail the phenomenological attitude, sometimes referred to as a scientific or open attitude. In choosing Vagle's approach, we chose a phenomenological attitude as one of openness, not of phenomenological reduction described by researchers such as Giorgi (1997), who uses transcendental notions of bracketing and Epoché to capture the essence of a phenomena. An "open attitude" as described by Dahlberg et al. (2008) means, "having the capacity to be surprised and sensitive to the unpredicted and unexpected" (p. 98). Bridling, described below, is one way both Dahlberg and Vagle suggest remaining open to the phenomena under investigation.

Bridling

Bridling is a way to negotiate complexity with data rather than being rigid and lockstep. The practice better encompasses the way towards an increased understanding of the phenomenon, and to us, the validity of the investigation. Bridling includes three main activities on the part of the researcher: (1) questioning pre-understandings (including assumptions), (2) remaining open, and (3) joining in an ongoing dialogue about the phenomenon. Therefore, bridling does not remove, set aside, or render the researcher non-influential but "animates and illuminates the researcher more fully" (Vagle, 2009, p. 592). The bridling entries in this study were made in three different formats: (1) as comments inserted alongside/in the margins of transcripts, (2) as individual, dated Word document files after analysis sessions, and (3) as notes inserted in books and articles.

Statement of Phenomena of Interest and Questions

The phenomena of interest in this study aims to show the reader "the lived quality and significance" of middle school students experience investigating dimensional relationships in a deep and meaningful way. The primary research question asks, "What is the lived-experience investigating dimensional relationships, including the fourth dimension?" Secondary questions ask: (1) What role does the video *Flatland* play in the experience for students? (2) Are there indications that persistent, or long-term learning occurred?

Data sources by Primary Research Question

All 43 students (21 fifth grade and 22 sixth grade at the time of instruction) from a K-12 independent school in the Northeastern United States were asked to participate in the study. Three fifth graders (1 female and 2 male) and seven sixth graders (6 female and 1 male) agreed to generate lived-experience descriptions one year after instruction. These were used to guide follow-up interviews where all 10 students articulated their experiences more fully and/or clarified their experiences. Second-round interviews were conducted as needed to clarify and expand upon key aspects of the experience. Although the 6th grade interviewees were mostly female, this was representative of the student body (18 females and 4 males).

Because we interviewed former students, we used a "conversational approach" to help alleviate possible power structures (Denzin, 1989). According to Denzin, interviewing "should *not* be a relationship where one party does all the talking and the other only asks questions. When interviews turn into this form, they become asymmetric, authoritarian social relations in which the power of social science determines the information given" (p. 43). To achieve this, we focused on having students describe their experience with as much detail and ability to capture multi-faceted aspects of the phenomenon.

During the instruction, student participation and work was captured through video and photographs. Additional data sources include lesson artifacts and bridling entries. These were used to guide interview questions, support/challenge interview data, and generate hunches as described by Glesne (2011). These additional data sources were also used to generate a multifaceted description of the phenomenon and answer secondary research questions.

Analysis: Whole-Part-Whole

We draw on both Vagle (2010) and Dahlberg et al.'s (2008) suggestions for data analysis, using a whole-part-whole approach. Following is a summary, describing with way data is analyzed both within and across sources:

- *Whole*. The first reading focuses on the whole data collection event, where all data pieces are brought together and read (not analyzed) as a whole.
- *Part*. This is followed by multiple line-by-line readings with researcher notes and follow-up interview questions for each participant. Subsequent line-by-line readings articulate meanings and consider notes, markings, follow-ups, and bridling entries. After saving documents for each participant, the last line-by-line reading articulates analytic thoughts for each part for each participant.
- *Whole*. Subsequent readings start to identify tentative manifestations across the data.

Findings

For this paper focused on becoming, we will explicate a particular manifestation of the learners' experiences *trying on* various dimensions (0D, 1D, 2D, 3D, and 4D). The phenomenon was evidenced through learners' conjectures about what it might be like to see, perceive, and move in the dimensions. Notable in the descriptions from learners was their attention to relationships within the dimensions. For example, when describing sight, perception, and motion in 3D space, they conjectured what it might be like to be a point, a line, a square, a cube, and even a hypercube in 3D space.

Learners described *Flatland* as the point when they started considering dimensional relationships. They related to the 2D Flatlanders' struggle to visualize three-dimensional objects. Students had an analogous struggle visualizing the fourth dimension. Students talked most about the last scene of *Flatland*. This is the point where they saw the fourth dimension animated for the first time and where they watched Spherius, the main 3D character in the movie, deny the existence of a fourth dimension. This denial prompted a lengthy discourse among students about close-mindedness. Students expressed being upset with Spherius for denying a higher dimension in a similar way that 2D characters in the movie denied a third dimension.

Taking on Sight and Perception from the Various Dimensions

Almost every student talked about what it might be like to *be* in a dimension. For some, this took the form of articulating what it actually means to be in 3D space and how they are only seeing parts of solids, but not all sides at once. They called this *seeing in 2D*. They talked about perceiving depth and *thinking* that we see in 3D. Somehow it was easy for them to then imagine by analogy what it must be like to see and perceive in the fourth dimension – they called this *seeing in 3D*.

Students engaged in their learning experience by trying on different dimensional lenses (or contemplating being other dimensions) and then visiting other dimensions, even the fourth. Alan is one student who attempted to generalize sight and perception for all dimensions and used Euclidean mathematical notions of the coordinate system in his discussion. He demonstrated an ability to perceive as “creatures” from different dimensions, related to the movie *Flatland*:

Pointland is a land of the zero dimension, where all that exists is a point. No axis. Lineland is only a line, the first dimension with the x-axis (or y). Flatland is the second dimension, with the x and y-axis. The creatures in the movie could only see a line wherever they looked. Spaceland is the third dimension, where we exist. With the x, y, and z-axis, it is most likely the largest ‘land.’ The creatures of this land can see a 2D image that can be perceived in 3D.

Interesting here was his distinction between sight and perception in 2D Flatland and 3D Spaceland. His ability to connect dimensions to the x, y, and z-axis demonstrated what Freudenthal termed *mathematization* (1973). He continued the process as he connected the fourth dimension into this same x, y, and z coordinate framework.

If someone asked me to describe the fourth dimension, I would say something along the lines of...basically a dimension more complex than ours as another axis or existence...starts with none in the zero, x in the first, y in the second, z in the third. We'll have to make a new letter in the alphabet for the fourth dimension. Or 'a'...it would probably work, yeah a different sort of direction. The fourth axis is probably where the idea of how I perceived hypercube came from. Cause it's a cube connected by another axis...like a cube in the center connected by the line coming from the outer corner. And I was thinking of those lines as being a fourth axis.

Alan was the only student who talked about the x, y, z, and “a” axis in relation to the dimensions, attributing *Flatland* as his starting point. He said previously he had only learned about the second dimension, but not the zero, first, third, or fourth dimensions. He talked about his experience as a whole:

Investigating the fourth dimension is to learn what the world is like mathematically and also to teach us more about the dimension we exist in. The fourth dimension was to explain why things happen, why the world is like it is. Well, before that, I didn't really understand. I had a vague understanding of the second dimension, but not much of the first, zero, third, or fourth. The biggest thing I learned is that humans exist in the third dimension but we can only actually see the second dimension – we just perceive the third. As we evolved to – not evolved – as we exist, only able to see the dimension below it and not the actual dimension it's in.

Dimensions Lower than Three (0-2)

This section incorporates students' discussion of the zero, first, and second dimensions. Students described these in various ways, taking the perspectives of the different types of dimensional creatures (0D – Point; 1D –

Linelander; 2D – Flatlander; 3D - Spacelander). For example, taking on a Linelander's perspective of its own dimension, or a Linelander's perspective of a Flatlander. In addition, they described the dimensions either by appearance, motion, perception, and even the sight available to a particular dimension.

Students talked least about the zero dimension, but almost all students recalled the “me” song that the Point character from *Flatland* sang. In the following description, Annie talked about the perspective of a point:

I had never really thought about it before, like a line can only see side to side of where it is or a point only notices itself because it can't - it's not really anything else, I thought that was really cool because I'd never really taken that in perspective before.

Students' discussion of the first dimension highlights their attention to limitations of sight and motion available to a line. For example, Susie expressed, “their [Linelanders] only way of traveling was right and left and only it couldn't have any sort of up and down. It was only just in the two directions.” Edward, Sabrina, and Daria also described Lineland, but they talked more about the Linelanders' point of view when other dimensional beings came into their line. For example, Edward talked about one Linelander's point of view towards Arthur Square, a Flatlander visiting the line in the movie:

In Lineland it could only see down it's line so it like it couldn't see up or down so it wouldn't be able to see him...it would only be able to see him if his eye or if his - if he was right on it's line, so I thought that was kind of cool how it had no other power to see anywhere else.

Students talked even more about the second dimension. Mary felt like the movie gave her a better understanding of the second dimension. “If you had a piece of paper, it helped me understand what it would be like if you had a circle that could, or a sphere that could sink into it.” Although this description may seem simplified upon first reading, Mary started articulating a foundational calculus concept of iteratively slicing a sphere. Another student, Susie, described the Flatlander's point of view and motion. Here, she started to insert herself into the second dimension showing how she took on the role of dimensional motion:

It's kind of crazy for me thinking about actually only living in the flat surface. I sort of remember always thinking about the idea of how do they move? Do they slide along on the floor or what? Or crawl? I remember when we started watching, that was basically one of the first things that popped out to me. I was like, how are they moving? Its kinda' cool to think about how – wasn't he saying that they only have northwards and southwards or something? That they don't have up and down? Yeah cause it's flat.

The second dimension is typically taught as a concrete dimensional plane. The movie allowed her to contemplate motion in this dimension, whereas her prior experiences limited her to only seeing two-dimensional drawings in books. It created another way of conceptualizing the second dimension, an unexpected (from the teacher's point of view) experience of taking on the role of seeing and moving in different dimensions.

The Third Dimension

Students' descriptions of the third dimension reflected a problematizing of their own 3D space. These students had little formal instruction concerning the third dimension at the start of the lesson. Edward gave a prototypical example of problematizing sight from his own third dimension, “We don't see circles, we see spheres. I mean we see circles but...and you can't really get anything completely flat except for like if you're looking at a screen. I guess that like what's inside would be totally flat.” In his description, there is tension as he tried to reconcile seeing “flat” but at the same time seeing, or perceiving, space. Although he did not use the word perceive, many students did. For example, Alan talked about seeing and perceiving. “The third dimension is something we can only perceive, and that we can't really understand further than that in a visual way.”

The Fourth Dimension

Students conjectured most about the fourth dimension. *Flatland* only gave a glimpse of a rotating tesseract at the end of the movie, but this seemed to be the scene that affected them most. During class, student questions about the existence of a fourth dimension were persistent and eventually led to the hands-on investigation with straw and bubble shapes. This led to several conjectures about the fourth dimension, elaborated below.

Sight in the Fourth Dimension

Mary talked about sight and being in the fourth dimension. It's as if she not only learned about the fourth dimension but also strengthened her understanding of the 0, 1st, 2nd, and 3rd dimensions and their relationship. She talked about what it would be like to see and perceive from each of these dimensions:

I want to know more about what it would actually be like to be in a fourth dimension. I have an idea. You'd be able to see all the sides of an object at once, but it would be like in Flatland. The square guy could only see the sphere when he was in his field of vision. So it would be weird to think that there would be things that we can't see...you can see all sides of something. I'm continuously getting the vision of your eyes popping out of your head and like curving around an object so you could see the back of it.

Mary's curiosity led her to think about seeing in a fourth dimension. Other students experienced this as well. For example, Edward said:

To see every side of a cube, your eyes would have to be...probably have to have one like, like coming out of like something that's like this [makes motion of hand making a hook coming from forehead curving out to front of face] I guess so you could see like this way toward...like you could see yourself...and you could see like anywhere I guess. That would be very difficult to see ourselves, like our eyes.

Alan also talked about sight in the fourth dimension, distinguishing between sight and perception:

Assuming that any creature can perceive in, can perceive a dimension below it – can SEE the dimension below it, but perceive its own. Well, that would imply the fourth dimensional creature would have to be able to see a third dimension. So basically what I imagine a fourth dimensional creature being like is a sphere with eyes on the inside of it, so it could look at something from every angle. And it would have to have some way of getting, getting things in there to look at, but I hadn't gotten past that. Yeah, you'd have to be able to see all sides at once. First my...could imagine seeing all three sides of a cube – all six sides of a cube at once.

Motion in the Fourth Dimension

In addition to conjecturing about the sight of four-dimensional beings, students also discussed motion. For example, Annie inserted the human form into the fourth dimension. "I think it would be pretty insane to live in the fourth dimension, especially if we were 4D humans, cause that'd mean like our guts would be like rotating inside of us." Edward used motion to think about what it must be like to eat a hamburger. "It will be forever going in and out of itself so you'd never get a bite with both [the bun and the meat] in it cause sometimes you'd get just the meat and sometimes you'd get just the bun." When Annie was asked how she would describe the fourth dimension to someone else, she said, "I would tell them that it's kind of like...I'm still not like 100% sure, but it's kinda' like an object...it...attached kind of inside an object, but it's kind of rotating so it's always on the inside or the outside." It's interesting that she was okay admitting not being 100% sure about describing the fourth dimension, yet continued trying. This seemed to happen for all the students interviewed.

Fourth Dimension Appearance

In addition to talking about sight and motion in the fourth dimension, students made conjectures about the appearance of fourth dimensional shapes. In these descriptions students used terms like "hyper-human." For example, Susie drew upon her science class to conjecture about "a human inside of a human" as a hyper-shape. Susie's remarks indicated that she was trying to picture humans as they might exist in these conjectured spaces.

An Open Mind

Students did more than describe the fourth dimension. They also talked about being open-minded and close-minded in relation to dimensions other than our own. For example, Mary said:

Before [learning about the 4th dimension] I always thought the 3rd dimension is as far as you can go in the dimensions or like, I don't know, not really the best, but we're the highest you can go or whatever. But now it's like...there might be something above us...not above us, but something different, or more, or whatever.

Mary talked about thinking that the third dimension was the highest. It's interesting that she chose the phrase, "not really the best, but we're the highest you can go or whatever." Edward did a similar thing:

He's [Spherius] like but there could never be a fourth dimension. So I think they can believe like numbers lower than them but they don't think there's anything higher than them because

they all want to be the best, I guess. Well, I thought that it would be hard to believe there was a fourth dimension.

This idea seems provocative and indicative of most human thinking, similar to people believing they are smarter than animals. Several students discussed Flatland characters' denial of higher dimensions, arguing about whether you could perceive other dimensions if you could never actually experience them. These discussions indicated a tension regarding whether students' perception of the world was closed or open.

Discussion and Conclusions

We chose these particular manifestations of the phenomenon to show how the experience led them to think about the various dimensions in provocative ways (motion, sight, perception, etc.). The findings indicate that *Flatland* served as what Zeck et al. (1998) refer to as a video anchor, a place to start conjecturing about dimensions. The fact that their conjectures persisted one year after their experience suggests that experiencing dimensional relationships ontologically may be a way to sustain student learning over time. Students were able to articulate specific scenes from the movie, as well as use them to continue conjecturing during the interviews. It may be that the video anchor, which provided a powerful analogy, was able to promote thinking and allow learners to *try on* a being outside their own dimension.

Although we asked non-phenomenological questions, such as conjecturing questions, we feel like their responses give researchers and teachers insight into the capability of middle school students' ability to reason deeply about space and dimension. We admit not attempting to formalize their learning by connecting the investigation to abstract symbols, although we did look at patterns between the dimensions (e.g., number of vertices, edges, faces). The main focus was encouraging them to engage in a conjecturing activity related to their interests following the film. Allowing the space for this to happen gives students an opportunity to engage in mathematical processes, much like a scientist problematizing their world. This seems in line with many of the Realistic Mathematics Education (RME) principles (e.g., Cobb, Zhao, & Visnovska, 2008; Gravemeijer, 1998; van den Heuvel-Panhuizen & Wijers, 2005).

Future Directions

We have already designed and are currently analyzing a second iteration of an instructional program that builds on this study. The revised program incorporates 70 cases of space and perspective in a hypermedia format and collaborative forms of reflective discourse to support the development of robust, flexible concepts among learners (Jacobson, 2008; Kolodner, 2006; Spiro et al., 1988). As part of our research, we have developed a framework grounded in the constructivist perspectives of Cognitive Flexibility Theory (e.g., Spiro et al., 1988) and RME.

At this point, we have developed a media rich set of cases for learners to examine multiple perspectives of space (e.g., noticing surprising similarities and surprising differences between variational instances of a concept) as a way to understand their inherent complexity. One investigation asks students to consider the various ways we capture and represent space. Cases include emergent video techniques that capture sporting events using slow motion or gyro cameras. Concerning invisible spaces, students are asked to consider fractal and hypercube animations representing dimension and spatial relationships we were unable to visualize until recently. At this still emergent stage of design, we are most interested in the ways students experience their engagement with the cases and the learning environment as a whole. Guided by the primary question, what is it for students to find themselves perceiving space as mediated in a variety of ways through technology, we are using post-intentional phenomenological methods to continue our investigation.

References

- Banchoff, T. F. (1990). Dimension. In L. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 11–59). Washington, D.C.: National Academy Press.
- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, 13(1), 1–14.
- Carroll, W. M. (1998). Geometric knowledge of middle school students in a reform-based mathematics curriculum. *School Science and Mathematics*, 98(4), 188–197.
- Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching*, 30(10), 1241–1257.
- Clement, J. (2008). The role of explanatory models in teaching for conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 417–452). New York, NY: Routledge.
- Clements, D. H. (1998). *Geometric and spatial thinking in young children*. Opinion paper, Arlington, VA. Retrieved from <http://eric.ed.gov/?id=ED436232>
- Cobb, P., Zhao, Q., & Visnovska, J. (2008). Learning from and adapting the theory of realistic mathematics education. *Éducation & Didactique*, 2(1), 105–124.

- Dahlberg, K., Dahlberg, H., & Nyström, M. (2008). *Reflective lifeworld research*. Lund, Sweden: Studentlitteratur AB.
- Davis, B., & Sumara, D. (2006). *Complexity and education: Inquiries into learning, teaching, and research*. Lawrence Erlbaum Associates, Inc.
- Denzin, N. K. (1989). *Interpretive interactionism*. Applied Social Research Methods Series (Vol. 16). Newbury Park, CA: Sage Publications, Inc.
- Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht, Holland: D. Reidel.
- Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht, Holland: D. Reidel.
- Giorgi, A. (1997). The theory, practice, and evaluation of the phenomenological method as a qualitative research procedure. *Journal of Phenomenological Psychology*, 28(2), 235–260.
- Glesne, C. (2011). *Becoming Qualitative Researchers: An Introduction* (4th ed.). Boston, MA: Pearson Education, Inc.
- Gravemeijer, K., & Cobb, P. (2006). Design research from a learning design perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 17–51). New York, NY: Routledge.
- Gravemeijer, K. P. (1998). From a different perspective: Building on students' informal knowledge. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space*, Studies in mathematical thinking and learning (pp. 45–66). Mahwah, NJ: Lawrence Erlbaum.
- Heidegger, M. (2008). *Being and time*. (J. Macquarrie & E. Robinson, Trans.). New York, NY: HarperCollins Publishers.
- Husserl, E. (1970). *Logical investigations*. (J. N. Findlay, Trans.) (Vols. 1-2, Vol. 2). London: Routledge & Kegan Paul.
- Jacobson, M. J. (2008). A design framework for educational hypermedia systems: Theory, research, and learning emerging scientific conceptual perspectives. *Educational Technology Research and Development*, 56(1), 5–28.
- 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.
- Kolodner, J. L. (2006). Case-based reasoning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 225–242). New York, NY: Cambridge University Press.
- Lehrer, R., Jenkins, M., & Osana, H. (1998). Longitudinal study of children's reasoning about space and geometry. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space*, Studies in mathematical thinking and learning (pp. 137–168). Mahwah, NJ: Lawrence Erlbaum.
- Merleau-Ponty, M. (2002). *Phenomenology of perception*. (C. Smith, Trans.). New York, NY: Routledge Classics.
- Monaghan, F. (2000). What difference does it make? Children's views of the differences between some quadrilaterals. *Educational Studies in Mathematics*, 42(2), 179–196.
- Piaget, J., & Inhelder, B. (1967). *The child's conception of space*. (F. J. Langdon & J. L. Lunzer, Trans.). New York, NY: Norton.
- Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). *Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains* (Technical Report No. 441). Champaign, IL: University of Illinois, Center for the Study of Reading.
- Travis, J., & Johnson, D. (2007). *Flatland*. Flat World Productions LLC.
- Vagle, M. D. (2009). Validity as intended: "bursting forth toward" bridling in phenomenological research. *International Journal of Qualitative Studies in Education*, 22(5), 585–605.
- Vagle, M. D. (2010, May). *A post-intentional phenomenological research approach*. Paper presented at the American Educational Research Association, Denver, CO.
- van den Heuvel-Panhuizen, M., & Wijers, M. (2005). Mathematics standards and curricula in the Netherlands. *ZDM - The International Journal on Mathematics Education*, 37(4), 287–307.
- van Hiele, P. M. (1986). *Structure and insight: A theory of mathematics education*. Orlando, FL: Academic Press.
- van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy*. Albany, NY: State University of New York Press.
- Zech, L., Vye, N. J., Bransford, J. D., Goldman, S. R., Barron, B. J., Schwartz, D. L., Kisst-Hackett, R., et al. (1998). An introduction to geometry through anchored instruction. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space*, Studies in mathematical thinking and learning (pp. 439–464). Mahwah, NJ: Lawrence Erlbaum.

Towards the Facilitation of an Online Community of Learners: Assessing the Quality of Interactions in Yammer

Marcela Borge, Pennsylvania State University, mborge@psu.edu
Sean Goggins, University of Missouri, goggins@missouri.edu

Abstract: This paper focuses on evaluating a theoretically informed approach to using social media as a means to support the development of a community of learners in an introductory human computer interaction course at a major US research university. Social network and communication analysis were used to examine the form and function of social interactions in Yammer. The data suggests that the theoretically informed approach to using Yammer succeeded in encouraging processes associated with a community of learners. The methods used in this paper may contribute to more effective ways of assessing the quality of discourse in socio-technical environments and the findings provide potential models for using social media as new contexts for developing conceptualization and discourse practices.

Introduction

User-Centered Design is a methodological approach within the field of Human Computer Interaction (HCI) that requires diverse knowledge and expertise. Both user-centered design and science share an inherent complexity, and the need for practitioners to make sense of many facts in order to advance solutions. Specific similarities include the requirement for a deep understanding of practice, higher order reasoning, and the ability to apply theory in concert with the accepted methods of the field. The inherent problem in teaching an introductory course in user-centered-design in higher education is that students expect a traditional lecture-based pedagogy, a form of instruction that is ineffective in helping beginning designers to develop necessary knowledge and abilities. In order to become designers, students need opportunities that require them to start thinking like designers: to critically evaluate artifacts in the real-world, think about user needs, cognition, emotion, resources, markets, and how all of these variables lead to the design of innovative solutions. Progressive instructional methods that emphasize student-centered learning have the potential to allow students to become more active participants in the learning process. The problem is that senior college students come with rigid mental models of what classroom interactions should be like in college settings. These student expectations have been shaped by past experiences and create obstacles for those aiming to introduce more progressive instructional methods, such as the development of a community of learners. If previous experiences create obstacles, then perhaps there would be fewer obstacles to progressive instructional methods in environments where students have had less educational experiences. Therefore, we wanted to examine the utility of using a social media environment, called Yammer to create learning environments outside of the classroom. We applied educational theory to guide the use of Yammer and examined the extent to which it could provide an engaging socio-technical context for learning. Yammer was used as a means to facilitate back channel discourse and communication in an introductory HCI Course at a major US research university. We then examine whether the form and function of discourse activity coincides with expected outcomes derived from learning theory.

Related Literature

Becoming Human-Centered Designers

Human-centered design (HCD) is a complex approach to developing design solutions that requires deep understanding of psychological theories and design practices. A central consideration of HCD is the identification of the needs, goals, and limitations of end-users. In order to become an effective user centered designer, students need to develop a wide range of content-specific and domain-general knowledge and skills. Students have to understand multiple theories related to how people think, learn, and interact with their environment as well as how these theories can be articulated through design (Carroll, 1997; Norman, 2002). Students also require domain general skills. They must learn to collaboratively illustrate their understanding of design problems and potential design paths, create representations to summarize findings, and communicate their ideas and reasoning to diverse audiences (Dym et al., 2005).

Unfortunately, students in HCD courses have been shown to lack domain-general skills: they do not reflect on their practice, negotiate or evaluate ideas with others, and also have problems applying course concepts to real world examples (Borge & Carroll, 2010). Students learn how to replicate methods and practice, but without a full understanding of why they are necessary or when they are useful. Such a findings are likely when students are not provided with opportunities to develop content specific and domain general knowledge in synchrony (Schunk, 2012).

Characteristics of Learning Communities and Implications

Students have been historically treated as peripheral in the learning process (Cuban, 1984). Most classrooms follow a model in which an instructor transmits knowledge to students and the main instructional goal is to motivate students to be receptive to the transmission (Rogoff, 1994). These types of classrooms are typically lecture-based, where the teacher is the primary source of verbal activity (Dunkin & Biddle, 1974).

Many researchers have advocated against traditional instructional practice in support of a more student-centered learning model. A student-centered learning model requires the instructor to shift from the role central knowledge authority to a facilitator of knowledge building activities, while students move from the periphery towards becoming central contributors of ideas for the community (Lave & Wenger, 1991; Papert, 1993). *Fostering a Community of Learners* (FCL) is one such model with certain key features: (1) students engage in individual research in order to (2) complete an important task and (3) share information in order to help each other complete the task. Meanwhile, students are also trying to deeply understand the disciplinary content that is inherent to the task (Brown & Campione, 1996). Another important aspect of FCL is for the instructor to make time to model thinking processes (Brown, Collins, & Deguid, 1989; Collins, Brown, & Houlm, 1991). For this reason metacognitive reflection is also an important element of this learning environment. Brown and Campione (1996) argue that it is not enough to have scripts or rules for how to behave, but that it is necessary to develop a system comprised of the previously mentioned components to provide a meaning and philosophy to support the culture of a community of learners.

The literature presented suggests potential methods that could be used to examine whether a learning community is actually developing in a virtual environment. For example, certain patterns of social interaction would need to be present. The instructor (expert) would likely be a central participator towards the beginning of a course with more posts and more interaction with students. In the beginning, students would likely be legitimate peripheral participators, participation in regular, low-risk behaviors that still aid the learning process (Lave & Wenger, 1991). As the course progresses, we would expect to see the expert slowly move to the periphery, while the students move towards central participation; the students would gradually become primary verbal contributors and participate in more sophisticated and risky social interactions. High quality and riskier interactions for design students in particular would include evidence of students reflecting on their learning practice or course concepts, evaluating course concepts, and applying course concepts to real-world examples in order to explore ideas in different contexts (Borge & Carroll, 2010). We can also infer from Brown & Campione (1996) and Brown et al. (1989) that it is crucial for the majority of a learning community's discussions to center on disciplinary content, where students use the language of the course, and try to develop their understanding by sharing resources, opinions, or interpretations meaning and implications in order to effectively apply course concepts to a task or practice.

Analysis of online participation data opens an opportunity to visually represent the patterns of interaction in an online community and examine the extent to which they coincide with theory. The ontology and methodological approach of group informatics (Goggins, 2013) will help us to systematically connect quantitative, visual indicators of movement from a network core to a network periphery with a qualitative understanding of how participants interact through Yammer. This can help to identify the primary ways that students discuss disciplinary content and to what extent this "talk" shows evidence of collaborative analysis and interpretation of course content.

Study design

Design-Based Research Methods and Research Questions

The overall methodology was design-based research, a method used to evaluate the potential of an educational technology that is grounded in theory or prior work and evaluated in real classroom settings (Brown, 1992). The main research questions were, (1) to what extent do patterns of interaction support or negate the development of a digital community of learners and (2) to what extent would students demonstrate sophisticated levels of thinking in the environment. The first question explores the extent to which the instructional approach succeeded in pushing students to communicate with each other about course content and become the primary providers of information for the community. Whereas, the last question focuses on ensuring that the discourse is productive from a learning perspective, as content related interpretation is more conducive to learning than off-task discourse (Wienberger & Fischer, 2006).

Course Content, Learning Objectives, and Student Assessment

The introductory course to user-centered design is a requirement for students in the design and development track of an information sciences and technology college in a large, US University. The course introduces students to fundamental concepts and practices of interaction design. Interaction design bears a resemblance to HCI, but goes beyond HCI's traditional emphasis on interaction with computers to include designing for a variety of human experiences (Rogers, Preece, & Sharp, 2007). Nonetheless, interaction design and HCI share

many core concepts and techniques. The course builds on a foundation of knowledge from cognitive and developmental psychology in order to help students understand “users”, the people they design products for. The course presents users as thinking beings with specific cognitive limitations that learn to interact with objects by synthesizing present experiences with past experiences and the knowledge, tools, and values associated with those experiences. These psychological theories serve as the foundations for design heuristics, additional theories, and methodological frameworks that evaluate the products that designers and developers create in their attempts to solve everyday problems.

The introductory 16-week user-centered design course was divided into two parts: Part 1 (knowledge comprehension) and Part 2 (knowledge application). It was taught by one instructor and supported by an undergraduate learning assistant. During each week of Part 1 of the course, students worked on a design challenge connected to course concepts. During Part 2, students picked their own design challenge and were expected to complete a project to demonstrate their ability to apply the core concepts and techniques covered in the course. The class met three times per week for 50-minute sessions. On the first day of the week the instructor would take 20 minutes to go over difficult concepts, on the last day of the week the instructor would use 15 minutes to review the week and answer questions. During the rest of class time, students worked on design challenges with a team of five-to-six members. During these work sessions the instructor and the learning assistant would walk around the class and check in on teams. In place of additional homework, students were required to post on a professional social media site, called Yammer, outside of class each week and discuss course topics. Students would receive credit for starting new discussion threads by creating an original post. They would also receive credit for engaging in an existing discussion by replying to an original post or by replying to a reply. Students could also “like” posts or replies, but would not get full credit for this kind of activity. This was the primary form of discussion for students with other class members that were not in their immediate team. Course participation accounted for 25% of the students’ overall grade and consisted of attendance, posting to and moderating Yammer, and providing one peer review of another team’s design challenge.

Participants

The study participants were junior and senior college students enrolled in an introductory human-centered design course at a large, US university. Students were divided into nine teams of five-to-six students in part one of the course and a different team in part two of the course. There were 38 participants included in the study, four of which were women; this is fairly representative of the gender distribution in the college.

Introducing Social Media as a Learning Tool

When integrating Yammer into classroom, there were design principles derived from the learning theory that the instructor followed. These were used as a means to increase the likelihood that Yammer would serve as an effective learning tool. For example, a modified version of cognitive apprenticeship was used as a means to acculturate students to the online community and present it as a situated learning environment (Brown et al., 1989; Collins et al., 1991). The Instructor would model desired posting behaviors, coach students on their posts, encourage sense-making activity, provide guides to help students learn how to moderate Yammer, slowly shift moderating responsibility to student teams, and then slowly fade from the environment. The goal was for the students to eventually take ownership of the environment.

Students were introduced to Yammer in the first week of class and the instructor modeled a variety of different original posts: reflecting, sense making, resource sharing, coaching, and polling students. Students were told that posts should relate to weekly course content. Throughout the next four weeks, the instructor used Yammer to model how to use and evaluate a range of outside resources, (i.e., academic articles, professional UX design blogs, magazine articles, and Wikipedia) and thinking (how sources could be used as a means to further develop understanding). The instructor maintained full responsibility for moderating Yammer for four weeks and then assigned each team to moderate thereafter. When moderating, students were responsible for encouraging each other to connect course content, respond to posts, and keep track of participation. The moderation of Yammer was a large part of their participation grade.

Data Collection

The data from the Yammer environment was extracted and exported to excel spreadsheets. There are three types of posts in Yammer: original posts, replies, and replies to replies. Each *original post* was the start of a new conversation and was time stamped. The data also provided the name of each poster, a unique ID, and created a Thread ID matching the unique ID of the poster. When students *replied* to original posts, the replier would be identified by a new unique ID and associated with a Thread ID that matched the original poster’s ID. This allowed us to see who replied to whom and identify different threads of conversation. The data also allowed us to see *replies to replies*. These posts contained the users unique ID, as well as a Replied to ID that matched the unique ID of the person they replied to; the thread ID still match the unique ID if the original poster. The term

“post” refers to original posts, replies, and replies to replies. Whereas, the top-level post that starts a new conversation thread refers to an original post. The data also included the entire content of what was being shared and links to any additional articles or materials that students attached to the post. This allowed us to analyze the quality of content.

Examining Social Interaction Patterns

Group informatics methods were used to examine patterns of interaction that emerged over time (Goggins, Mascaro, & Valetto, 2013). The data was cleaned and exported to a social networking analysis tool, called Gephi. In total, 503 posts were analyzed in Gephi. This tool facilitated the creation of visualizations to capture activity: people are represented as nodes and lines between nodes represent connections. Connections indicate events where students respond or are responded to in the Yammer. Two separate students exported the data into Gephi and produced independent visualizations and ran analysis in order to ensure that the system and subsequent analysis was reliable. Besides aesthetic differences, the results of the analysis were the same.

Classification and Assessment of Posts

The objective of this analysis was to assess the quality of discourse that students engage in while using the environment, for this reason the 98 instructor posts were excluded from the analysis. The remaining 405 student posts were classified according to the topic of the contribution and whether it was connected to course content. Posting behavior was categorized according to the type of cognitive activity represented in each post. In order to develop a valid and reliable coding construct, coding activity was initially based on differing levels of cognitive activity that coincided with previous work (Borge et al., 2012). When possible behaviors were exhausted, a construct map was developed with levels of cognitive behaviors, definitions, and concrete examples. The different types of activity were then compared to research from communication analysis and sense-making literature (Convertino et al., 2009; Dyke, Howley, Adamson, Rosé, 2012; Pirolli & Card, 2005). It is important to note that the levels of did not assess writing quality, but rather the level of cognitive activity that students displayed in the post. Utilizing methods similar to interaction analysis (Jordan & Henderson, 1995), meetings were held with student researchers to pick out behaviors and distinguish between codes in order to refine the coding construct. Level of cognitive activity was used as a means to assess the quality posts in the social media environment: informal, practice oriented conversations. Once the coding construct was finalized, the first author and a research assistant analyzed 20% of the data. The inter-rater reliability was substantial (Landis & Koch, 1977): $r = .89$, $p < .001$; Kappa = .78, $p < .001$. Disagreements were discussed and resolved. The research assistant then coded the full data set.

Results

Social Interaction Patterns

Social network analysis revealed three major patterns in the data. These patterns support the claim that there was a high-level of student participation and connections between students during discourse, and that students eventually took responsibility over the learning environment. Our findings use the language of social network analysis and group informatics. Conceptually, lower centrality measures correspond with what Lave & Wenger (1991) describe as “peripheral” participation. High “degree centrality” corresponds with membership in the core. The first pattern is related to the density of participation and level of connectedness of students. Findings indicate that there was a high degree of participation from the class as a whole and students were very connected to classmates through posting behavior. This is characterized by the visualizations in Figure 1, which shows a high degree of participation and ties between students at each four-week time interval. Nodes indicate students who contributed posts and the size and color of the nodes indicate frequency of posting behavior. Lines between nodes indicate connections between people, meaning they received a post from or sent a post to another contributor. Of the 38 students enrolled in the class, only 2 did not participate in the environment at all. Of those that participated, four did not establish any connections to other students. On average, students made a total of 10.28 posts to the environment over the 12-week period, $SD = 7.18$, $Min = 0$, $Max = 29$. Total posts include original posts and replies. The total number of student posts was 405, compared to 98 for the instructor.

The second pattern identified is related to ownership and cognitive presence in the environment (Garrison, 2003). The instructor moves from a central to peripheral participator over time, while students become more central contributors. Figure 1 shows this pattern over three time intervals. Degree centrality, or the impact a member has in the discussion, is shown by the relative position in the network: the more impactful a person is relates to how centrally they are located in the network. At P1 the instructor (with the alias of II) is in the central left with a relatively high degree of posts and connections. At P2, “II” shows movement towards the periphery of the graph, with fewer connections and posts than at P1. At this time, several students start to become more central than the instructor. By P3, “II” has moved to the periphery with very few posts and far fewer connections, whereas students’ posting behavior becomes more equitable and central to the network. We

also find a shift in type of participation from P1 to P3, adding analysis of the direction of degree centrality and weight of connection, shown in figure 2. “In degree” centrality (x-axis) is expressed as replies to a person; “out degree” centrality (y-axis) as replies from a person. Figure 2 illustrates increasing diversity of participation across the time periods, with more people favoring “in degree” or “out degree” centrality. To calculate weight of a connection, we assign less weight to communications engaging a larger group and greater weight to direct replies to another student, following benchmarks established by Goggins, Laffey, & Gallegher (2011), and a systematic methodological approach for analysis of trace data, called Group Informatics (Goggins et al., 2013).

Types of Posts

Posts were initially classified as containing course content talk or “other” talk. Course content talk is defined as posts or replies related to core concepts and techniques of the class, as well as project related discussions. “Other” talk included course management and metacognitive behaviors related to awareness and reflection of students’ own learning needs, practices, or experiences. Of the 98 Instructor posts, 58 (59.2%) were related to course management, a type of “other” talk. In contrast, only 55 of the 405 student posts (13.6 %) were “other” talk and 86.15 were directly tied to course content. Some of this “other” talk was included posts where students shared resources that challenged traditional learning models and would discuss their own learning experiences. Only 0.25% of total talk was considered off-task, not dealing with content or learning.

Content-related posts were examined for level of cognitive behavior. There were five levels of exhibited cognitive behaviors for content related talk: sharing, extending, checking/ rephrasing, synthesizing and interpreting (see table 1 for a list of exhibited behaviors, definitions, and frequency counts). Though extending posts were not ranked as highly as interpretation posts, these were used a great deal by students. Students would regularly go back to search for content materials or resources other students shared and would refer to them in class or use them for their projects. In fact, the most common student complaint of Yammer was that it was hard to find previous posts.

Findings indicate that the quality of posts were relatively high. Sixty-six percent of the posts were classified as synthesis or interpretation of course content. However, there was a range of quality within each level. For example, a post would be classified as interpretation if students made a content related claim and (1) supported this claim with evidence or rationale or (2) considered alternative viewpoints and weighed options. There were better and worse examples of rationale, evidence, and weighing of concepts in

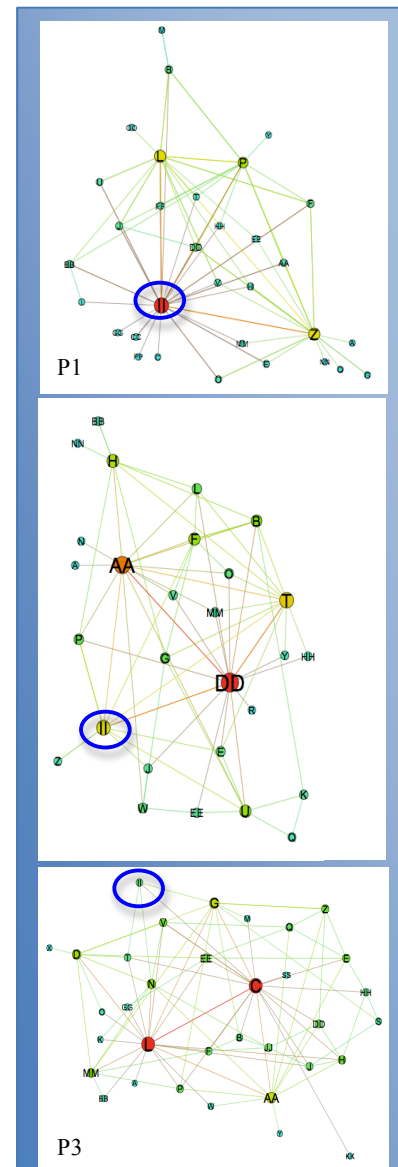


Figure 1. Social network diagrams of Yammer activity at three time periods: P1, P2, and P3. The instructor is labeled “II” and is circled for easier identification. Nodes are color-coded based on centrality from least to most central.

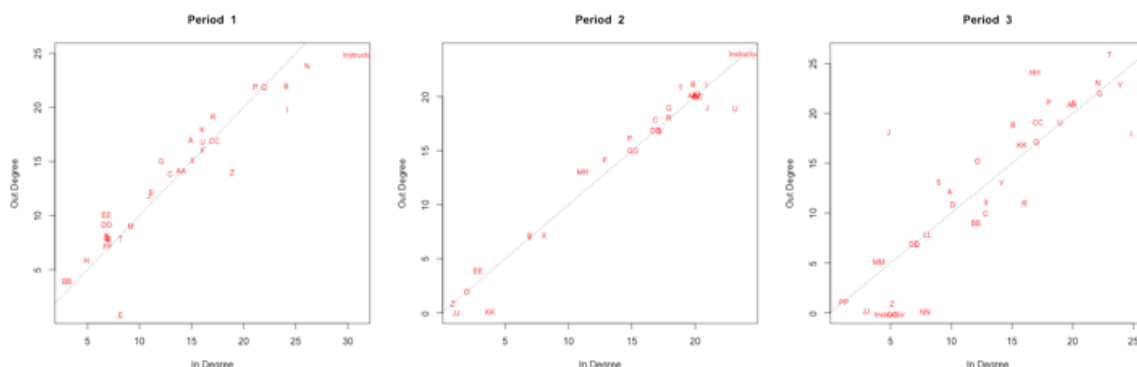


Figure 2. Increasing diversity of participation role across three time periods

Table 1: Types and levels of talk exhibited by students in the Yammer environment.

	Level/Type	Definitions	% of Posts
Content-related Talk	1- Sharing (SHR)	Discussing content or facts without making claims that are supported by rational or evidence. Cannot include concrete examples related to real-world use, links or attachments to additional material, reference to previous reply, or reflections on learning or thinking processes.	9.1%
	2- Extending (EXT)	Extending available readings or content material by providing links or attachments to additional content (video, article, concrete example, etc.) to help think about or understand course content. May include level 1 behaviors, but cannot include rational, concrete examples related to real-world use, reference to previous reply, or reflections on learning or thinking processes.	17.3%
	3- Checking/ Repeating (C/R)	Checking understanding, clarifying what someone previously said, or repeating/rephrasing previous post without adding new ideas. May include level 1 or 2 behaviors, but cannot include new fact, claim, rational, evidence, connections to the real world, or reflections on learning or thinking processes.	1.7%
	4- Synthesizing (SYN)	Connecting info to other info - making connections to the real world, bringing multiple ideas together. Evidence that post is referring to previous post and adding to the idea, or that poster is connecting a previous idea to a real-world example that no one else has mentioned. May include level 1-3 behaviors, but cannot include evidence or rational for claims, or reflections on learning or thinking processes.	31.6%
	5- Interpreting (INT)	Connecting course content info to other info and making a judgment about it or evaluating a claim or work. Claims or opinions MUST be supported with rational or evidence, or they must show that they are weighing ideas against each other. May include level 1-4 behaviors, but cannot show evidence of reflecting on learning or thinking processes.	26.4%
Non-content Related Talk	MC- metacognitive Reflection/Awareness	Thinking about or sharing awareness of learning experiences as an object of thought. Reflecting on learning process, or demonstrating awareness of learning process by articulating learning patterns, or needs as learners. May include level 1-5 behaviors.	4.7%
	CM- Class Management	Posts related to organizing, coordinating, documenting, or deciding on classroom activity.	8.9%

posts. For example, a project team was gathering requirements for a design idea aimed at college students, posted a survey, and asked the class to help by providing data. A fellow classmate responded to the post by saying, “Hey I really liked the way your survey was designed, in terms of asking questions that all relate to each other in a series, and having repeating variables. Good luck with everything!” In this example, the student shares his opinion and supports that opinion with a rationale. For that reason it was coded as a level five, but it was one of the least sophisticated examples found. The following example shows a higher range of sophistication of thinking. The chapter students read, the week of the post, focused on different user interfaces and their affects on user experience. In Yammer, a student started a new thread, asking students to think about different user interfaces, trade-offs and design ideas related specifically to air-based gestures. The student presented Kinect for Xbox360 as an example from the book and asked, “Do you think adding a hand-held motion sensor control option would help to improve the Kinect's interface? Yes, it would make gameplay easier, No, the Kinect is fine how it is, Indifferent, I dislike the Kinect and prefer controller-based games?” Many students responded to this post and what follows is a response rated as a higher range level five post:

“I agree with [Stanley \[tagged\]](#), [Don \[tagged\]](#), and [Jack \[tagged\]](#). I prefer controllers over motion sensor gaming, and I think adding a "sensor controller" would make it way too similar to the Wii. I think the Kinect is a good concept, and is perfect for casual gamers or families who have younger kids, since I find that a lot of Kinect games are geared towards that age group. I think it helps to balance the Xbox360 as a whole to be more marketable to a wider ranged audience. It allows hardcore gamers to continue using controllers while allowing their younger siblings (or possibly children) participate in game time. In response to [Julie \[tagged\]](#), I was able to find statistics that say in 2011, 59 million Xbox 360's were sold and 18 million Kinect's were sold. Not the exact statistics we're looking for, but I think that it shows not all gamers who have 360's are strictly one type of gamer, although I think it does show that more users tend to use the 360 for hardcore over casual gaming.”

The student begins by choosing which students he agreed with, indicating that he read their posts and begins to synthesize their input with his own. He repeats one of the claims made in a previous post, “adding a "sensor

controller" would make it way too similar to the Wii" and "Kinect is a good concept, and is perfect for casual gamers", but then adds his claim about appealing to younger kids. The student supports this claim, stating that Kinect games are mainly marketed to younger audiences. He further supports the claim that adding the Kinect appeals to different audiences, citing facts about the number of Xbox consoles and Kinects sold. He interprets the difference in sales as meaning that only small portions of Xbox users add the Kinect option and concludes that this is indicative of different markets. Here there is evidence of sharing, synthesizing, and interpreting. Interpreting is the highest level of cognitive activity; this post is categorized as a Level 5, interpretation post.

Discussion

Our approach examined the form and function of the social interactions that took place over time in the Yammer environment. Group informatics methods provided visual tools that allowed us to see changes in the structure of social interactions and changes in member activity. Communication analysis informed us as to the type of discourse activity that occurred and whether it matched to the designer's intent. Combining these research methods provided us with a clear understanding of how the instructional approach for using Yammer functioned in practice. Our findings shed light on the use of social technologies as a means to provide a place for students to share, explore, and think about course concepts with others. The learning goals for such a discourse environment are comprehension and application, not pure problem solving. This fills an important gap in research, as previous studies examining collaborative discourse, are directed at assessing the quality of joint problem-solving (Roschelle & Teasley, 1995), creation of common ground (Convertino et al., 2009), or conceptual change (Gunawarda, Lowe, & Anderson, 1997). These are important contributions, but they are not focused on more casual forms of knowledge application and the quality of discourse in a learning community.

Our data suggests that the combination of theoretically informed pedagogical design paired with the use of lightweight social media technology can provide students with opportunities to engage in learning processes associated with a community of learners approach to instruction without requiring unsustainable management practices from instructors. This is an important finding because it paves the way for exploring whether such approaches might scale to larger online learning environments. As online learning environments become more prevalent, a critical consideration should be what kinds of learning experiences do the environments model and support and how will these experiences shape students' understanding of important learning processes and expectations for student-instructor interaction. Our approach enabled students to learn how to become members of a design community, by connecting course content to real-world examples, seeking out ways to better understand course content, and engage in discussions with other students about core concepts and techniques related to human-centered design. Also of importance, it planted a seed in students about what it means to learn that challenged traditional instructional models. Students were also extremely connected within the environment, discussed disciplinary content in fairly sophisticated and meaningful ways, and eventually replaced the instructor as the central information giver and evaluator. Such processes may be even more critical for purely online courses as it may help students to feel more connected to each other and to the course content.

Another important contribution of this work is that presents social media environments as learning environments in their own right, spaces for learning through discourse rather than supplements for transmitting or checking information. These learning environments will likely never be the same as face-to-face interactions and this fact has both costs and benefits. Rather than emphasizing the ways in which these environments fail to emulate real-world learning activity, we must start figuring out how to use the inherent differences of these environments to extend learning opportunities. We can leverage the lack of students' experiences with these environments to create a "space" for students to engage in processes they might avoid in face-to-face interactions. Previous findings related to students' lack of ability in college courses, may be the product of a context that prioritizes task completion over quality of discourse. When groups face time pressures they tend to prioritize task completion over the processes they use to complete tasks (Kerr & Tindale, 2004). This implies that the pressure of time constraints and required activities imposed by college course structures, combined with students' previous experiences, may interfere with students' ability to exercise and improve processes of conceptualization and discourse. Providing students with a less stressful context, new structures, and theoretically informed interaction designs may help to mitigate this problem.

Future studies will examine student perceptions of the Yammer as a learning tool and include a more fine grain analysis of the strengths and weaknesses of the pedagogical design. We also plan to compare differences between resident and online student populations. Through this work we aim to develop a better understanding of the variables associated with rich discourse environments so as to better meet the needs of students and incorporate the application of learning theory into the design of new digital learning environments.

References

Borge, M., & Carroll, J. M. (2010). Using collaborative activity as a means to explore student performance and understanding. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the Disciplines: Proceedings*

- of *ICLS 2010: 9th International Conference of the Learning Sciences*. Chicago, IL: International Society of the Learning Sciences.
- Borge, M., Ganoë, C., Shih, S., and Carroll, J. (2012). Patterns of team processes and breakdowns in information analysis tasks. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA.
- Brown, A. L., & Campione, J. C. (1996). Psychological Theory and the Design of Innovative Learning Environments : On Procedures , Principles , and Systems. *Innovations in Learning: New Environments for Education* (pp. 289–325). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- 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, J. S., Collins, A., & Deguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–42.
- Collins, A., Brown, J. S. & Holum, A. (1991). Cognitive Apprenticeship: Making Thinking Visible. *American Educator*, p.6-11, 38-46.
- Carroll, J. M., (1997). *Minimalism: Beyond the Nurnberg Funnel*. MIT Press, Cambridge, MA.
- Convertino, G., Mentis, H. M., Rosson, M., Slavkovic, A., and Carroll, J. M. (2009). Supporting content and process common ground in computer-supported teamwork. In *Proc. CHI 2002*, ACM press (2002), 2339-2348
- Cuban, L. (1984). *How teachers taught: Consistency and change in American classrooms 1890-1980*. New York: Knopf.
- Dunkin, M. J., & Biddle, B. J. (1974). *The study of teaching*. New York: Rinehart & Winston.
- Dyke, G., Adamson, D., Howley, I., Rosé, C. P. (under review). Enhancing Scientific Reasoning and Explanation Skills with Conversational Agents, submitted to *IEEE Transactions on Learning Technologies*.
- Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education*, 94(1), 103-120.
- Garrison, D. R. (2003). Cognitive presence for effective asynchronous online learning: The role of reflective inquiry, self-direction and metacognition. *Elements of quality online education: Practice and direction*, 4, 47-58.
- Goggins, S., Mascaro, C., & Valetto, G. (2013). Group Informatics: A Methodological Approach and Ontology for Understanding Socio-Technical Groups. *JASIS&T*, 64(3), 516–539.
- Goggins, S. P., Laffey, J., & Gallagher, M. (2011). Completely online group formation and development: small groups as socio-technical systems. *Information Technology & People*, 24(2), 104–133.
- Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of educational computing research*, 17(4), 397-431.
- Jordan, B., & Henderson, A. Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4, 1 (1995), 39-103.
- Landis, J. R., Koch, G. G. The measurement of observer agreement for categorical data. *Biometrics* 33 (1977), 159-174.
- Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annu. Rev. Psychol.*, 55, 623-655.
- Norman, D. A. (2002). *The design of everyday things*. Basic books.
- Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer*. New York: Basic Books.
- Pirolli, P., & Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of International Conference on Intelligence Analysis* (Vol. 5, pp. 2-4).
- Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning* (pp. 69-97). Springer Berlin Heidelberg.
- Rogoff, B. (1994) Developing understanding of the idea of communities of learners. *Mind, Culture, and Activity*, 1(4), 209 – 229.
- Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction design: beyond human-computer interaction*. John Wiley & Sons.
- Schunk, D. H. (2012). *Learning Theories an Educational Perspective* (6th ed.). Boston, MA: Pearson Education, Inc.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & education*, 46(1), 71-95.

Communities of Learning Practice: Balancing Emergence and Design in Educational Settings

Filitsa Dingyloudi and Jan-Willem Strijbos, Ludwig-Maximilians-University Munich, Department of Psychology, Leopoldstr. 13, 80802
Email: Filitsa.Dingyloudi@psy.lmu.de, Jan-Willem.Strijbos@psy.lmu.de

Abstract: This paper explores the aspects of emergence and design in the theoretical notions and models of Communities of Practice (CoPs) and Communities of Learners (CoLs), along with the potential of a new notion and model to epitomize emergent communities in an educational setting. The need for a new notion and model emerged from a longitudinal study of communities in higher education that, among others, aimed to explore the communities' design principles and stages of development. The characteristics of emergent communities in an educational setting were not fully congruent with either CoP or CoL and their theoretical principles. This paper delineates the need for an emergent notion and model, termed "community of learning practice", which incorporates elements from both CoP and CoL, but simultaneously aims for balance between emergence and design in an educational setting to an extent that allows communities to grow through members' voluntary participation and negotiated community structure.

Introduction

A number of notions, including communities of practice, communities of learners, knowledge building communities and learning communities, have been introduced by theoreticians and researchers in the fields of education and knowledge management since the late 1980s and early 1990s to describe social settings within which individuals share and co-construct knowledge, expertise and learning experiences towards a shared enterprise (e.g., Barab & Duffy, 2000; Bielaczyc & Collins, 1999; Brown, 1992; Brown & Campione, 1990; Lave & Wenger, 1991; Scardamalia & Bereiter, 1994; Wenger, 1998a). Although it would have been insightful to theoretically analyze all existing community concepts and their principles, for the purposes of this paper only the concepts of Communities of Practice (CoPs) and Communities of Learners (CoLs) are considered given their increasingly dominant profile supported by theoretical claims in contemporary research literature. Many contributions refer to the notions of CoP and/or CoL, but refer less consistently to the model and underlying theoretical framework that they represent.

CoPs were originally coined as an analytical notion by Lave and Wenger (1991) to describe already existing phenomena in craft production within a situated learning framework. A theoretical analysis of the concept was further developed by Wenger (1998a) highlighting the foundational principles of its meaning, structure and value. The popularity of the concept and its overuse both by researchers and practitioners in different contexts – although not always consistent with its foundational principles and initial conceptualization – has led to conflicting, contradictory, misleading or at least superficial treatment of the term on a theoretical and implementational level (Kimble, 2006; Roth & Lee, 2006; Wenger, 2010). This has contributed either to an instrumental interpretation of the concept used to design educational or organizational settings, or one-way-fit-all interpretations of the term to refer to groups of learners or co-workers (Hughes, 2007; Kimble, 2006; Roth & Lee, 2006; Vann & Bowker, 2001; Wenger, 2010). Wenger (2010) also reflected on the status of the community of practice concept, realizing that it is out of control, since practitioners and researchers have been using it without taking into consideration its theoretical framework and principles.

Apart from CoPs, the notion of CoLs has been also widely used by educators and researchers to refer to communities that aim for the advancement of knowledge and learning how to learn on the classroom level in educational settings (Bielaczyc, Kapur & Collins, 2013; Brown, 1992; Brown & Campione, 1990). Its initial conception by Brown and Campione (1990) aimed to describe the practical implementation of a set of theoretical principles in classrooms, based upon Vygotskian premises and Dewey's (1897) discovery learning, towards a reform movement in education through an innovative classroom design (Brown & Campione, 1994). CoLs in education are associated with a shift of the students' role in schools from passive knowledge recipients to active co-constructors of knowledge, while being responsible for their own learning and its design (Brown, 1992). Therefore, CoLs have been used as an educational technique to enhance knowledge sharing and distribution of expertise and responsibility among students and teachers aiming to build a collective learning culture in which students respect and value each other's contributions (Bielaczyc, Kapur, & Collins, 2013; Brown & Campione, 1994, 1996).

According to the foundational theoretical principles of the notions of CoP and CoL, the importance that has been attached to the aspects of emergence and design varies across the two community notions and models. CoPs are built on the idea of self-generation, self-directedness and voluntary participation, drawing their energy

and motivation from within the community itself (Janson, Howard, & Schoenberger-Orgad, 2004), whereas CoLs are designed and facilitated by educators, researchers or a combination, and typically applied into a classroom setting with fixed participants (i.e., students and teachers/adults) and pre-defined pedagogical objectives based on the curriculum (Rogoff, 1994). This paper aims to examine the emergence and design aspects in both community notions and models, which are represented in different ways in each one of the two models, and introduce the need for a new model to describe communities that emerged in an educational setting through the voluntary participation of learners, while recombining elements from both existing models – yet not fully fitting to either of both precursors.

Before the recombinant model can be discussed, it is of vital importance to understand the theoretical framework from which both community notions emerged and which elements are represented in the theoretical principles of both concepts. Despite the differences between both notions of communities, which will be explored later in this paper, they share similar theoretical underpinnings grounded in sociocultural approaches to learning.

Sociocultural Approaches to Learning: Situated Learning and Situated Cognition

Sociocultural approaches to learning gained increasing attention in the 1980s by researchers to explore alternative ways of understanding learning beyond behaviorism and cognitivism (John-Steiner & Mahn, 1996; Wertsch, 1991, 1998). Sociocultural approaches to learning are often associated with Lev Vygotsky and his collaborators' works on learning and development of human behavior in the 1920s and 1930s (Vygotsky, 1978). Vygotsky, in his conceptualizations of development, learning and knowledge construction from a sociocultural perspective, emphasized the dynamic interdependence of socially shared activities and internalized processes in the co-construction of knowledge, since the internalization that takes place is situated within socially shared activities (John-Steiner & Mahn, 1996).

Situated learning and situated cognition developed as sociocultural approaches from Vygotsky's sociocultural theory. Both approaches emphasize that learning is a socially situated practice and learning, thinking, and knowing constitute relations among individuals, who engage in authentic collective activities, and arise from the surrounding sociocultural setting (Brown, Collins, & Duguid, 1989; Lave, 1988, 1990; Lave & Wenger, 1991). Yet, they differ with regard to how they could be implemented in educational settings.

Situated Learning

Situated learning is interconnected with the notion of CoP and refers to learning as “not merely situated in practice – as if it were some independently reifiable process that just happened to be located somewhere; learning is an integral part of generative social practice in the lived-in world” (Lave & Wenger, 1991, p. 35). The constituent characteristic of engagement in social practice, within which learning occurs, is legitimate peripheral participation (LPP) and refers to the process by which newcomers move towards becoming more central participants of a CoP (Lave & Wenger, 1991). According to Anderson, Reder and Simon (1996), the foundational claims of situated learning are often misimplemented in educational settings, however, Lave and Wenger (1991) specifically highlighted that

“(…) legitimate peripheral participation is not itself an educational form, much less a pedagogical strategy or a teaching technique (...) this view makes a fundamental distinction between learning and intentional instruction (...) this is very different from attributing a prescriptive value to the concept of legitimate peripheral participation and from proposing ways of “implementing” or “operationalizing” it for educational purposes.” (Lave & Wenger, 1991, pp. 40-41)

This quote depicts that situated learning's constituent characteristic of legitimate peripheral participation cannot be operationalized as an instructional approach with pre-defined educational purposes.

Situated Cognition

Situated cognition, proposed by Brown, Collins and Duguid (1989), argues that meaningful learning can only occur when embedded in the social context within which it is used to “(...) enculturate students into authentic practices through activity and social interaction” (Brown, Collins, & Duguid, 1989, p. 37). Situated cognition, as a theoretical model, can be translated into educational practice through the instructional approach of cognitive apprenticeship, and aims to foster learning in a domain by promoting students' knowledge acquisition and development through the use of cognitive tools in authentic activities (Brown, Collins, & Duguid, 1989; Collins, Brown, & Newman, 1989). Therefore, situated cognition positions itself closer to innovative educational reform and intentional instructional approaches, and is implicitly connected – through their focus on instructional purposes – with the notion of CoL.

Two Contemporary Notions of Communities

Over the past two decades, the notions of CoP and CoL have flourished in research, and educational and work-place settings. Despite their popularity, their slightly different theoretical foundations have slipped out of focus. The following section will describe both notions of community – in close alignment to their foundational elements and theoretical principles – and the degree of emergence and design.

Communities of Practice

Within the framework of sociocultural theory and situated learning, Lave and Wenger coined the multifaceted concept of CoP to refer to “(...) a set of relations among persons, activity and the world, over time and in relation with other tangential and overlapping communities of practice” (Lave & Wenger, 1991, p. 98). In other words, a CoP is formed by individuals who have a common interest in a domain of human endeavor and mutually engage in a process of social learning, working together and sharing ideas in order to collectively solve problems and co-construct knowledge over a period of time (Wenger, 1998a).

CoPs share a common structure, which is constituted of three elements: domain, community and practice (Wenger, McDermott, & Snyder, 2002). The element of domain refers to a shared domain of interest that matters to members (Wenger, 1998a), contributes to the identity of the community which reflects “a set of issues, challenges and passions through which members recognize each other as learning partners” (Wenger, White, & Smith, 2009, p. 5). Individuals’ membership in a CoP is related to their dedication to that domain along with a communal competence that diversifies the members of the community from others, without necessarily requiring high expertise in a specific field (Wenger, 1998a). Members of the community tend to appreciate their collective competence and learn together, even though people outside of the community might not value or identify with their collective experience (Wenger, 1998a). The element of community is evident when individuals engage in joint activities and discussions, help each other, and share information, while building relationships that enable members to interact and learn from each other (Wenger, 1998a). This does not imply that members of a CoP work together on a consecutive basis, although interactions are essential in the constitution of a CoP (Wenger, 1998a). The element of practice refers to a shared repertoire of resources, like experiences, tools and ways of addressing issues (Wenger, 1998a). Hence, a CoP should not be conceptualized merely as a community of people with common interests, but rather as a community whose members develop a shared practice. The process of a sharing practice requires time and interaction and seems to be a more or less self-conscious process (Wenger, 1998a). Wenger (1998a) further defined the relationship between practice and community, within the framework of CoPs, by three core dimensions: mutual engagement, joint enterprise and shared repertoire (see Table 1).

Table 1: Dimensions of practice as the property of a community

Core dimension	Description	Implication	Value
Mutual engagement	Engagement in collectively negotiable actions	Does not imply homogeneity as a prerequisite for community development	Recognition of the importance of others’ competences, knowledge and contribution
Joint enterprise	Negotiated response to address their common situation	Moves beyond a common goal to involve relations of mutual accountability	Mutual negotiation of actions and sharing responsibility for the common enterprise
Shared repertoire	Developed routines, activities, artifacts and stories	Can be used as reference points for negotiation of meaning	Negotiation of new meaning and dynamic development

The mere existence of these dimensions does not imply a coherent community, but rather whether they are collectively used by the participants for a common enterprise (Wenger, 1998a). The major challenge of a CoP is to collectively identify the community, continuously negotiate the reasons of its existence and sustain a social space for learning along with being involved and committed over time (Wenger, Trayner, & DeLaat, 2011).

Emergence and Design in CoPs

CoPs do not constitute formal structures among individuals, but informal entities within the minds of its members, built around the relationships that members develop with each other while sharing common problems and interests (Ardichvili, Mauer, Li, Wentling, & Stuedemann, 2006). As Kirschner and Lai (2007) explain, CoPs are not things, but processes in which social learning takes place. A CoP is perceived as a self-generated and self-organized system, an organism or an autonomous group that needs time to form itself, exists for the benefit of its members, and may continue its existence even after the completion of a project or task (Lave &

Wenger, 1991; Kimble, 2006; Wenger, 1998a; Wenger, McDermott, & Snyder, 2002). According to Wenger (1998b), communities go through various stages of development characterized by different kinds of activities and different levels of interaction among community members. Therefore, a CoP differentiates itself from other sets of people who operate in groups, because it neither starts its existence and its action the exact moment that a task or project starts, nor disappears at the end of it (Wenger, 1998a).

Since learning occurs in the sphere of experience and practice, CoPs cannot be designed in the sense of producing an “organizational unit” or being implemented as a pedagogical technique (Wenger, 1998a, p. 228), but they can be “recognized, supported, encouraged, and nurtured” (Wenger, 1998a, p. 229). However, as Wenger (2010) explains, “(...) it is indeed difficult to find the right balance between enough formality to give them legitimacy in the organization and enough informality to let them be peer-oriented, self-governed learning partnerships” (¶ A co-opted concept: on the instrumental slippery slope). In contrast to self-generated and self-developed CoPs, the emergence and design of CoLs in educational settings underlies intentionality and particular objectives.

Communities of Learners

Based upon the premise of sociocultural theory and situated cognition, the conception of CoL emerged in the early 1990s as part of the *Fostering Communities of Learners* project by Brown and Campione (1990), which aimed to transform traditional classrooms into innovative learning environments that foster the distribution of expertise among adults and children underlying a need for educational reform (Brown, 1992, 1994; Brown & Campione, 1996). According to Rogoff (1994), learning emerges in CoLs when individuals participate in shared enterprises with others by having active, but often different, roles in sociocultural activities. According to Rogoff, Goodman-Turkianis and Bertlett (2001), a community is constituted of relationships among individuals, who collectively try to accomplish common enterprises, while being stably involved, and considering those multifaceted relationships as fundamental for the evolution of the community. The aim of a CoL moves beyond successfully accomplishing tasks to developing collective practices that transcend the particular individuals involved (Rogoff, Goodman-Turkianis, & Bertlett, 2001).

A variety of CoL models have been operationalized and implemented by researchers over the years (Bielaczyc, Kapur & Collins, 2013; Brown, 1992; Brown & Campione, 1996; Rogoff, 1994; Rogoff, Goodman-Turkianis, & Bertlett, 2001; Scardamalia & Bereiter, 1994). Regardless of the differences in operationalization and implementation of the models, which move beyond the scope of this paper, they share several underlying principles. According to Bielaczyc, Kapur and Collins (2013), the constituent characteristics of CoL models are: diversity of expertise among participants through valued contributions, a shared objective of developing collective knowledge and skills over time, learning how to learn, and mechanisms for sharing. The main objective of a CoL is to promote a learning culture, in which individuals and the community as a whole are learning how to learn and develop disciplinary knowledge, while respecting and valuing different contributions of its members (Bielaczyc, Kapur, & Collins, 2013).

Emergence and Design in CoLs

CoLs have been used as a community-based instructional approach by educators who aim to foster student-centered learning, therefore CoL is associated with a shift in the learners’ role in schools from a passive knowledge recipient to an active co-constructor of knowledge responsible for their own learning (Brown, 1992). Intentional CoLs by educators often start with a course, operate on the classroom level, assume that the classroom is one community, and the participants are students whose individual achievement is evaluated and rewarded and prioritized over the collective success (Barab, Kling, & Gray, 2004; Roth & Lee, 2004). CoLs are supported and guided by a teacher-facilitator or other adults (Brown, 1992). However, the implementation of the CoL notion by educators in classrooms has been a challenging process and often misrepresented the notion itself (Barab, Klinef, & Gray, 2004), which has been often employed by educators “(...) as a slogan rather than as an analytical category” (Barab, Kling, & Gray, 2004, p. 3).

A Recombinant Model of Communities: Communities of Learning Practice

Despite the differences between the notions of CoP and CoL, their commonalities have contributed to theoretical and implementational confusion. In an attempt to clarify theoretical contradictions in the notion of CoP, Roth and Lee (2006) claim that “(...) the notion of community in the context of classrooms is inappropriate or even false – unless the students concretely realize the collectively defined motive and have some choice and control in the matters” (p. 32). Indeed, it is misleading to consider whole classrooms as CoPs, because the notion of CoP entails emerging elements and not pre-defined, forced attributes to students in classrooms. However, several CoPs may exist within the context of the classroom or the broader educational setting, in the same way that they may exist in any other social context. In contrast, CoLs have been used in educational settings as an instructional approach by educators to foster students’ control and responsibility in collectively developing knowledge through sharing of contributions, which are respected and valued by others

(Rogoff, 1994). Therefore, classrooms were transformed into CoLs as part of an educational reform movement. The CoL notion does not include emergent elements, but instructionally designed and defined ones. Is there room for a middle path between those two notions by employing theoretical values from both? What could a middle path be for a community model to emerge in an educational setting and be “designed” to an extent that such a “design” does not interfere with its underlying “emergent” values and principles?

Communities of Learning Practice

During a longitudinal study in higher education, in which we explored the aspects of design and emergence, developmental stages (Dingyloudi & Strijbos, 2014), sharing practices and value creation in multiple extra-curricular emergent communities in parallel to a Learning Sciences master’s program, we initially expected to be able to describe our communities under study by either adopting the CoP or the CoL model. However, as our research progressed it proved impossible to provide a description of our emergent communities that would be compatible with either the CoP or CoL model. How our communities under study emerged, were structured and evolved generated the need for a new conceptual descriptor, which operates in-between and beyond the existing models. This third notion aims to describe emergent communities by incorporating elements from both CoP and CoL models, but simultaneously creating a unique profile without being a mere variation of either of the two models.

We have termed this third notion “Communities of Learning Practice” (CoLP) (See Figure 1). The inclusiveness of this notion is initially depicted by using the terms “learning” and “practice” in its conception. The relationship between learning and practice seems to be realized at two levels in the CoLP model. It can be either perceived as (a) “learning how to practice” (e.g., academic skills) that has the potential to contribute to the broader educational context, or (b) “practicing how to learn”, through the practice of skills within the community setting. How participants may experience the relationship between learning and practice may vary, as well as how they value this relationship.

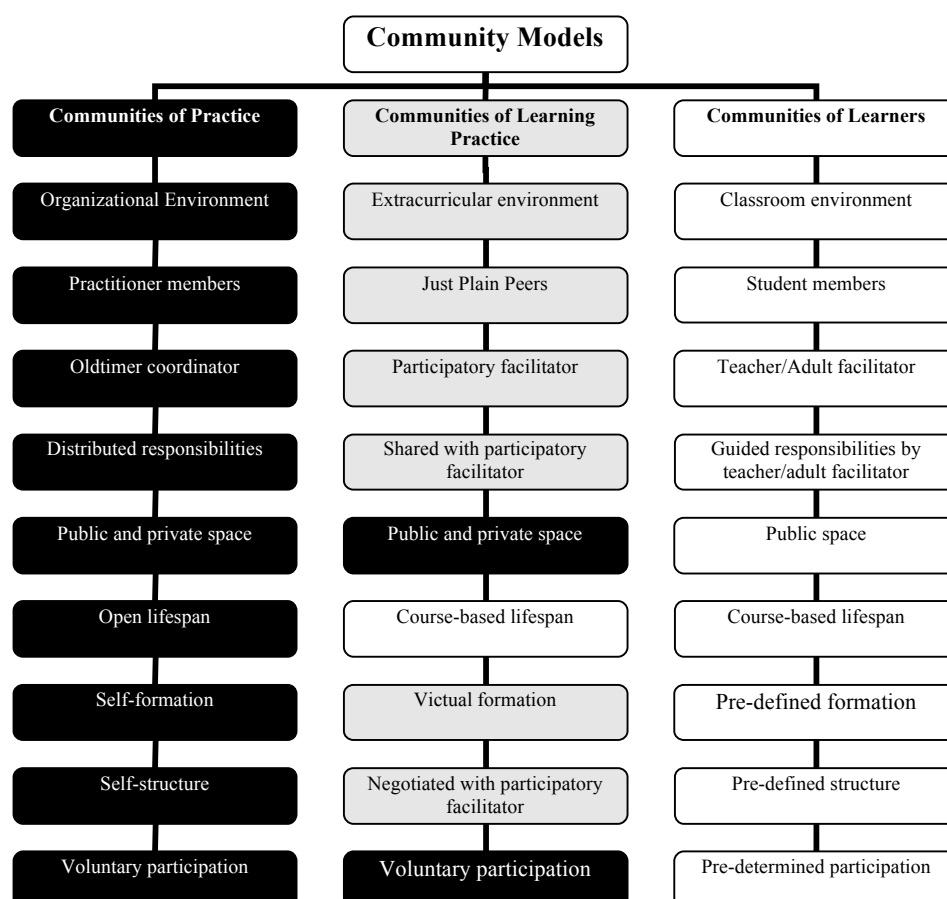


Figure 1. A recombinant model of Communities of Learning Practice (CoLP)

CoLPs derive from and operate in educational settings, and in parallel with – but not integrated into the curriculum – by having no pre-defined pedagogical objectives. Therefore, CoLPs are extra-curricular entities that emerge from students’ common needs and are not used as an instructional approach by educators,

researchers or stakeholders to foster curricular learning objectives. With respect to the environment within which communities grow, CoLPs and CoLs are both formed in an educational setting, but the CoLP lives *in parallel* as opposed to *in* the curricular setting. Although CoPs could also be formed in educational settings, they typically find better ground for development in organizations in an attempt to contribute to organizational success through transfer of professional skills and profitable work place practices (Wenger & Snyder, 2000).

The potential participants of CoLPs are peers, who are students in the broader educational setting, and gather together to address identified common needs (e.g., academic challenges) that derive from the broader educational setting. Their participation throughout the CoLP lifespan is voluntary and members are free to withdraw whenever the value of their participation fades away. CoLPs are open to any student who wishes to act as a peer among peers and share, negotiate and co-construct learning experiences through the sharing mechanism of peer feedback. The participants of CoLPs can be described as *Just Plain Peers* (JPPs) as a context-specific instantiation of the term “Just Plain Folks” (JPFs) introduced by Lave (1988). Following Brown, Collins and Duguid (1989), the issues and problems that Just Plain Folks face “(...) arise out of, are defined by, and are resolved within the constraints of the activity they are pursuing” (p. 35). Table 2 provides a summary of the features characterizing JPFs/JPPs, students, and practitioners. CoLs typically involve students as participants, although CoLs are designed to enhance a shift in students’ activity in classrooms and move beyond laws, symbols and fixed meanings. CoPs may have as participants either practitioners or JPFs. The JPPs in a CoLP resemble the activity of practitioners, although JPPs are more closely related to everyday activity.

Table 2: JPFs/JPPs, Students and Practitioners (adapted from Brown, Collins, & Duguid, 1989)

	JPFs/JPPs	Students	Practitioners
reasoning with:	causal stories	laws	causal models
acting on:	situations	symbols	conceptual situations
resolving:	emergent problems and dilemmas	well-defined problems	ill-defined problems
producing:	negotiable meaning and socially constructed understanding	fixed meaning and immutable concepts	negotiable meaning and socially constructed understanding

CoLPs act as a support structure for students, are generated by the students themselves but facilitated by a *participatory facilitator*, who is neither an expert nor a peer-member. The participatory facilitator and the JPPs *share responsibility* for coordinating and co-structuring the *public space* and reinforcing social interactions in the *private space*. The public space refers to informal events which are open to community members to exchange ideas, and find solutions to their problems, attributing a ritualistic and substantive dimension to the community by offering its members the experience of participation and meeting other members. The private space refers to a web of relationships among members who can be involved into one-to-one informal discussions that inform any design decision of the public space (Wenger, McDermott, & Snyder, 2002). In contrast, the facilitator in CoLs is the teacher or another expert (Brown & Campione, 1996), whereas in CoPs the coordination role can be taken on by any of the community members (Wenger, McDermott, & Snyder, 2002). CoLs and CoPs further differ regarding the spheres within which they operate, since the CoL operates in the public space (e.g., classroom) whereas the CoP operates in both the public and private space.

CoLPs are *virtually formed*, since the participatory facilitator communicates the invitation for a CoLP to be initiated (if relevant to the JPPs) to the potential JPPs. The participatory facilitator also supplies the community with ritualistic sustainability (i.e., a consistent rhythm of public community events), enhances the informality of the public community events, facilitates the “learning how to practice” and “practicing how to learn” aspects of the community on a public level, and builds private social bridges with the JPPs with the aim to reinforce community aliveness (Wenger, McDermott, & Snyder, 2002). Therefore, the community formation is not entirely self-generated (as is the case for a CoP), but there is a virtual approach to this formation by the participatory facilitator, i.e. supplying the community with mechanisms that facilitate community sustainability and members’ social bounding. Along with the virtual formation, the structure of the community is also *negotiated* among the JPPs and the participatory facilitator. Structure refers to the ritualistic and structural dependence of the members on each other, which is negotiated throughout the lifespan of the CoLP. Contrasting the formation and structure of CoLPs with both two existing community models – CoLs are instructionally pre-defined, pre-structured and guided, whereas CoPs are self-formed and self-structured – it becomes evident that the aspects of emergence and design in CoLPs do not reflect either the CoL or CoP model, and move beyond them by employing a balanced approach to emergence and design.

The lifespan of CoLPs is course-based, given the broader socio-educational context within which they evolve. Since the relevance of forming the CoLP derives from the broader educational context, this context constitutes the defining element of its lifespan as well. However, the CoLP might regenerate or transform itself

as long as the JPPs continue to identify potential values of participation in their CoLP. Regarding lifespan the CoLP resembles a CoL and differentiates a CoLP from a CoP, which does not start or end with a task or project (Wenger, 1998a).

Balancing Emergence and Design of Communities

The rationale for another community notion relies on the need for balancing emergence and design of a community in an educational setting. According to Wenger (1998a), the balance between emergence and design in the community context is depicted in the paradox that “(...) no community can fully design the learning of another (...) no community can fully design its own learning.” (p. 234). The emphasis expressed by “fully” in Wenger’s statement is what makes the relationship between emergence and design complex and asks for an equilibrium between the two.

Wenger, McDermott and Snyder (2002) suggest a set of design principles that could foster a sense of community aliveness without designing an intervention to achieve this. Hence, designing a CoP would be contradictory with the underpinning theory of self-organized systems. In contrast, CoLs do not allow for much freedom on behalf of the community members, since the participation and objectives are pre-defined and pre-structured for pedagogical purposes. Therefore, in an attempt to achieve an equilibrium between emergence and design – without being conflicting or incompatible with any of the CoL and CoP notions – the CoLP constitutes a middle path that incorporates community values from both notions *and* stands on its own as well.

Implications and Conclusion

This paper explored the aspects of emergence and design that are attached to the dominant notions of CoP and CoL and presented the need for a new community notion to characterize emergent communities in education. This new notion of “Communities of Learning Practice” (CoLP) emerged from our longitudinal study on emergent communities in higher education, during which we aimed to achieve a balance between emergence and design. CoLP incorporates elements from both community models, constituting an entity in-between and beyond them. Metaphorically, this resembles the reproduction of living organisms, in which a new living entity is generated by recombining different chromosomes from two other living organisms.

The emergence of the CoLP notion has theoretical and practical implications. First of all, the emergence of the notion by itself calls researchers’ and educators’ attention to the aspects of emergence and design within either the CoP or the CoL model, highlighting the importance of considering them before attributing the label of CoP or CoL to the communities under study or being implemented, to avoid theoretical misinterpretations. In addition, a CoLP and the community model it represents, offers a theoretical model for a community that is neither self-generated and self-developed, nor instructionally designed and pedagogically guided. This model has the potential to describe emergent communities in educational settings that move beyond the curriculum and the constraints of the classroom supported by a participatory facilitator. With respect to the practical implications in education, CoLPs suggest an even more autonomous community model compared to the CoL that educators should allow to emerge out of the voluntary participation of the students and let them be JPPs for their own emergent needs and problems, which might not necessarily be addressed in the curriculum. A CoLP model cannot be used as an instructional approach, but it should be provided the ground, space and time, to flourish.

References

- Ardichvili, A., Mauer, M., Li, W., Wentling, T., & Stuedemann, R. (2006). Cultural influences on knowledge sharing through online communities of practice. *Journal of Knowledge Management*, 10, 94-107.
- Barab, S. A., & Duffy, T. M. (2000). From practice fields to communities of practice. In D. H. Jonassen & S. M. Land (Eds.), *Theoretical foundations of learning environments* (pp. 25-55). Mahwah, NJ: Lawrence Erlbaum Associates.
- Barab, S. A., Kling, R., & Gray, J. H. (Eds.) (2004). *Designing for virtual community in the service of learning*. Cambridge: Cambridge University Press.
- Bielaczyc, K., & Collins, A. (1999). Learning communities in classrooms: A reconceptualization of educational practice. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory* (pp. 269-292). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bielaczyc, K., Kapur, M., & Collins, A. (2013). Cultivating a community of learners in K-12 classrooms. In C. E. Hmelo-Silver, A. M. O’Donnell, C. Chan & C. A. Chinn (Eds.), *International Handbook of Collaborative Learning* (pp. 233-249). New York: Taylor & Francis.
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, 2, 141-178.
- Brown, A. L., & Campione, J. C. (1990). Communities of learning and thinking, or a context by any other name. *Contributions to Human Development*, 21, 108-125.

- 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. 229-270). Cambridge, MA: MIT Press.
- 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.), *Innovations in learning: New environments for education* (pp. 289-325). Mahwah, NJ: Erlbaum.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated learning and the culture of learning. *Educational Researcher*, 18, 32-42.
- 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: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dewey, J. (1897). My pedagogical creed. *School Journal*, 54, 77-80.
- Dingyloudi, F., & Strijbos, J. W. (2014). Evolution of communities of learning practice in higher education: Collective unites of analysis. In *Proceedings of the ICLS 2014 conference "Learning and becoming in practice"*. Boulder, CO: International Society of the Learning Sciences.
- Hughes, J. (2007). Lost in translation: the journey from academic model to practitioner tool. In J. Hughes, Jewson, N., & L. Unwin (Eds.), *Communities of practice: critical perspectives*. New York: Routledge.
- Jackson, J. (2008). *Language, identity and study abroad: Sociocultural perspectives*. London, UK: Equinox.
- Janson, A., Howard, L., & Schoenberger-Orgad, M. (2004). The odyssey of Ph.D. students becoming a community of practice. *Business Communication Quarterly*, 67, 168-181.
- John-Steiner, V., & Mahn, H. (1996). Sociocultural approaches to learning and development: A Vygotskian framework. *Educational Psychologist*, 31, 191-206.
- Kimble, C. (2006). Communities of practice: Never knowing undersold. In E. Tomadaki & P. Scott (Eds.), *Innovative approaches for learning and knowledge sharing: EC-TEL workshop proceedings* (pp. 218-234). Knossos, Crete: Sun SITE Central Europe.
- Kirschner, P. A., & Lai, K.-W. (2007). Online communities in education. *Technology, Pedagogy and Education*, 16, 127-131.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.
- Rogoff, B. (1994). Developing understanding of the idea of communities of learners. *Mind, Culture, and Activity*, 1, 209-229.
- Rogoff, B., Goodman-Turkanis, C., & Bartlett, L. (2001). *Learning together: Children and adults in a school community*. New York: Oxford University Press.
- Roth, W.-M., & Lee, Y.-J. (2006). Contradictions in theorizing and implementing communities in education. *Educational Research Review*, 1, 27-40.
- Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *The Journal of the Learning Sciences*, 3, 265-283.
- Vann, K., & Bowker, G. (2001). Instrumentalizing the truth of practice. *Social Epistemology*, 15, 247-262.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Wenger, E. (1998a). *Communities of practice: Learning, meaning and identity*. Cambridge: Cambridge University Press.
- Wenger, E. (1998b). Communities of practice: Learning as a social system. *System Thinker*, 9, 1-5.
- Wenger, E. (2000). Communities of practice and social learning systems. *Organization*, 7, 225-246.
- Wenger, E. (2010). *Communities of practice and social learning systems: The career of a concept*. Retrieved November 8, 2010, from <http://wenger-trayner.com/wp-content/uploads/2012/01/09-10-27-CoPs-and-systems-v2.01.pdf>
- Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Boston: Harvard Business School.
- Wenger, E. C., & Snyder, W. M. (2000). Communities of practice: The organizational frontier. *Harvard Business Review*, Jan-Feb, 139-145.
- Wenger, E., White, N., & Smith, J. D. (2009). *Digital habitats: Stewarding technology for communities*. Portland, OR: CPsquare.
- Wenger, E., Trayner, B., & De Laat, M. (2011). *Promoting and assessing value creation in communities and networks: A conceptual framework*. Heerlen, the Netherlands: Ruud de Moor Centrum.
- Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.
- Wertsch, J. V. (1998). *Mind as action*. New York: Oxford University Press.

Dynamic Visualization of Motion for Student-Generated Graphs

Jonathan M. Vitale, Kevin Lai and Marcia C. Linn, University of California, Berkeley
jonvitale@berkeley.edu, laikevin@gmail.com, mclinn@berkeley.edu

Abstract: Graph construction and interpretation are critical 21st-century skills. In this study we investigate how 8th grade students construct graphs in the context of a week-long online curriculum unit that links dynamic visualizations to graphical data. We test two forms of visualization: dual animation depicting both the student's graph and the correct graph in terms of a narrative context and single animation depicting only the student's graph. Quantitative results indicate that both forms of animation supported understanding, but dual animation facilitated construction of more accurate graphs earlier in the unit. Case studies reveal unique graphing patterns associated with each form of animation.

Introduction

In this research we compare two forms of graph visualization to explore how new technologies can improve graph understanding. Graphing is a critical mathematical and scientific skill, as well as a key practice used in personal and policy decisions. Growth in the field of data science suggests that the ability to apply graphing knowledge in a broad range of contexts will be an increasingly valuable 21st-century skill. As such, new standards assert that students translate between quantitative, graphical and narrative forms (NGACBP, 2010). Given documented student difficulties with graph understanding (Shah & Hoeffner, 2002), we seek tools that can help students interpret graphs. We test a graph visualization tool under two conditions to explore how feedback on graph construction can improve graph understanding.

Graph understanding involves a variety of thinking skills, ranging from basic perceptual processing to highly complex cognitive reasoning (Shah & Hoeffner, 2002). Successful graph construction and interpretation requires students to coordinate between multiple representations (e.g. text, tables, and graphs in various formats) and sources of knowledge (e.g. perceptual, contextual, and mathematical). Given the complexity, iterative refinement of instruction and testing of comprehensive alternatives is valuable (Quintana et al., 2004).

To promote graph understanding we designed instruction using the Web-based Inquiry Science Environment and the knowledge integration (KI) framework and tested alternative uses of our visualization tools (Linn & Eylon, 2011). The KI framework has proven useful for design of instruction featuring dynamic visualizations (Ryoo & Linn, 2012). The pattern guides students to articulate predictions, add new ideas, distinguish between predictions and new ideas, and reflect on the problem. When students articulate their ideas, including non-normative ideas, they are prepared to compare and contrast these ideas with alternatives introduced in the curriculum. By facilitating reflection on contrasting ideas, technology can support robust conceptual change.

While the general pattern of the KI framework is widely applicable, the specific activities associated with each aspect of the KI pattern are dependent on the conceptual domain and available technologies. For graphing, we can elicit ideas by asking students to interpret a graph or to construct a graph about a complex situation. Research shows that students often generate a "pictorial" interpretation of the graph rather than capturing a relationship (Leinhardt, Zaslavsky, & Stein, 1990). For position vs. time plots (position-time), students often make a pictorial interpretation when they report that a flat line segment represents an object moving forward along a straight path, or that a non-zero slope represents an object moving up or down (Mokros & Tinker, 1987). Additionally, students may generate a graphical representation of an object "moving back" to a reference point by drawing a line that essentially goes back in time. This response does not recognize that a position-time graph will always move forward in time (i.e., from left to right).

Dynamic visualizations (e.g. animations) have successfully promoted acquisition of ideas about graphs (Moreno, 2001) and specifically about motion graphs (Imhof, Scheiter, Edelmann, & Gerjets, 2013). For example, in the SimCalc software MathWorlds (Roschelle, Kaput, & Stroup, 2000), students view the animation of an elevator that moves according to the shape of a segmented line graph. Likewise, Ploetzner, Lippitsch, Galmbacher, Heuer, and Scherrer (2009) depict line graphs through an animation of a runner. By observing these linked representations students have the opportunity to discover relationships between abstract spatial features (e.g. slope) and narrative events.

Our research extends this line of investigation to activities that help students distinguish among ideas about motion graphs. Distinguishing requires that students develop and apply criteria to evaluate ideas and then compare multiple ideas based on these criteria. This multi-step process represents a significant design challenge, given inherent limitations on students' working memory (Sweller, 1988) and ability to attend to multiple visual phenomena within the same sensory modality (Mayer, 2001). A general solution is to reduce cognitive load by distributing information to the environment so that the learner has less information to maintain in working memory (Zhang & Norman, 1994).

For narrative-based position-time graphs, ideas may be elicited by prompting students to graph the narrative. Animated feedback depicting the motion inherent in these graphs (“student animation”) then provides an opportunity to evaluate ideas by comparing observed motion to the imagined motion of narrative events. While this may be straightforward for simple narratives, for complex narratives the burden of remembering the sequence of events may overwhelm students. Furthermore, for subtle changes in motion the student may not clearly recognize any incongruity with the narrative text, as he or she interpreted it. In this case the mental reference by which the animation is evaluated is challenging to construct and difficult to maintain in memory.

A potential remedy is to provide an additional concrete reference animation that accurately depicts the narrative (“normative animation”). While this design would seem to reduce the burden on students, there may be unintended consequences of introducing a new visual representation. For example, splitting attention between the two animations may inhibit comprehension (Mayer, 2001). Additionally, by relying solely on the normative animation, rather than generating a mental simulation of the narrative, students may miss an important learning opportunity (Black, Segal, Vitale, & Fadjo, 2012; Vitale, Black, & Swart, 2013). Furthermore, reliance on visual feedback has been associated with application of opportunistic trial-and-error strategies, which can interfere with learning in a digital environment (e.g., Logo, Cope & Simmons, 1994).

In this research we compare the use of dual animations (Student+Normative condition) to a single animation (Student-only condition). We hypothesize that the Student+Normative condition will facilitate the production of more accurate graphs but anticipate that each condition has some advantages. We take an exploratory approach to analysis of student strategies, and how they reflect affordances of the graphing tools. We take an in-depth look at individual students’ graphing artifacts to evaluate how emergent strategies may promote or inhibit learning.

Methods

Participants and Procedures

Three teachers from two suburban middle schools chose to participate in this study. A total of 384 eighth grade students participated in some part of this study. Out of these students 333 students completed the pretest, (some part of) the curriculum unit, and the posttest. The populations served by school these schools differed markedly in income-levels and home language [School A: N = 319, 7% reduced lunch, 3% ELL; School B: N = 65, 63% reduced lunch, 25% ELL]. Both pretest and posttest were administered to children individually. Pairs of students were assigned to collaborative workgroups by their teacher to work on curriculum. In one class within School B students worked individually. Workgroups were randomly assigned to a condition [S+N or S-only] by the software and received one of two sets of similar activities. While graphing, students seeking help were directed by either the classroom teacher or researcher to complete the graph and observe the animated feedback.

Curricular Materials

This study was conducted in the context of a curriculum module entitled *Graphing Stories*. The goal of the unit was to familiarize students with the process of constructing graphs and interpreting graphs in terms of narratives. The curriculum was developed within WISE (Web-based Inquiry Science Environment), utilizing a variety of instructional and assessment tools (Linn & Eylon, 2011). Within a WISE module each page displayed on screen is referred to as a “step”. Groups of related steps are referred to as “activities”.

Table 1: Curriculum layout for “Graphing Stories”

#	Activity Title	Description/Goals	Relevant Steps
1.	<i>Graphs Tell a Story</i>	Intro to graphing concepts.	
2.	<i>Retell the Story with Graphs</i>	First narrative graphing exercise with animated feedback.	2.2/2.3/2.4 – Table/Graph “Bear Story”/Feedback (<i>Bear</i> item)
3.	<i>Graphing Movement</i>	Direct instruction on the relationship of slope to speed and direction.	3.3/3.4 – Graph person standing still/Feedback 3.7/3.8 – Graph person moving forward and back/Feedback 3.11/3.12 – Graph person moving fast and slow/Feedback
4.	<i>Trading Graph Instruction</i>	Peer review of narratives based on graphs.	
5.	<i>Drive Your Car</i>	Second narrative graphing exercise, with student control of axes to adjust scale.	5.2/5.3/5.4 – Table/Graph “Sam’s Journey”/Feedback (<i>Sam</i> item) 5.6 – Graph “Rita’s Journey”
6.	<i>Your Story</i>	Personal narrative and graphing exercise.	

Table 1 displays the general layout and features of the *Graphing Stories* curriculum unit. In the context of this study we highlight the steps that feature table and graph construction, followed by an animated feedback. Figure 1 display the feedback layout for the “Bear story” and “Sam’s journey,” respectively.

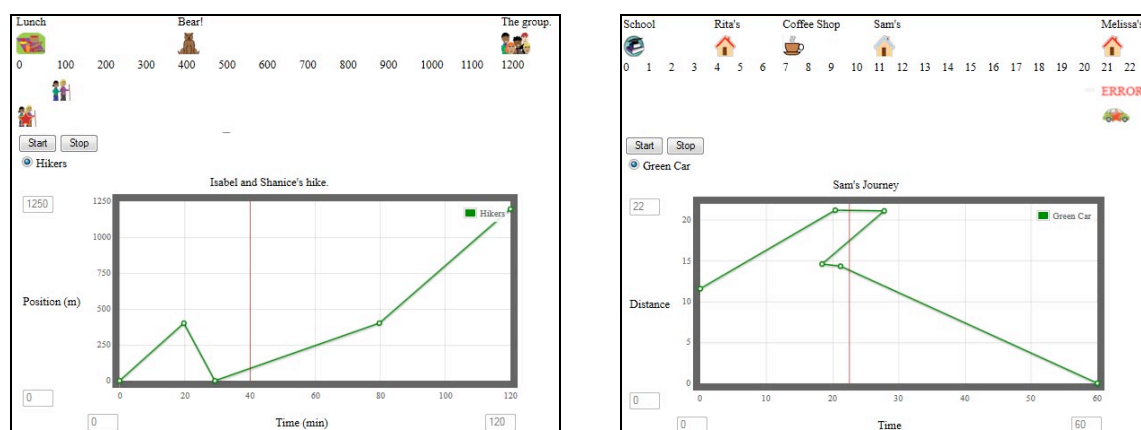


Figure 1. Screenshot of animated feedback tool (Student+Normative), *Bear* item (left) and *Sam* item (right).

For feedback on the *Bear* item, displayed in Figure 1 (left), the lower half of the image displays a student-constructed graph imported from the previous step. Above the graph, animated “hikers” move adjacent to a number line representing the range of the narrative. The lower “hikers”, marked with a red star, move according to the normative solution, while the upper “hikers” move according to the student graph. The vertical red line in the graph indicates the current time in the animation and moves from left-to-right. At the current state in the animation (40 minutes), the story indicates that the hikers should be resting at the “lunch spot”; however, this student did not include a resting segment, resulting in a visual discrepancy. For the *Sam* item (Figure 1, right), the animation depicts the motion of cars. In this case the graph incorrectly represents “going back” as a reversal in time, thereby producing a non-functional graph. In all versions of the animation, during those periods of time when the hikers or car is predicted to be at multiple positions at once, the student-based image of the hikers or car is replaced by the word “ERROR” to indicate that the graph cannot be animated appropriately.

Test Materials

The pretest and posttest included graph interpretation items in mixed multiple choice and open response formats. For this study we focus on two open response items. In “The Race”, students were asked to interpret an incomplete graph representing two runners racing towards a finish line from different starting positions (Figure 2). Students were asked to predict the winner and provide a scientific explanation for their choice. For “A Journey” students were asked to select and explain which of three graphs of distance over time – including two non-functional graphs – could represent a bicycle ride. Both items were intended to elicit students’ non-normative concepts about slope and speed. Items were coded independently by two researchers, differences were resolved by consensus (inter-rater agreement: *Race*: kappa = .88; *Journey*: kappa = .80).

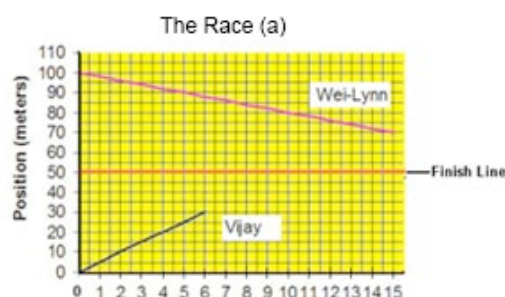


Figure 2. Graph for “The Race” item given in pretest and posttest.

Analysis Approach

All graphs generated by students for the *Bear* and *Sam* items in *Graphing Stories* were scored using an automated algorithm according to the rubric displayed in Table 2. First, student graphs were evaluated for the presence of up to four specific features corresponding to concepts addressed in the unit, including: where to place the initial point and how to represent “moving back”, “not moving”, and a change in speed. The presence of a feature was assessed either, in the case of a single point, by evaluating whether the point deviated from the

expected coordinates less than a threshold or, in the case of a slope, by evaluating whether the rise and run of the segment deviated from expected less than a threshold (in both cases approximately 5% of axis length). Second, a score was constructed to reflect the number of detected features.

Table 2: Graph scoring rubric: features and scores

Feature	Description
Initial Point	Initial point (time = 0) is placed at expected coordinates
Backwards (neg. slope)	“Moving backwards” is represented with a negative slope segment
Motionless (zero slope)	“Standing still” is represented with a zero slope segment
Speed change	“Slow” to “fast” is represented with two segments with differing slopes
Score	Description
3 - 6	Total # of features present in graph (at least once) + 2
2	If no features detected, were there at least two points?
1	Less than 2 points were detected

For both *Race* and *Journey* items a KI rubric was applied to assess the extent to which student response supported multiple normative ideas, linked in a coherent manner (Linn & Eylon, 2011). In the case of graph items normative ideas may include numerical evidence, relevant spatial characteristics of the graph (e.g. slope), formal concepts about graphs (e.g. functionality), and relevant connections to the narrative context (Table 3).

Table 3: KI scoring rubric for “The Race” pre/post open response item.

Score	Level	Description	Examples
0	No Answer		
1	Offtask	Does not address question	<i>I don't know.</i>
2	Irrelevant/Incorrect	Wei-Lynn is faster than Vijay or other incorrect statements	<i>I think Wei-Lynn will win the race because she is already ahead of Vijay.</i>
3	Partial Normative isolated ideas without a valid link	General observations of the graph OR statements about speed with no reference to the graph	<i>Vijay is running faster, he will win</i>
4	Basic Elaborate a scientifically valid link	Connects observations from graph to speed or compares lines based on distance or time	<i>I predict Vijay to win the race because he was closer to the finish line 6 seconds after the race began, meaning he has a higher rate of speed.</i>
5	Complex Elaborate two or more scientifically valid links	Connects observations from graph to Vijay's speed AND Connects velocity or slope to distance and time	<i>Vijay is traveling faster, because his line has a greater slope. His speed is 5m/s, while Wei-Lynn's speed is 2 m/s. Though Wei-Lynn's line seems longer, Vijay will reach the finish line faster.</i>

Results

An ANOVA of the pooled pre- and posttest data revealed a significant effect of testing session [$F(1, 331) = 18.6, p < .001$], demonstrating that across both conditions students made gains from pre- to posttest (Figure 3). However, neither a main effect of experimental condition [$F(1, 331) = 0.3, p > .1$], nor an interaction between condition and testing session emerged [$F(1, 331) = 1.4, p > .1$].

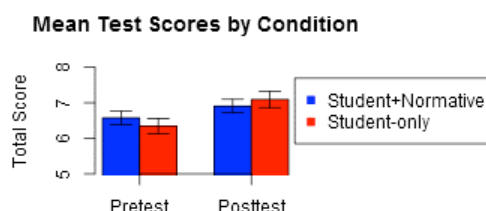


Figure 3. Mean scores for pretest and posttest, by condition

These results suggest that both conditions were equally effective in terms of promoting student graph understanding. This could either indicate that the addition of the normative animation did not affect students' graph construction strategies, or that the emergent strategies had similar effects on student learning, on average. To address this distinction we investigated initial and final graphs for *Bear* and *Sam* graphing items (Figure 4).

For the *Bear* item, while students in both conditions began with similar scores for their initial graphs, students in the S+N condition completed the step with more accurate graphs than students in the S-only condition [$t(202) = 2.6, p < .05$]. In contrast, while the final graphs of the *Sam* item did not differ by condition, students in the S+N condition produced more accurate graphs initially [$t(141) = 4.5, p < .001$].

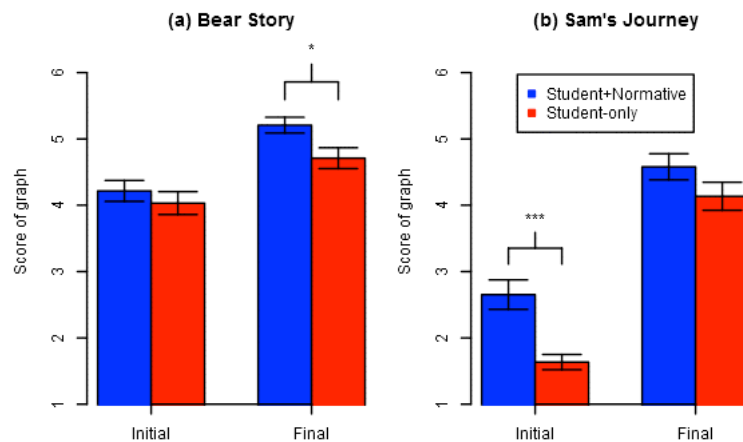


Figure 4. Comparison of initial and final graph scores for two curriculum items, by condition. (Note: for all figures and tables p-values denoted as such: *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .1$)

Across both conditions, for both items, there was a clear improvement from initial to final graphs; however, this could have been due to students taking an iterative strategy of deliberately plotting partial graphs (which would not attain high scores, initially). To determine whether the feedback tool promoted more accurate revisions, rather than simply more complete revisions, we compared the accuracy of the first graph produced with at least 4 points (sufficient to achieve max score) to the final graph. For both *Bear* and *Sam* items accuracy increased from first 4-point graph to final graph [*Bear*: $t(197) = 6.3, p < .001$; *Sam*: $t(126) = 4.9, p < .001$].

What differences in strategies may have driven the difference between conditions in accuracy of graphs? To address this question at a broad level we tallied the number of unique graphs produced for both items by counting each step visit in which the student created, deleted, or moved a point (Figure 5a). In most cases revisions of graphs were completed by revisiting the graphing step following the feedback step. For the *Bear* item students in the S+N condition made significantly more revisions than students in the S-only condition [$t(203) = 3.3, p < .01$]; however, by the final graphing item (*Sam*) an opposite trend emerged in which students in the S-only condition revised marginally more often [$t(146) = 2.0, p = .05$]. We also analyzed the mean time spent graphing each item by summing duration across all step visits (Figure 5b). For the *Bear* item, differences in conditions reflect number of unique graphs: students in the S+N spent more time graphing than students in the S-only condition [$t(203) = 4.1, p < .001$]; however, no differences emerged on the *Sam* item.

Perhaps due to the large discrepancy on time for the *Bear* item, more students in the S+N condition who began the *Bear* item did not attempt the *Sam* item than in the S-only condition [S+N: $n = 42$; S-only: $n = 19$; $X^2(1) = 8.7, p < .01$]. This attrition is reflected in the different degrees of freedom reported in previous analyses. Of the students remaining by *Sam* there were no differences between groups on pretest “The Race” scores [$t(186) = 0.05, p > .1$], suggesting that these groups were equivalent in terms of prior knowledge.

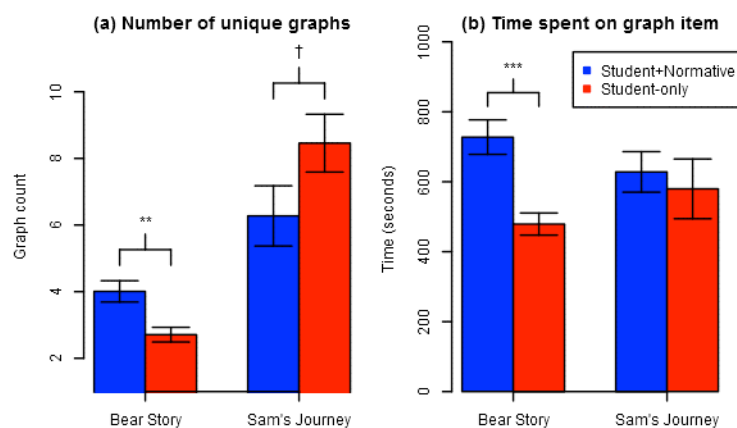


Figure 5. Comparisons of number of unique graphs (a) and time spent on graphing items (b) by condition.

Case Studies

The summary measures discussed above indicate that initially (i.e., on the *Bear* item) students in the S+N condition spent more time graphing and viewed feedback more often than students in the S-only condition. We analyze specific cases from log data to demonstrate how initial student ideas and the affordances of the graphing tools impacted graphing strategies. Cases were selected from a subset of students who showed improvement from pretest to posttest to illustrate how the graphing activity may have supported learning.

For pair A, in the Student-only condition, both students' incorrect response to "The Race" pretest item revealed a misunderstanding of the relationships between slope, direction and speed. Figure 6 displays a series of plots that this pair produced for the *Bear* item, initially (plot 1) and following feedback (plots 2-8). While the initial plot does include a segment representing "moving back" with a negative slope, this may not have been purposeful. In the 2nd through 7th attempts the pair tests multiple ideas about how to represent "moving back", including segments that progressed backwards in time (plots 3 and 5). While the final plot correctly represents "moving back", it misses a subsequent "resting" segment; however, this was congruent with their preceding tabular representation, which also missed this feature of the narrative.

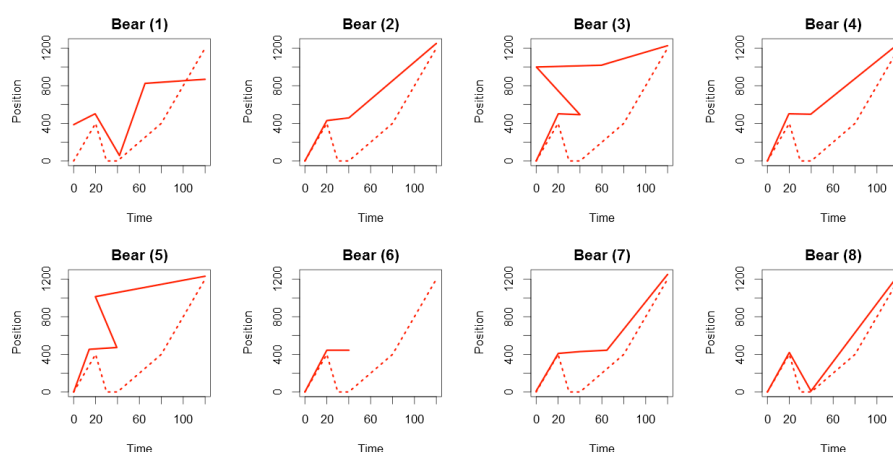


Figure 6: Pair A, *Bear* graphing item, S-only condition. Solid lines indicate student-constructed graph. Dotted lines indicate normative graph (students did not view this graph).

Likewise, Pair B (Figure 7) in the S-only condition also missed several key features of the narrative in their table: neither "moving back" nor "resting." While this pair took a successful iterative approach to represent data in their table, by missing relevant features in the table this pair lost the opportunity to test ideas about how to represent "moving back." However, in a subsequent, simpler narrative ("forward-back") these students were able to recognize their mistake and construct an appropriate graph.

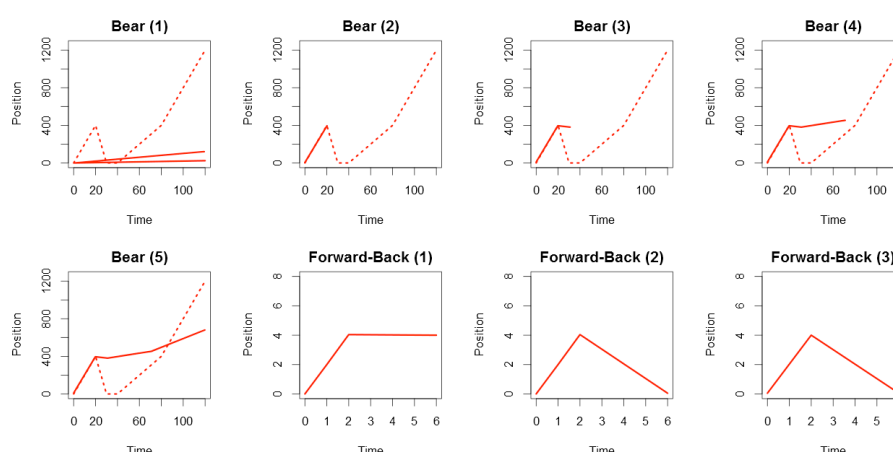


Figure 7: Pair B, *Bear* graphing item and *Forward-Back* graphing item, S-only condition.

In these two examples students were able to produce graphs that (nearly) accurately depicted the information in their tabular representations. Therefore the animation that these students viewed likely fit their interpretation of the narrative. This prevented these students from recognizing missing features of their graph. On the other hand, the S+N condition feedback was not dependent on students' interpretation of the narrative in advance of viewing feedback, thereby affording easier evaluation of graph accuracy.

For pair C (Figure 8), these students produced a graph (plot 4) with a transposed resting and backwards segments, as well as a missing final speed change. Because, in this case, the students produced a correct table prior to graphing, the errors in plot 4 likely represent a simple mistranslation of the tabular data. While the resulting animation would be discrepant from the narrative, the difference – particularly in regards to the speed change – would be subtle. Yet, with the addition of the normative animation the discrepancy is clear. In this case the pair recognized the errors and successfully revised their graph (plot 5).

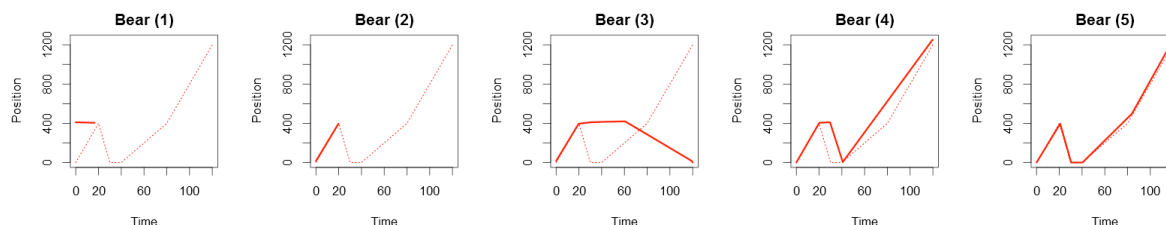


Figure 8: Pair C, *Bear* graphing item, S+N condition.

The S+N condition, however, also has the unanticipated consequence of supporting a process of trial-and-error to achieve an accurate result. For example, Figure 8 shows 8 of 15 unique graphs that pair D produced in the *Bear* item. While these students understood how to represent both “moving back” and “resting” events graphically as shown in their early graphs, they appear to use the affordances of the feedback to rapidly guess and check, without deliberate planning, to complete the speed change feature of their graphs.

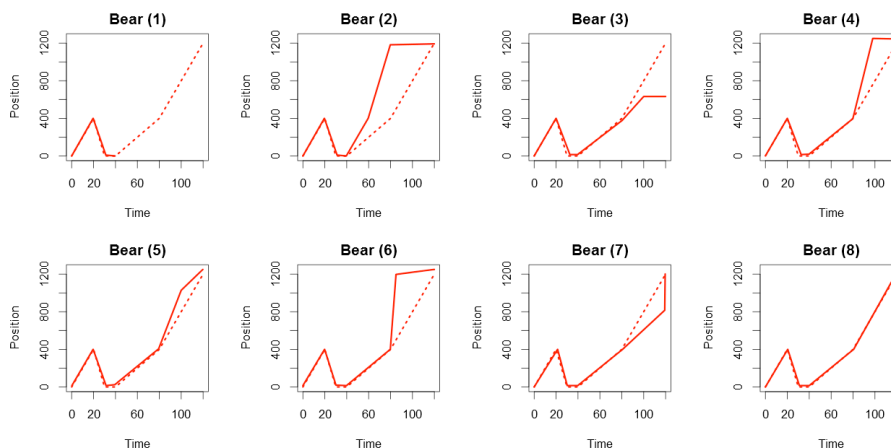


Figure 9: Sample of Pair D plots, *Bear* graphing item, S+N condition (7 plots removed).

Conclusions and Implications

While graphing instruction is often situated in an abstract mathematical context, students need to interpret quantitative data represented in graphs in multiple professional fields and everyday life. Thus, activities such as constructing graphs from a narrative provide students with an opportunity to engage in authentic professional practices. Additionally, by situating activities in personally-relevant contexts we invite students to make connections between formal concepts addressed in school and everyday experiences.

The *Graphing Stories* curriculum unit in WISE has been designed and refined with the collaboration of multiple expert teachers and researchers to help students develop ability to link narratives to graphs and graphs to narratives. The curriculum seeks to help students test and refine their ideas and to develop the ability to interpret graphs in new complex situations. Technology can help students add new ideas and also engage in the complex and difficult task of distinguishing between and reflecting on multiple ideas (Linn & Eylon, 2011).

While dynamic visualizations are common in online tools, this investigation reveals that students need extensive opportunities to test and refine their ideas with guidance that helps them sort out conflicting interpretations. When instruction promotes distinguishing between predicted and actual outcomes of a visualization students are more likely to revise their own thinking. The manner by which technology supports distinguishing is an important subject of learning sciences research. We encourage additional research that builds on these findings.

In this study the Student-only [single] condition supported evaluation of an animation in reference to the imagined sequence of the narrative. This required the student to establish a clear interpretation of the sequence. As the case studies demonstrate, when students misinterpreted important sections of the narrative, as revealed by the table, the feedback did not promote revision. On the other hand, the Student+Normative [dual]

condition alerted students of an error independently of their understanding of the narrative. In some cases this affordance led students to adopt a trial-and-error strategy that circumvented more rigorous planning.

While we suspect that the similarity in posttest gains between groups is partially due to the tradeoffs illustrated in the case studies, there are several alternative explanations that require future research. First, it may be the case that over the course of six learning activities students in both conditions achieved a personal ceiling on the material. This is demonstrated by the initial and final graphs for the *Sam* item. While students in the Student-only condition arrived at the item more prepared to construct accurate graphs initially, their final graphs were generally of equal quality as those in Student+Normative condition. Future research will explore more complex tasks.

While neither condition showed a clear advantage we suspect that some mix of dual and single animations would be beneficial. Specifically, while the normative animation could represent an initial scaffold for early items, it could be removed in later activities to help students develop new strategies. Additionally, alternative forms of feedback (e.g. text), in conjunction with the student animation, could facilitate evaluation and planning for graph revisions. More generally, this study represents an important direction for learning sciences research on the use of visualizations to advance graph understanding. While visualizations are a valuable tool for conveying ideas, the specific features and affordances require critical analysis. As the case studies presented here illustrate, emergent student strategies are often surprising and ultimately informative.

References

- Black, J. B., Segal, A., Vitale, J. M., & Fadjo, C. L. (2012). Embodied cognition and learning environment design. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of learning environments* (2nd ed., pp. 198–223). New York: Routledge.
- Cope, P., & Simmons, M. (1994). Some effects of limited feedback on performance and problem-solving strategy in a logo microworld. *Journal of Educational Psychology*, 86(3), 368–379.
- Imhof, B., Scheiter, K., Edelmann, J., & Gerjets, P. (2013). Learning about locomotion patterns: Effective use of multiple pictures and motion-indicating arrows. *Computers & Education*, 65, 45–55.
- Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, 60(1), 1–64. doi:10.2307/1170224
- Linn, M. C., & Eylon, B.-S. (2011). *Science learning and instruction: Taking advantage of technology to promote knowledge integration*. New York: Routledge.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mokros, J., & Tinker, R. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, 24(4), 369–383.
- Moreno, R. (2001). Designing for Understanding: A Learner-Centered Approach to Multimedia Learning. *Cognition*, (1999), 248–250. Retrieved from <http://www.unm.edu/~moreno/conferences.htm>
- National Governors Association Center for Best Practices, C. of C. S. S. O. (2010). *Common Core State Standards*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.
- Ploetzner, R., Lippitsch, S., Galmbacher, M., Heuer, D., & Scherrer, S. (2009). Students' difficulties in learning from dynamic visualisations and how they may be overcome. *Computers in Human Behavior*, 25, 56–65.
- Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., & Duncan, R. G. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, 13(3), 337–386.
- Roschelle, J., Kaput, J., & Stroup, W. (2000). SimCalc: Accelerating students' engagement with the mathematics of change. In M. J. Jacobson & R. B. Kozma (Eds.), *Innovations in science and mathematics education: Advanced designs for technologies of learning* (pp. 47–75). Lawrence Erlbaum Associates.
- Ryoo, K., & Linn, M. C. (2012). Automated feedback in a computer-based diagramming tool to help students distinguish among energy ideas in life science. In *Poster presented at the annual meeting of the American Education Research Association in San Francisco, CA*.
- Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, 14(1), 47–69.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257–285.
- Vitale, J. M., Black, J. B., & Swart, M. I. (2013). Applying grounded coordination challenges to concrete learning materials: A study of number line estimation. *Journal of Educational Psychology, Advance on*.
- Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, 22, 87–122.

Supporting Teacher Learning for Pedagogical Innovation through Collaborative Co-Design: Issues and Challenges

Nancy Law, Johnny Yuen and Yeung Lee, University of Hong Kong, Pokfulam, Hong Kong
Email: nlaw@hku.hk, jyuen@hku.hk, yeunglee@hku.hk

Abstract: This study reports on an in-depth analysis of two teachers' implementation of one collaboratively designed curriculum unit in the context of the school's participation in an e-learning initiative to integrate ICT use in the grade 5 General Studies curriculum. The analysis reveals significant differences in the actual implementation as well as in the students' learning outcomes achieved. While both teachers ventured beyond their comfort zone to provide more opportunities for student interactions and explorations, the enacted implementations reflect differences in the learning goals they targeted and their knowledge of practice. Visualizations of learning analytics did not trigger further exploration among teachers, though these reveal strong links between learning outcomes and the enacted curriculum. Findings raise questions about collaborative instructional design as a model for organizing teacher learning for pedagogical innovation, and the paper proposes design principles and further research to better facilitate innovation-focused professional development.

Introduction

There is much in the literature about the challenges to changing teachers' pedagogical practices, whether ICT is involved or not (Fullan, 2007; Hargreaves, 2000). The present study is concerned with promoting teacher adoption of more student-centered, inquiry-oriented pedagogies in their everyday practice, using ICT integration as a key leverage for pedagogical change. It is a design research conducted in the context of a government funded e-learning pilot project a school was awarded to carry out, and for which the research team was contracted by the school concerned to provide professional development and consultation service. A team design approach (Koehler, Mishra & Yahya, 2007) was adopted to foster the teachers' development of technological pedagogical content knowledge (TPACK, Koehler & Mishra, 2005) needed for the implementation as the teachers had been using rather traditional, teacher-centered delivery methods, and ICT use was primarily confined to teachers' use for whole class presentation purposes. In this design study, in line with the focus of the government pilot scheme to foster pedagogical transformation through e-learning implementation, the professional development focus was on learning design—the identification of learning goals and the design of appropriate teaching and learning activities that provide learning experiences conducive to student attainment of those goals (Goodyear, 2005). The knowledge and skills needed for the development of digital learning resources/online activities is not included in the professional development plan.

An important principle for design research is the need to be sensitive to the local context and to respect and nurture the agency of the teachers concerned. The teachers involved here were novices with regard to e-learning, and they did not have a clear understanding or expectation of the transformative potential of ICT. While the teachers were willing to take part in the pilot project, they were worried about their own technological competence and the workload involved. Hence the design principles adopted and the actual socio-technical design for teacher learning took account of the contextual constraints.

The focus of the present study is on identifying evidence of teacher learning throughout the design and enactment process of the technology-enabled, blended lessons in order to assess the effectiveness of the design principles for teacher learning adopted and how these could be improved. In this study, all five teachers were enacting the same curriculum unit design agreed upon after participating in a collaborative co-planning process. Given this scaffolded collaborative design process, would the teachers' implementation be largely the same, and how may they differ? What are the conditions, if any, for such teams to engage in sustained teacher learning and continuous improvement in professional practice? These are the research questions that this study addresses.

Literature Review and Conceptual Framework

Mishra and Koehler (2005) identify seven types of teacher knowledge for effective integration of ICT use in teaching and learning: content knowledge (CK), pedagogical knowledge (PK), technological knowledge (TK), pedagogical content knowledge (PCK), technological pedagogical knowledge (TPK), technological content knowledge (TCK) and technological pedagogical content knowledge (TPCK or TPACK). While there is no common agreement on the nature of TPCK as a theoretical construct, nor is there a commonly adopted set of measurement for TPCK (Voogt et al., 2012), there is general consent that the intersection and interaction between these three areas of knowledge and performance are important in the development of teacher competence to integrate ICT use in their teaching.

For teacher learning provisions to be effective, they have to be coherent and meaningful within the teachers' professional contexts. Koehler and Mishra (2005) reported significant advances in teachers' TPCK when they engage in the collaborative design of online courses. The design process provides participants with an authentic and engaging context to exercise their own creative agency, and the designed course materials become a reification (Stein & Coburn, 2008) that can mediate further learning through teachers' implementation practice. Design research (Brown 1992) has been found to be an appropriate model for supporting teacher learning as it can be designed to address three key challenges encountered in designs for teacher learning: the need to collaborate within the institutional context of the school, the need to trace and document individual as well as group level learning, and the need to provide an authentic context to link the teacher learning with their classroom practices (Cobb, Zhao & Dean, 2009). Design research is particularly suited when the teacher learning for technology integration aims to foster pedagogical change in the context of school reforms (Holland 2001). Penuel et al (2007) further found that the incorporation of lesson planning time and the provision of technical support were important for promoting implementation.

In the present study, the university-based research team was invited by the school to provide professional development support to help teachers integrate ICT use in their teaching of the school curriculum, and to conduct formative evaluation of the school's e-learning pilot implementation. Design-based research was adopted. Each cycle of learning is centered around the design and implementation of an e-learning curriculum unit. e-Learning in this context refers to the integration of online student learning activities to be completed individually or in groups within or outside of the timetabled class teaching time. To facilitate this mode of e-learning for novice teachers and learners, the researchers provide a Moodle-based learning management system (LMS) customized for young learners for the school to implement their blended learning curriculum units.

In designing the present study, the research team is guided by the following design principles:

1. The ultimate decisions on the design of the e-learning curriculum unit to be implemented rest with the school teachers to ensure its intelligibility, feasibility and acceptability to the culture and norms of the school (Fishman et al., 2000).
2. The teacher learning interactions should focus on the conceptual aspects of learning design (i.e., the selection and deployment of learning activities and associated online resources and tools for the targeted learning objectives) and skills in classroom implementation. To reduce the cognitive stress and workload for teachers, the installation of the designed learning activities and digital resources on the LMS are to be carried out by the research team.
3. Accessibility of technical support—hotline and email support for LMS operations.
4. Analyses of students' learning processes and outcomes will be used as a main form of feedback to scaffold teacher reflection on practice and further learning.

The Design Study

This design-based research was conducted in the form of a school-university partnership initiated by a primary school requesting school-based professional development and technology support for the implementation of the school's e-learning pilot project. The school was located in an area of relatively low SES and did not have a strong record of e-learning implementation. The predominant pedagogy was teacher-centered instruction, but both the principal and the teachers were keen to take advantage of the funding support provided to pilot schools to make advances in the pedagogical adoption of ICT. Each pilot project funding application had to identify the pedagogical focus, targeted outcomes and an implementation plan. This school identified the fostering of students' information literacy skills as recommended by the Education Bureau in Hong Kong (EDB 2005).

Grade 5 General Studies (GS) was selected by the school for e-learning implementation in the 2012-13 school year. All of the five in-service teachers teaching grade 5 GS in this school participated in the study. The collaborative design took an iterative process, led by the pilot project coordinating teacher, who was also one of the five grade 5 GS teachers. The teacher team selected the curriculum unit on *Fostering a Positive Attitude Towards Life* for e-learning implementation in the 2012-13 school year. Figure 1 is a graphic representation of the design study process undertaken, based on the design principles described above.

Based on design principle 1, the research team invited the teachers to provide an initial e-learning curriculum unit design they would like to implement based on their prior curriculum design. Three cycles of interaction took place between the coordinating teacher and the university team. During the first iteration, the researcher input focused on converting the essentially teacher directed curriculum plan to include inquiry-oriented individual and group-based student activities to foster student engagement, after which the coordinating teacher provided a revised unit plan. A second iteration of coordinating teacher-university researcher interaction took place that focused on providing ideas for further e-learning refinement. The coordinating teacher then submitted a finalized e-learning unit plan to the research team, who then implemented the e-learning component of the curriculum unit design into five separate but identical courserooms on the LMS, one for each of the five classes. A face-to-face team planning meeting involving all five GS teachers and the research team member then took place, during which the courseroom features and activities were introduced. The courseroom activities

were further refined after the meeting for actual deployment, based on the teachers' concerns and suggestions. The curriculum plan involved two double periods (each class period was 35 min) in the computer laboratory, and an optional single period for extension activities to be decided by the class teacher concerned.

The university-based researchers observed all the five teachers' classroom enactment of the final curriculum plan. The courseroom activities for all five classes were identical, which also served as a reification to mediate the implementation of the agreed curriculum plan. Videos were taken during the lesson observations. A student focus group interview and a teacher interview were conducted for each class right after the second double period. A short debriefing meeting was also held with all five teachers shortly after the completion of the curriculum unit. Three months after the completion of the teaching, a meeting was scheduled between a research team member and the teachers to review the graphic visualization of analyses of data from the LMS server that provide information about students' learning processes and outcomes for the five classes.

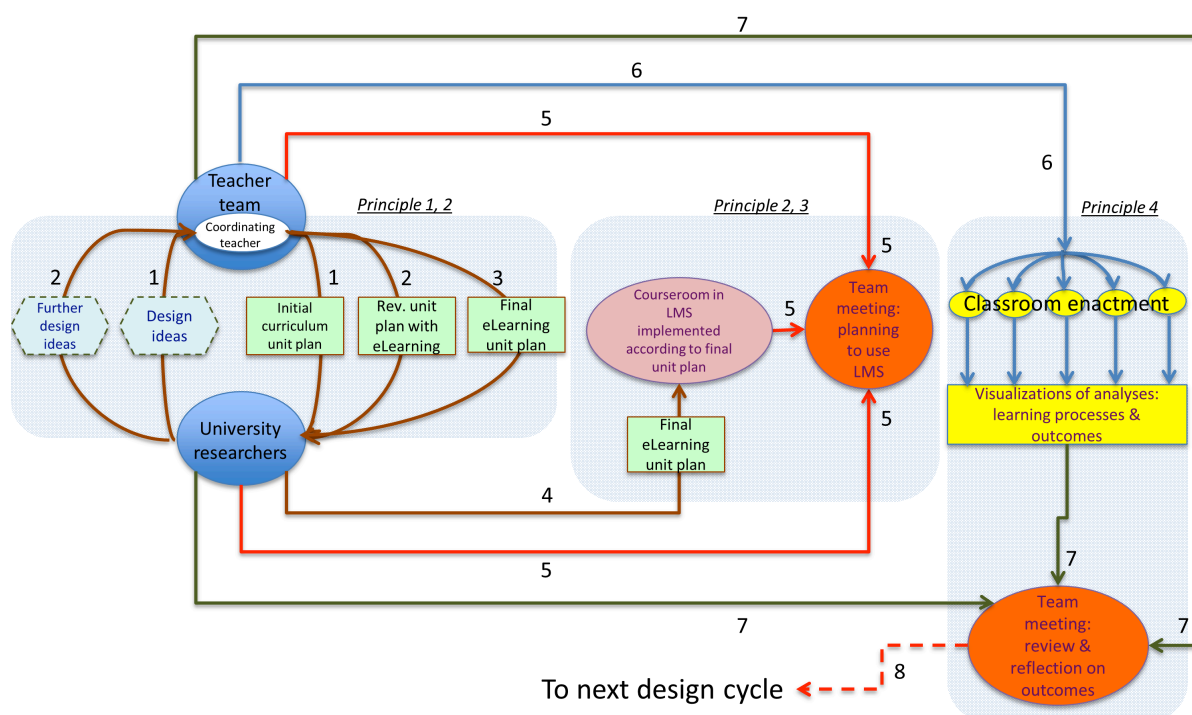


Figure 1. A graphical representation of the design research cycle implemented by the research team.

Design of the Curriculum Unit and Data Sources

The curriculum unit submitted by the teachers included a listing of the specific learning objectives for the unit, which remained unchanged throughout the study: I. set personal goals and develop an executable plan to achieve the goals, II. adopt a positive attitude towards stress and setbacks, and III. help others to adopt a positive attitude towards life. The basic learning design for the finalized curriculum unit was to use two contrasting cases, a hikikomori youth (1) and a “reborn hero”(2), to stimulate students' discussion and reflection, in order to foster in students a positive attitude towards difficulties and challenges in life. Table 1 presents the designed learning activities for the first of two 70 min. lessons in this curriculum unit. The plan clearly shows that the planned use of ICT was to support student sharing and exploration of ideas. Four types of digital tools are involved: a YouTube video for viewing by the whole class, an online discussion forum for students to post and share ideas, an online voting tool that provides an immediate summary of the class responses and a wiki for each group to review and consolidation what they have learnt. At the end of the second lesson, students were asked to share their thoughts and reflection on their learning experience using a blog in the LMS used in this course.

Analysis and Findings

In this section, we report on teachers' learning as reflection during each of the three phases of the design study process: the curriculum design phase, the enactment phase and review of learning analytics visualization of students' learning processes and outcomes.

Teacher Learning During the Design Phase

The design phase includes steps 1, 2, 3 and 5 of the design cycle in Figure 1. During steps 1 to 3, the research team interacts with the teachers through the coordinating teacher. Hence it is not possible to differentiate

learning at the individual level from the group level. The initial curriculum unit was based on the teaching plan used in the previous school year, which was strongly influenced by the textbook. The design focus was on including computer-based activities to provide opportunities to learn information literacy skills such as defining, accessing, managing, integrating, evaluating and communicating information. There were three major learning activities in the initial curriculum unit that focused on getting the students to develop their own personal goal: learning about the life cycle of the butterfly as an introduction to changes in one's life journey; identify a personal goal and to developing a personal plan for its achievement; and to develop a better understanding of various stages in one's life journey through studying the life events of a reborn hero. The e-learning activities included were mainly searching, selecting and presenting information connected with these tasks. The researcher input encouraged the teachers to focus on the social affective aspects of the targeted learning objectives and to include more sharing and group interactions among students. An example of a unit with a similar set curriculum objectives designed and implemented by teachers in a local secondary school was introduced to the coordinating teacher for reference. That example unit used two contrasting cases as a pedagogical anchor, a hikikomori youth and a reborn hero, and details of the design were publicly available from a webpage.

Table 1. Sequence of learning activities for the two 70 min lessons based on the teaching plan finalized by the coordinating teacher after the co-planning team meeting.

Sequence	Activity type	Description
Lesson 1		
1.1	Investigate	Students asked to investigate unhappiness by sharing with the whole class their unhappy moments, and discuss how they handle unhappiness.
1.2	Watch	Students watch a TV documentary (YouTube) on a hikikomori youth, Ar Yuen.
1.3	Discuss—online forum	Students share their feelings about the video, discuss why Ar Yuen isolated himself, whether and how Ar Yuen should connect with the outside world.
1.4	Investigate	Students vote online on the problem that troubles them most from a list: academic results, relations at school, relations with family, friends, money matters. Teacher review voting results in class, invite students to elaborate on the problems they experienced, and whether they had tried to isolate themselves under those circumstances.
1.5	Produce—online	Students work in groups of three on a wiki to describe: what problems trouble them most and how they solve those problems.
Lesson 2		
2.1	Discuss-whole class	Students asked to summarize the problems they encountered as reflected in group wiki entries.
2.2	Watch	Students watch a YouTube video on one “reborn” hero.
2.3	Discuss—online forum	Students to post their feelings about the reborn hero in the video.
2.4	Investigate & produce online	Students work in groups to search for further examples of reborn heroes online and select one to write about on their group wiki and reasons for their choice of hero.
2.5	Online peer assessment	Students peer assess other groups based on an online rubric and vote for the best group wiki.
Lesson 3—optional		
3	Reflection blog	Students post reflections on their learning in the unit using the LMS blog module.

The revised design submitted by the coordinating teacher was very close to the example design. The teachers decided to drop the activities related to the butterfly life cycle and the development of a personal goal and plan, and adopt the use of two contrasting cases as the design anchor. In the second round of interactions between the research and the teacher teams, the focus was on clarifying the intended function of the e-learning activities included and to reduce the number of different online activities within the same lesson to ensure that the plan was feasible. Hence, the finalized plan did not differ much from the plan previously submitted.

Step 5 in Figure 1 was the final step in the design and planning process. During this meeting the teachers were shown a sandbox courseroom on the LMS with the planned online activities set up for them to explore. During this meeting, the main concerns expressed by the teachers were how to navigate and manage the different activities, and whether the students would be able to handle the various online activities according to the plan. There was no discussion on the appropriateness of the overall pedagogical design. The teachers spent much time on mastering the technical skills of navigating and using the LMS, which was totally new to them. Towards the end of the meeting, several teachers expressed their lack of confidence in helping students to use

the e-portfolio system provided, Mahara, for their students to create their group report on a reborn hero in the second lesson. The research team then eliminated the need for students to work on two platforms (though linked through single sign-on), and created a group wiki in the LMS course room for the task.

To summarize, the teachers exhibited openness to trying out new pedagogical designs and technology tools, while still harboring much anxiety in the process. This is an important first step in teacher learning.

Teacher Learning During the Enactment Phase

As the curriculum design required all students to have access to an internet-connected computer during lessons, it was necessary to schedule the classes at times when the computer labs were available. In the end, the teachers decided that the best time was to conduct the classes during the post-examination period when teachers were able to flexibly schedule different activities. They booked both computer labs available in the school for two whole days and schedule the two double period lessons on the two days. This was very convenient for the research team as all lesson observation and interviews could be completed in two days. On the other hand, the teachers did not attempt to engage in peer lesson observation because of the tight time schedule, which in hindsight was an obstacle to more productive professional learning (to be further discussed in a later section).

All five teachers implemented the online activities according to the plan and were positive about the experience and outcomes in the post-observation interviews. The debriefing meetings with the teachers, whether individual or in groups, were rich and productive. While the implementation details and competences of the teachers differed, all were pleased with the ability of the students to complete the designed activities. Much of their initial fears and worries were allayed and most of the teachers found students' levels of engagement to be beyond their expectation. The most encouraging aspect of the teachers' conversation was its focus on matters that span both technology and pedagogy. For example:

My concept of wiki was different from what I experience in this unit.... It is a learning approach.... Watching the YouTube video and working on wiki iteratively is very good.

I could not imagine my students' response to be like this. I want to suggest .. modification .. For example, voting for the best [group] wiki [on reborn hero]. We should ask them to vote not only for the best group, but to give a score for each group. When asked whether we should ask students to do peer or self assessment earlier, I thought for a long time I have not seen peer- or self-assessment and I worried that the children would not be able to handle....

It is difficult to get low achievers to collaborate ... I find that their engagement in this activity to be much better than before. ... online voting is a kind of guidance, encouraging them to do better. For example, one of the groups in the weak class was not putting in their best After the voting, the group was surprised that none in the class voted for them and asked why.

The teachers' talk exhibits a fluid and fluent integration of pedagogical ideas and concerns with the functionality and use of different online tools. The observation of their own students' engagement in the e-learning activities served as an important experiential grounding that connects/bridges the teachers' prior pedagogical skills and understanding with new technological tools and new pedagogical methods. The bridging was so successful in helping the teachers to make sense of the new tools and activities that they were able to suggest improvements to their enacted designs that would go beyond what they were initially prepared to "risk". Clearly, these teachers gained much in TPCK in their learning through design and enactment of those designs. What is also very encouraging was that the teachers were also aware of this venture as a productive professional learning experience, as so well articulated by one of the teachers in the post lesson interview:

I did not plan for the children to come out and talk, but since there was time available ... the [LMS] platform provide flexibility for teaching. You can think as you go along. You can never plan perfectly in advance ... we need to accumulate experience. So I think this time we are not really teaching, but learning. We are learning how to use this [LMS] platform

It is clear from the above quote that the fear of not being able to navigate and manage activities on the new platform is replaced by confidence and sense of control, as communicated through the term "flexibility".

Variations in Teachers' Enactment of the E-Learning Plan

While all five teachers implemented the learning activities as planned, a detailed analysis of the classroom videos revealed subtle and important differences between them, which could be linked to differences in patterns of students' engagement and learning outcomes achievement as revealed through analyses of the online log data. Due to space limitations, two of the teachers' classroom interaction and orchestration are selected for

examination in this section as they represent different foci in their enactment. A brief summary of the activity sequence for the first double period lesson for Teachers A and C is presented in Table 2. Neither of these two teachers was the coordinating teacher.

Table 2 A summary of two teachers' implementation of the first lesson in the co-planned curriculum unit.

Sequence	Teacher A	Teacher C
1.1 Investigate	Did not perform this activity.	<ul style="list-style-type: none"> Recapped a similar topic from last year Asked for <u>the meaning of "hidden" as in hikikomori</u> Ask for an adjective to describe Ar Yuen
1.2 Watch	Showed video.	Showed video.
1.3 Discuss	Ask whole class how the video described Ar Yuen, the meaning & attitudes of hikikomori. Strict regimented instruction to students: <ul style="list-style-type: none"> 10 min to post own ideas 3 min to <i>read others' posts</i> 5 min to <i>discuss in group about notes read</i> 8 min to <i>post group discussion results and respond to others</i> 	<ul style="list-style-type: none"> Checked students can login Asked students to post responses to above questions on forum Gave feedback on students' postings & refer to hikikomori <u>as a social phenomenon</u> Guided students to post causes of hikikomori in terms of <u>individual, family and social factors</u>
1.4 Investigate	<ul style="list-style-type: none"> Administered online voting on key difficulty <i>Randomly chose 2 students to report how they voted</i> 	<ul style="list-style-type: none"> Administered online voting on key difficulty Reviewed results and explained why academic results had highest votes Contrasted Ar Yuen's attitude with students'
1.5 Produce	<ul style="list-style-type: none"> Asked students how they solve own problems 10 min for <i>group to discuss f-t-f their problems & solution</i> Groups post results in wiki Read some students' wiki entries to conclude 	<ul style="list-style-type: none"> Asked each group to discuss solutions to most concerned problem & post on wiki Read out from wiki some of the students' solutions

N.B. Italicized texts indicate careful scaffolding of student interactions structures (individual, group and whole class) using face-to-face and online means. Underlined texts indicate conceptual scaffolds for student discussion and reflection.

Teacher A followed the plan very closely in the form of instructing students on the activities they need to go through. He did not attempt to provide a conceptual focus or guide the discussion. Teacher A followed the plan very closely in the form of instructing students on the activities they need to go through. He did not attempt to provide a conceptual focus or guide the discussion. On the other hand, he paid much attention to structuring of the discussions. In sequence 1.3, he gave time for the students to read others' postings, followed by a face-to-face group discussion, and then posting the group discussion outcomes as a response to the forum postings. Here the teacher modeled how online discussions is meant to be an interaction of ideas, and encouraged students to take each others' postings seriously. The italicized text in Table 2 highlights the instances Teacher A provided social-technical structuring of the discussions, which demonstrated his TPCK fluency in relation to facilitating student discussions.

Teacher C demonstrated a strong conceptual focus on her learning design enactment. She was the only teacher who started the unit with recapping for the students a similar topic in last year's curriculum, and to guide students to discuss the causes of hikikomori from individual, family and social perspectives; both of these moves were not mentioned in the curriculum plan. She also ensured that all students were able to login to the online learning platform before she gave instructions for the first discussion forum. The instances where she provided a conceptual scaffold for the students are underlined in Table 2.

Variations in Students' Learning Outcomes

Some preliminary analysis of the students' online discussion and reflection blog data reveal interesting differences in the learning outcomes for the different classes. Figure 2 presents an analysis of the quality of the forum postings from the different classes using an argumentation-based framework.

The quality of forum postings are often analyzed based on the quality of argument presented. As the children involved in this study were only grade 5 children and were not used to online discussion as a learning activity, explanations were not further categorized. The analysis presented in Figure 2 shows that class C has the highest number and percentage of high quality postings (i.e. explanation, personal judgement and suggestions). It is not clear whether this outcome is a consequence of the cognitive scaffolding provided by the teacher as Class C also happens to be the high achievers at this grade level. Another measure of quality of forum

discussions is the level interactivity among participants. Figure 3 is the Social Network Analysis (SNA) display for the two classes. The class C SNA result is a typical star shape indicative of students' postings serving primarily as responses to the teacher's posting. There was not real discussion or exploration of ideas among students. On the other hand, the class A SNA display is a well connected network showing much interaction among the students.

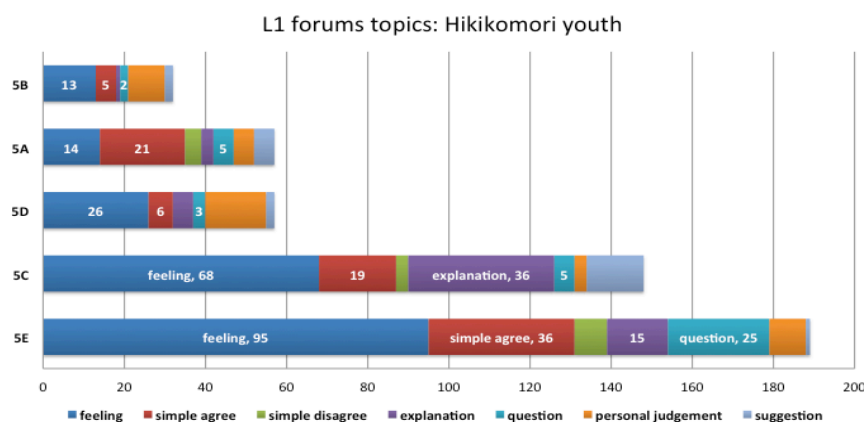


Figure 2. Analysis of the nature of the forum posts contributed by the five classes of students in lesson 1.

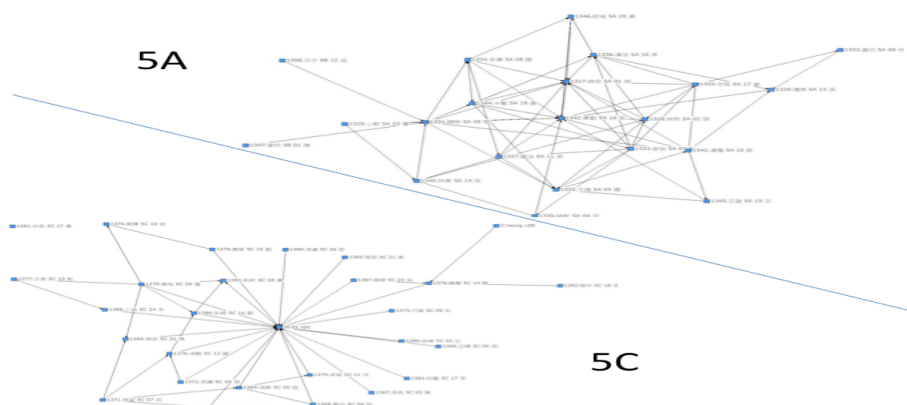


Figure 3. Social Network Analysis results for the discussion forum writing interactions for classes A and C.

A further analysis was conducted to categorize the level of reflection of the students' end of unit blog posts: 1) a simple listing of ideas learnt, 2) include positive suggestions on handling difficult situations, and 3) connect personal experience in the reflection. Contrary to results presented in Figure 2, class A students' levels of reflection in their posting (1:13:6) was much higher than those from class B students (10:5:1).

Variations in Students' Learning Outcomes

Three months after the curriculum unit enactment, the research team had an opportunity to present the above findings to the teachers, hoping that they will find these outcomes intriguing and trigger another conversation that allows for further probing of the unit design and its enactment. Disappointingly, the teachers did not show much interest and the graphs were taken as rather academic artefacts. On reflection, the research team identified several reasons why this may be so: 1) the analyses shown were rather unfamiliar to the teachers and do not really carry meaning in terms of learning outcomes as the teachers understand them, even though they are familiar to the learning sciences research community; 2) lack of recency to effectively trigger discussion due to the time lag between curriculum enactment and the presentation of the analyses outputs; and 3) the teachers did not have an opportunity to observe each others' lessons and were hence unaware of the differences in enactment. All three conditions limited greatly the opportunity for this design study to contribute to deeper teacher learning.

Conclusion

Results from this study indicate that collaborative design based on the design principles presented under the conceptual framework section is a non-threatening approach to support teacher learning for e-learning implementation. The collaborative design that emphasizes teacher agency and teachers as the ultimate decision-maker provides a safe environment for teacher learning. The actual enactment of the learning design and post-enactment reflection is crucial to effective learning. The present study demonstrates that teachers were able to

master the interactive aspects of TPACK through practice. On the other hand, the findings also reveal the inadequacies of the design principles in fostering progressive inquiry and learning through reflecting on the variations in teacher enactment and associated student outcomes. Curriculum enactment reveals an extra, more nuanced level of teacher knowledge that cannot be adequately reflected in reified design documents or curriculum artifacts, be they digital or otherwise. To leverage this extra level of teacher learning and becoming in and through practice requires three additional design principles: 1) promote assessment and learning analytic literacy so that teachers can interpret and make use of the tools and indicators developed in the learning sciences community for understanding and improving their students' learning; 2) artifacts and opportunities to support reflection on practice should be provided as soon as possible after enactment to ensure recency and impact; and 3) provide peer lesson observation opportunities to teachers to support a teacher learning discourse on knowledge in and through practice.

While design principles clearly matter, it is important to note that there are school and system level conditions that are necessary for these principles to be feasibly implemented. For example, peer observation of teaching would only be a rarity unless explicitly supported by school level policy. Assessment and data analytic literacy and the availability of automatic data analytic and visualization tools need system level investment and policy support to be realized.

Endnotes

- (1) A Japanese term referring to **reclusive** adolescents or young adults who withdraw from social life, often seeking extreme degrees of **isolation** and confinement, from <http://en.wikipedia.org/wiki/Hikikomori>.
- (2) A term that refers to those who is determined not to succumb and continue to lead a productive life even in the face of great atrocities. In this curriculum unit, the “reborn hero” is a radio broadcaster (DJ) who returned to his former job after undergoing surgery as treatment for tongue cancer.

References

- 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.
- Cobb, P., Zhao, Q., & Dean, C. (2009). Conducting design experiments to support teachers' learning: A reflection from the field. *the journal of the learning sciences*, 18(2), 165-199.
- Education Bureau (EDB) (2005). *Information Literacy Framework for Hong Kong: Building the Capacity of Learning in the Information Age*. Hong Kong: Government printings.
- Fishman, B., Best, S., Foster, J., & Marx, R. (2000). *Fostering teacher learning in systemic reform: A design proposal for developing professional development*. ERIC Clearinghouse.
- Fullan, M. (2007). *The new meaning of educational change*. New York: Teacher's College Press.
- Goodyear, P. (2005). Educational design and networked learning: Patterns, pattern languages and design practice. *Australasian Journal of Educational Technology*, 21(1), 82-101.
- Hargreaves, D. H. (2000). Schools and the Future: the Key Role of Innovation. In Centre for Educational Research and Innovation (CERI) (Ed.), *Innovating Schools* (pp. 45-57). Paris: OECD.
- Holland, P. E. (2001). Professional development in technology: Catalyst for school reform. *Journal of Technology and Teacher Education*, 9(2), 245-267.
- Koehler, M. J., & Mishra, P. (2005). What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *Journal of Educational Computing Research*, 32(2), 131-152. doi: 10.2190/0ew7-01wb-bkhl-qdyv
- Koehler, M. J., & Mishra, P. (2005). What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *Journal of educational computing research*, 32(2), 131-152.
- Koehler, M. J., Mishra, P., & Yahya, K. (2007). Tracing the development of teacher knowledge in a design seminar: Integrating content, pedagogy and technology. *Computers & Education*, 49(3), 740-762.
- Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, 44(4), 921-958.
- Stein, M. K., & Coburn, C. E. (2008). Architectures for learning: A comparative analysis of two urban school districts. *American Journal of Education*, 114(4), 583-626.
- Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J., & van Braak, J. (2013). Technological pedagogical content knowledge, a review of the literature. *Journal of Computer Assisted Learning*, 29, 109-121.

Acknowledgment

This research was conducted on a set of data from the research “*Pedagogical Design for Effective e-Learning*” commissioned by the Hong Kong Applied Science and Technology Research Institute (www.astris.org).

Inexplicable Silence: An Uncomfortable Analysis of the Social Silences

Daniel Steinbock, Keio University, KMD Office, 4-1-1 Hiyoshi, Kohoku-ku 223-8526, Japan
daniel@steinbock.org

Abstract: *Social silence* refers to silences that arise in face-to-face social interaction. Despite having a reputation for being nothing at all, close examination of social silence reveals various *kinds* of silence, arising under different conditions, playing a variety of roles in the temporal order of social interaction. What distinguishes one silence from another? For instance, why are some moments of silence experienced as uncomfortable and others as unremarkable? Here I introduce a conceptual framework for describing a variety of social silences, built on the work of conversation analysis and an ethnographic study of silence in the religious practices of Quakers. A focus on social silence naturally makes the unspoken dimensions of interaction salient and offers new confirmation for the view that face-to-face learning and communication are always partly constituted by embodied (non-verbal) actions. The included framework, descriptions and empirical data serve as platform for two theses: 1) calibrating to the acoustic environment is a basic way that individual behavior is situated in its social context; 2) group silence regulates collective attention by directing it to the present moment of social interaction.

Keywords: social silence, social interaction, embodied practice, conversation analysis, Quakers

Introduction

Social silence refers to silences that arise in face-to-face social interaction. In the West and in the social sciences generally, silence is commonly perceived to be a particular kind of nothing, empty of words and structure, offering no material for attention or analysis. On the other hand, certain kinds of silence in social interaction are routinely cause for discomfort, embarrassment, and other charged emotions — hardly what we'd expect of 'nothing' — demonstrating that social silence is, in fact, *something* and worthy of our attention. This paper offers a way of distinguishing one silence from another and thus can answer why, for instance, some moments of silence are experienced as uncomfortable and others as unremarkable.

My aim is to describe a few varieties of social silence, with attention to the different situating conditions in which they arise, take form, and end. I have found this micro-analysis of situated silence helpful for understanding the unspoken side of a range of social behavior, including but not limited to public speech, cocktail party conversation, awkward silence, and mystical experience (Steinbock, 2012). An investigation of social silence offers balance against the overwhelming bias towards linguistic phenomena — talk and text — in the social sciences. As ethnographers of social interaction and speaking have repeatedly asserted, face-to-face learning and communication are always partly constituted by embodied (non-verbal) practices (Birdwhistell, 1970; McDermott, et al., 1978; Goodwin, 2011). Social silence naturally makes these unspoken dimensions salient, without artificially separating them from the total communicative system (Tannen & Saville-Troike, 1985).

In these few pages, I offer both the wide view — a way of distinguishing and ordering kinds of social silence — and a close look at one culturally-situated practice of silence. First, I introduce a conceptual framework for categorizing different kinds of silence in social interaction, built on the work of conversation analysis. I then share results from a four year ethnographic study of group silence I conducted in a Quaker community. Quakers have practiced deliberate group silence for religious purposes for over three and half centuries. In addition to participant-observation, I recorded panoramic video of Quakers' embodied practice of group silence *in situ*. These materials provide a basis for discussing the significance of social silence within and beyond the Quaker context.

Social Silence

Like all social phenomena, each silential situation is unique in its particulars. But are there, perhaps, recurrent forms that arise under comparably similar conditions? Here I'll describe three distinct categories of social silence, each having variations that are 'comfortable' (situationally-appropriate) or 'uncomfortable' (situationally-inappropriate): 1) intra-speaker silence, i.e. pauses in individual speech, 2) inter-speaker silence, i.e. gaps between speaker turns in conversation, and (3) group silence. Group silence refers to those seemingly inexplicable silences, familiar to any casual observer of social life, that occasionally arise in rooms full — now

suddenly empty — of conversation. Group silence is the least understood of the three forms and thus will earn the most attention in this paper.

This three-fold framework distinguishes silences by the person(s) responsible for producing them. What does it mean to assign 'responsibility' for a silence? Unlike talk, spilt milk, and left hooks, silence seems to be no one's doing in particular. Although commonsense, this view misconceives silence as meaningless emptiness. In response, let us examine a number of silences in their social context and take careful account of who is responsible.

Individual Pause (Intra-speaker Silence)

Individual silence refers to breaks in speech while a single speaker 'has the floor.' Yet in what sense is individual silence *social*? Mark Twain, expounding on the use of silence in public speaking, illustrates how an individual is not only socially responsible for his own silences, but capable of deploying them with great communicative precision:

...the *pause* — that impressive silence, that eloquent silence, that geometrically progressive silence which often achieves a desired effect where no combination of words howsoever felicitous could accomplish it. The pause is not of much use to the man who is reading from a book because he cannot know what the exact length of it ought to be; he is not the one to determine the measurement — the audience must do that for him. He must perceive by their faces when the pause has reached the proper length... For one audience the pause will be short, for another a little longer, for another a shade longer still; the performer must vary the length of the pause to suit the shades of difference between audiences. (Twain & De Voto, 1940, p. 226)

Twain observes the subtle social contingencies of delivering a single break in public speech, noting three elements that will be recurrent themes in this paper: 1) mutual attendance between speaker and listeners, 2) attention directed to expressive bodies (faces, in this case), 3) the normative claim that a silence should fit its social context. Twain's main point is that the length of a pause is gauged by keen social awareness of the audience's developing reaction. Too short and it won't have its desired 'impressive' effect, being indistinguishable from an ordinary break between utterances. An over-long silence, on the other hand, may have *too* impressive an effect — an experience terrifyingly familiar to any public speaker who has momentarily forgotten her speech. In other words, when further talk is expected but none is offered, the resulting silence may take on an uncomfortable hue. Indeed, the impressive effect of a skilled speaker's dramatic pause may be understood as playing deliberately close to the edge of uncomfortable silence — like a tight rope walker flirting with danger while we all watch, thrilled and anxious. One can imagine that this mixed emotion is precisely what Twain's orator looks for in the faces of his audience.

As for *who is responsible*, in the case of public speech-giving, the orator is the only possible candidate: for any break in speech, he is locked into the position of next-speaker (Sacks & Jefferson, 1995, v.2:521). Generalizing this principle to ordinary conversation, if the next-speaker is determinate (e.g. when someone has just been asked a question), that person is responsible for any silence that ensues (Schegloff, 2007, p.19-20). So long as an individual remains in the position of next-speaker, any silences are her responsibility. The situation is different when there is more than one candidate for next-speaker and we get *shared* responsibility for silence — inter-speaker silence.

Lapses in Conversation (Inter-speaker Silence)

The field of conversation analysis, closely associated with the work of Harvey Sacks and Erving Goffman, pioneered the identification of 'rules' that implicitly govern the formal organization of conversational turn-taking and interaction rituals (Sacks, Schegloff & Jefferson, 1974; Goffman, 1967). According to Sacks and his collaborators, conversation is, properly speaking, no more than one person talking and no less than one person talking. Now we all know that everyday conversation is full of lulls, interruptions, and overlaps. What Sacks' definition does is establish the *unremarkable* norm against which gaps and overlaps can be seen as *accountable deviations*.

Talk can be continuous or discontinuous. It is continuous when, for a sequence of transition-relevance places, it continues (by another speaker, or by the same continuing) across a transition-relevance place, with a *minimization of gap and overlap*. Discontinuities occur when, at some transition-relevance place, a current speaker has stopped, no speaker starts (or continues), and the ensuing space of non-talk constitutes itself as more than a gap — not a gap, but a *lapse*... (Sacks, Schegloff & Jefferson, 1974) [emphasis mine]

As Sacks and colleagues suggest, accountable silences enter conversation as lapses, for they go against the expected norm of continuous talk — expected, at least, by members of Anglo-American culture. When an interstitial gap is over-long and no speaker takes the floor, it is transformed from an unnoticed structural marker into an acutely accountable moment lodged in the foreground of awareness — *uncomfortable silence* (McLaughlin & Cody, 1982). Case in point: the silence that falls between two people who have just met at a cocktail party (that canonical laboratory for social interaction and self-presentation). If a moment should arise when newly acquainted persons run out of supplies for conversational material and a shared silence emerges, it may be construed as an embarrassed or awkward silence, the embarrassment springing from conversationalists' mutually visible inability to keep to the expected 'rule' of continuous talk.

Undue lulls come to be potential signs of having nothing in common, or of being insufficiently self-possessed to create something to say, and hence must be avoided. (Goffman, 1967, p. 36)

Once individuals enter a conversation they are obliged to continue it until they have the kind of basis for withdrawing that will neutralize the potentially offensive implications of taking leave of others. While engaged in the interaction it will be necessary for them to have subjects at hand to talk about that fit the occasion and yet provide content enough to keep the talk going; in other words, safe supplies are needed. What we call "small talk" serves this purpose. When individuals use up their small talk, they find themselves officially lodged in a state of talk but with nothing to talk about; interaction-consciousness experienced as a "painful silence" is the typical consequence. (Goffman, 1967, p. 120)

Given its direct conflict with the implicit 'rules' of conversational turn-taking, it is no wonder that silence has come to have uncomfortable connotations. It evidences failure in communications — a breakdown in the otherwise continuous flow of talk. Discontinuities are often hastily masked or repaired by noise-making or meaningless filler talk: coughing, "ok, so...", "anyway..." (McLaughlin & Cody, 1982; Newman, 1982). Perhaps, like the author, one may take the uncomfortable quiet itself as a conversation topic to banish the silence.

Group Silence

Group silence is the least understood of the three silences discussed in this paper, partly because it transcends the boundaries of a single conversational turn-taking system, and so has not been a unit of analysis for conversation analysts. Yet any casual observer of social life has noticed these seemingly inexplicable quietings in a room full of parallel conversations, such as at a cocktail party. Suddenly the room is quieter than it was a moment earlier. Speakers' voices drop in volume or cease altogether. For a fleeting moment, the social boundaries between conversational huddles weaken or dissolve, and those present are drawn into one of the most transient of social organizations: group silence.

Who is responsible for such a silence? Unlike the previous silences, where responsibility lies with the candidate(s) for next-speaker, group silence appears to be an emergent social phenomenon, irreducible to particular persons who are or are not speaking. Also, whereas the previous silences are discrete — present or not present — group silence has variable magnitude: the proportion of present persons who are participating in its performance. The limit case, when all talk momentarily ceases, can only be understood as a collaborative achievement, for such a state of quiet requires the participation of everyone present; even a single non-conformer prevents the achievement.

Group silence cannot be explained by simple logical extension of conversation and discourse analysis for it passes beyond the analytic territory of these tools. What material explanation can account for group silences arising? That is, by what means are present persons synchronizing their sound-making behavior despite giving no indication that their attentions are oriented to anything other than their private conversations? In order to establish empirical ground on which to propose answers to these questions, I present the following ethnographic data, drawn from a study on group silence I conducted within a Quaker community. These materials offer insight into how such silences unfold in realtime and why we must attend to the *embodied practice* of group silence in order to adequately explain it. After analyzing group silence in the Quaker context, I will return to the topic of group silence in general.

Ethnographic Study of Quaker Silence

Study Setting and Methods

In order to investigate social silence ethnographically, I sought out naturalistic laboratories where silences routinely arise in the social soundscape. I marginalized myself at social gatherings in order to listen to the whole room at once. I let silences enter my conversations with others in order to (tactlessly) observe the consequences. The community that became the eventual focus of my research has made group silence their central religious

activity for over three-and-a-half centuries. The Religious Society of Friends, commonly known as Quakers, gather regularly for practices known as ‘Meeting for Worship’ and ‘Worship Sharing,’ where participants sit in a circle, in deliberate silence, for religious contemplation. Silent worship is punctuated by brief, spontaneous speech acts, known as ‘ministry,’ where participants share, if inwardly moved, insights that have just occurred to them in the silence.

I conducted a four year ethnographic study of Pacific Friends Meeting, an established Quaker community on the West coast of the United States. The primary method was participant-observation in all aspects of community life: religious, social, and committee gatherings. In the final stage of research, I captured the first-ever video recordings of Quaker worship. This video study was designed to capture the embodied practice of group silence in fine detail. By combining participant-observation with software-based video analysis, I was able to verify key ethnographic findings with video evidence. The present paper offers a glimpse of this data diversity in a highly abbreviated form. For the whole picture, see (Steinbock, 2012).

The video study was conducted over a six week period in weekly 90-minute sessions. I used a 360-degree panoramic video camera to record groups of Quakers practicing ‘worship sharing,’ with groups ranging in size from six to thirteen participants, most attending multiple sessions. Due to the unusually specific nature of the data, I wrote custom software to analyze the video using computer vision and modeling techniques. Movement detection algorithms identified who was moving at any particular instant and the magnitude of their movement from one frame to the next (1152 x 320 pixel resolution, 15 frames per second). Counter-intuitively, this *visual* data turned out to be more representative of the perceivable soundscape of group silence than the corresponding audio data. A video camera is a more sensitive detector of the nearly inaudible, small body movements that characterize Quaker silence, whereas a microphone is unable to distinguish such tiny sounds from background noise. More importantly, the video image locates the movement (and thus the sound) precisely in space, so that I could justifiably assign responsibility to a specific participant (see Figure 1, showing Quakers’ bodies delimited for movement analysis). On this basis, the structure and timing of participants’ movement-sound and stillness-silence was precisely mapped and measured.



Figure 1. A frame of panoramic video shows Quakers practicing group silence.

Embodied Practice of Quaker Group Silence

Over the course of four years participating in Quaker silent worship, I listened to the soundscape of this culturally-situated social silence and observed the embodied behaviors that constitute it. In result, I identified five distinct ways that Quakers ‘do’ silence (Steinbock, 2012). Three occur during the formal practice of worship, which I termed *settling*, *ministering*, and *gathering* silence. The remaining two, *arriving* and *integrating*, appear immediately before and after formal worship, respectively. I focus on *settling silence* in this paper for it illustrates the underlying embodied order for *all* the Quaker silences.

When Quaker worship officially begins and participants settle into deliberate silence, the room becomes much quieter than it was a moment earlier. Relative to its former clamor, the room is utterly silent. And yet, the disappearance of noise from the room makes it possible to hear what ‘silence’ actually consists of: the tiny sounds of movements that living bodies make (shifting posture, sighing, scratching, sniffing, and so on). Though ever-present, these sounds go unnoticed under the covering scuffle of ordinary activity and talk. But when silence falls over a Quaker gathering, these sounds become salient. We could say that the practice of group silence ‘re-calibrates’ auditory perception such that quiet body sounds become accountable, noticeable events. As I will attempt to show, by becoming accountable, they become *consequential*, modifying the social interaction field in observable ways.

Given the inevitable noises produced by embodied beings, what sort of soundscape should we expect to find in the silence of Quaker worship? Presumably individuals cough, sniffle and fidget on their own physiological schedules, so we would expect to find a more-or-less uniform distribution of individual sound-making against a silent background, sometimes a little louder or softer, as chance would have independent physiologies align in time.

This is a reasonable expectation but it is not to be found in the Quaker silence. Instead, individual body-sounds arise in definite *group formation*, clustered together in time, and separated by long gulfs of quiet. One person's cough, for instance, is almost inevitably accompanied by a cascade of sounds from several individuals around the room. After the cascade has run its course, the soundscape returns to prolonged quietude, that is, before another 'wave' of sound inevitably emerges. This ebb and flow pattern of embodied synchrony (Condon & Ogston, 1966) was consistently observed across different days and participants. A similar cascade pattern was observed among British Quakers (Dandelion, 1996), though only in connection to overt causes like the arrival of latecomers. When I attended to the soundscape of silence itself over scores of sessions, waves of spontaneous, synchronized, embodied sound and silence proved to be the defining feature. These observations were confirmed by quantitative analysis of movement in my video study of Quakers practicing group silence. See Table 1, showing movement events detected across all six 90-minute sessions, according to the number of people synchronously moving at the same time (event size). The data show that Quakers tend to move together, i.e. it is relatively rare for a single person to move without others following suit. Averaged across all sessions, Quakers moved together 88.5% of the time, lending quantitative support for the ethnographically-observed finding.

Table 1: Synchronous movement events detected in all video-recorded sessions of Quaker worship, sorted by event size. (Size is the number of participants moving synchronously (within 6 seconds of each other), from 1 to N. Bottom row shows the percentage of movement events where more than one person participated.)

Event Size (# of participants)	Session 1 (N=8)	Session 2 (N=11)	Session 3 (N=13)	Session 4 (N=12)	Session 5 (N=6)	Session 6 (N=13)
1	14	12	6	1	8	2
2	22	12	3	1	8	1
3	15	12	3	2	12	4
4	20	6	6	2	16	6
5	5	9	5	6	8	4
6	11	13	1	4	14	7
7	4	4	1	4		4
8	4	4	2	3		2
9		2	0	1		2
10		3	1	5		1
11		0	1	3		2
12			5	0		2
13			0			0
Total	95	77	34	32	66	37
Total > size 1	81	65	28	31	58	35
Percent of events > size 1	0.85	0.84	0.82	0.97	0.88	0.95

Calibrating to the Social-Acoustic Environment

In the deep quiet of Quaker worship, a sound as subtle as a scratch or as overt as a cough transforms the social-acoustic environment from one where a norm of quietude holds sway to one where a little sound is permitted. The initial transgression provides 'cover' for others to follow with apparently lower social cost, analogous to the 'broken windows' theory of social norms (Keizer, et al., 2008). In turn, the sound contributions of additional participants further transform the norm, making it even more permissive, and resulting in the cascade of sound described above. As participants come to complete their posture changes, coughs, etc., silence once again settles over the room, and this event, too, becomes something people normatively join in with. These dynamics could be described as complementary feedback loops, as seen in autopoietic complex systems (Maturana et al, 1973).

The synchrony of Quakers' sounds and silences illustrates a key idea: *calibrating to the social-acoustic environment is a basic way individual behavior is situated in its social context*. Furthermore, the acoustic environment is not an external given; everyone present co-participates in its creation. In this way, the group continuously negotiates a working consensus as to the appropriate volume level for sound-making behavior. Thus is seen the difference between the official rules for culturally-appropriate behavior in a situation (e.g. Quakers worship in silence) and the actual practice of persons participating in contingent interaction, sharing the responsibility of co-determining what constitutes appropriate behavior at any given moment.

Given such precise synchrony in the Quaker practice of silence, we might go looking for the communications medium individuals use to announce their respective plans of action and work to agree on what the consensus will be at any given moment. Scholars have repeatedly shown the multiplex communicative

power of embodiment in orchestrating realtime collaborative action (Schütz, 1951; Birdwhistell, 1970; McDermott, et al., 1978; Goodwin, 2011). In the Quaker case, we can say that the *embodied practice of silence* is the communications medium in that deliberate group silence establishes an environment in which the social order can be co-managed by the subtle bodily manipulation of sound and silence.

Is the embodied synchrony of Quakers merely an epiphenomenon of unconsciously negotiating a social behavior norm or it is possible that it plays a functional part in their purposeful practice? At the very least, Quakers' synchronized sounds and silences imply they are more aware of each others' bodily presence than their closed-eyed meditations might otherwise suggest. This inference was affirmed by study participants when they described how they experienced silence during worship. They spoke in various ways of the supportive role that the presence of others played in their personal meditative practice, specifically calling out bodily sounds and silences as reminders to focus on the group activity of worship:

In both Quaker contexts and in Buddhist contexts, I've experienced the much greater power of a group sitting together, worshipping or meditating as the case may be. There's just something about having a hundred people together in the room just being together in silence. (Sasha, 11-Oct-2010)

I am attuned to social cues of movement and body language — shifts in posture that speak of your restless searching, as my body says the same to you. I hear you with more than my ears... These social cues of the presence of others in worship — the rustles, breathing, coughs — remind me that I am not alone, that you are here too. (August, 10-Apr-2011)

We can begin to see what the bodily awareness of others affords to Quaker participants. Body sounds in close proximity remind the meditator that he or she is not alone and is gathered with others for the collaborative task at hand: group religious contemplation. Recall that the soundscape to which participants are referring when they speak of the reminding power of body sounds has a particular temporal structure, i.e. the very shape of a recurrent reminder, repeatedly calling participants back to the meditation by re-asserting the embodied presence of others. Here lies the link between the social practice of group silence and the psychological process of meditation.

[W]hat's different between an individual experience of prayer I have and one I have in Meeting for Worship is that there is an awareness of everyone else's intentionality... we're there with the same or very similar intentions and that brings the experience a certain power. But with that also comes a responsibility... that I'm there too and I need to be aware of why I'm there and keep coming back to it. It's okay if for a minute I get distracted by something that comes into my head, but out of that shared intentionality for the group comes this desire for me to really stay focused in a way that I find difficult... to be mindful for the group, who's also supporting me through their same intentionality as we all do this collective worship together. (Thomas, 11-Oct-2010)

Quaker silence is the sound of many people meditating in close proximity. While a person meditating alone may be prone to distraction, Quaker silence provides an acoustic environment that recurrently renews one's intentionality. If the rustling of others reminds the meditator that he or she is not alone and is gathered with others for a shared purpose, the ensuing silence symbolizes that purpose: attention to the felt presence of immediate experience, the present moment. Thus the social-acoustic dynamics of group silence are linked to the purposeful psycho-spiritual practices of those present. With the passing of each wave of body sound, participants gain the recurrent experience of a lasting silence settling over the room, reminding them of why they are gathered.

The Vietnamese Buddhist teacher, Thich Nhat Hahn, writes about how in his tradition there are temple bells that chime, and hearing them, people are encouraged to pause and think to themselves: "Listen, listen. This wonderful sound calls me back to my true self." And he points out that in Western societies we don't have so many temple bells that we can hear easily... In Friends Meetings, we worship in silence, and it's the *silence* that calls me back to my true self. (Andrew, 25-Oct-2010)

In summary, as Quakers calibrate their sound-making behavior to the local environment, moments are repeatedly organized for silence to become contagious. Participants are reminded to direct their attention to the present that silence exposes: people gathered together, alive to each other and an unfolding group experience. For Quakers and other contemplative practitioners, attention to the present is a practice for generating insight; by exposing the habitual patterns of mind that carry attention away from the present, insight emerges. Over time, distraction may decrease as the mind settles on the present. Over the course of a single Quaker worship

session, recurrent silence settles the congregation into more stable, longer-lasting periods of quiet. Only from these deeply quiet times can the other two silences of Quaker practice, *ministering* and *gathering*, emerge (Steinbock, 2012).

An Explanation for Inexplicable Group Silence

Two key insights have emerged from the preceding discussion: 1) Calibrating to the social-acoustic environment is a basic way that individual behavior is situated in its social context. 2) Group silence directs collective attention to the present moment. Now we're in a position to return to the canonical cocktail party and see both of these principles at work in the everyday phenomenon of spontaneous group silence. I'll describe three kinds of group silence in turn, each produced by the interaction of these two principles. All three are potential grounds for discomfort.

The first group silence has an obvious environmental cause. In any social setting, the social-acoustic environment determines an appropriate volume range for talk and other sound-making. This has both social and pragmatic factors. One must speak loud enough to be heard by one's interlocutors but not so loud as to expose one's talk beyond its intended audience. Appropriate speaking volume depends entirely on acoustic context: a rock concert will differ from an art museum. Furthermore, contexts are acoustically dynamic, requiring correspondingly dynamic responses from participants: at a party, when the background music suddenly goes silent, speakers instantly drop their voices in adaptation. Someone who fails to do so is inadvertently thrust 'on-stage' as their talk is exposed to a suddenly expanded audience. To be thus caught out of tune with the environment may be cause for embarrassment. Such is the importance of staying both socially and acoustically calibrated to the immediate context.

The second group silence has no obvious warrant and so people must go looking for it. As discussed early in this paper, the implicit rule of conversation is to maintain continuous talk. At a party, where many conversations are ongoing in parallel, clamor is the norm and silence a deviation for which warrant must be present. If, while one is speaking, other speakers in the room go quiet for no apparent reason, the appropriately-calibrated response is to follow suit. This adds further power to the signal that stimulated the response. Whereas a change in background music is an external cause for social silence, here the operant signal is the sound of silence itself. Those present may momentarily turn away from the members of their conversation and look around in search of warrant for the silence: perhaps the birthday cake is being carried into the room. When the inexplicableness is resolved by an object or outlet for collective attention, we can say the group silence is *indexical*: it calls attention to something. This is a useful silence, a signal for collecting and directing group attention, just as we saw in Quaker silence.

The contagious and functional qualities of indexical group silence may indicate its evolutionary roots. Several animal species signal alarm to their mates through silence and motionlessness (Dapper, et al., 2011; Pereira et al., 2012). Joining in with group silence also corresponds to the contagious human behavior of group looking (Milgram, Bickman & Berkowitz, 1969), where there's generally good reason to look where a group of others are looking: something dangerous or delicious this way comes.

The third and final kind of group silence is seemingly inexplicable: the case where no warrant is found. Spontaneous silences may emerge, for instance, when inter-speaker silences coincidentally align across different conversations, creating the impression of a warranted indexical silence. As above, conversations momentarily drop in volume or cease altogether. The acoustic boundaries between different conversations weaken or break down altogether, making it difficult to speak for one's intended audience alone, and further adding to the tendency toward not speaking. Drawn together into a shared awareness of their mutual silence, the group casts about for an object or outlet for its now-gathered attention. Something is expected but nothing is happening. Painful associations with past conversational lapses may be triggered, except this silence is multiplied by N. Participants are now caught in a state from which it is difficult to escape: any utterance or overt body sound/gesture produced in such a quiet environment will instantly become the outlet for pent-up attention, exposed on stage to a greatly expanded audience and to greater risk for embarrassment. Thus bodies tend to freeze into motionlessness as voices fall quiet, everyone waiting to discover who will break the silence. (Classroom teachers are intimately familiar with a similar situation, when a question put to students is answered with silent, collective fear of taking the floor.)

Fortunately for party-goers, repair is usually forthcoming, but not through individual heroics. A Quaker-like wave of synchronous noise or utterance is enough to pull everyone back into the comforting din of small talk. In practice, a momentary group silence may be uncomfortably obvious or so brief as to be unnoticeable. In fact, upon inspection, the soundscapes of social gatherings *always* contain brief, spontaneous group silences that exhibit the dynamics described above. This only goes to show how sensitive and responsive people are to their social-acoustic environments. That extreme instances occasionally rise to the level of conscious discomfort should not be taken as evidence of communicative failure, but rather as dazzling displays of social awareness and collective action.

Conclusion

Calibrating to the social-acoustic environment is a basic way individual behavior is situated in its social context. With intra-speaker silence, we learned that an effective dramatic pause is performed by sensitivity to the audience's embodied social cues. With inter-speaker silence, we learned that turn-taking conversationalists aim to keep talk continuous, with a minimum of gap and overlap; this, too, is achieved by sensing the embodied cues of others to anticipate the timing of their utterances. With group silence, we found a medium for collecting group attention, which serves Quakers' contemplative purposes but may be cause for discomfort in some everyday social settings.

Future research could make a fruitful study of social soundscapes, particularly to examine how embodied synchronies of sound and silence show up outside the Quaker context. I have observed, for instance, that speakers in adjacent conversations unconsciously synchronize the cadence of their utterances and silences, especially when intimate subjects are being discussed, to collaboratively maintain mutual privacy.

Social silence is an underappreciated element of everyday social interaction that deserves wider awareness, by both scholars and everyday conversationalists. Like laughter and applause, silence is a medium for sharing embodied experience: it says something important, without words, about the situation in which it occurs. Perhaps the next time you find yourself in a group where an inexplicable silence is arising, the sense that you are falling into an uncomfortable void may be replaced by the sense that you have fallen into a more intimate relationship with the people around you — something worthy of appreciation.

References

- Birdwhistell, R. L. (1970). *Kinesics and context: essays on body motion communication*. Philadelphia: University of Pennsylvania Press.
- Condon, W. S., Ogston, W. D. (1966). Sound Film Analysis of Normal and Pathological Behavior Patterns. *Journal of Nervous and Mental Disease*, 143(4): 338-347.
- Dandelion, P. (1996). *A Sociological Analysis of the Theology of Quakers: The Silent Revolution*. Lampeter, UK: Edwin Mellen Press.
- Dapper, A.L., Baugh, A.T., Ryan, M.J. (2011). The Sounds of Silence as an Alarm Cue in Túngara Frogs, *Physalaemus pustulosus*. *Biotropica*. 43(3): 380-385.
- Goffman, E. (1967). *Interaction Ritual: Essays on Face-to-Face Behavior*. New York: Doubleday.
- Goodwin, C. (2011). Contextures of action. In J. Streeck, C. Goodwin & C.D. LeBaron (Eds.), *Embodied Interaction: Language and body in the material world*. Cambridge: Cambridge University Press.
- Keizer, K., Lindenberg, S., Steg, L. (2008). The Spreading of Disorder. *Science*, 322(5908): 1681-1685.
- Maturana, H. & Varela, F. (1980). Autopoiesis and Cognition: the Realization of the Living. In Robert S. Cohen and Marx W. Wartofsky (Eds.), *Boston Studies in the Philosophy of Science* 42. Dordrecht: D. Reidel Pub. Co.
- McDermott, R.P., Gospodinoff, K., Aron, J. (1978). Criteria for an ethnographically adequate description of concerted activities and their contexts. *Semiotica*, 24(3-4): 245-75.
- McLaughlin, M. & Cody, M. (1982). Awkward silences: Behavioral antecedents and consequences of the conversational lapse. *Human Communication Research*, 8(4): 299-316.
- Milgram, S., Bickman, L. Berkowitz, L. (1969). Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology*, 13(2): 79-82.
- Newman, H. M. (1982). The sounds of silence in communicative encounters. *Communication Qlty.*, 30(2): 142-149.
- Pereira, A.G., Cruz, A., Lima, S.Q., Moita, M.A. (2012). Silence resulting from the cessation of movement signals danger. *Current Biology*, 22(16): R627-R628.
- Sacks, H., & Jefferson, G. (1995). *Lectures on conversation*. Oxford, UK: Blackwell.
- Sacks, H., Schegloff, E.A., Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language* 15(4): 696-735.
- Schegloff, E.A. (2007). *Sequence Organization in Interaction*. Cambridge: Cambridge University Press.
- Schütz, A. (1951). Making Music Together: A Study in Social Relationship. *Social Research*. 18(1): 76-97.
- Steinbock, D. (2012). *Making Silence Together: Collaboration in the silent gatherings of Quakers*. (Doctoral Dissertation). Retrieved from www.steinbock.org/pubs/steinbock_dissertation.pdf
- Tannen, D., & Saville-Troike, M. (1985). *Perspectives on silence*. Norwood, N.J.: Ablex Pub. Corp.
- Twain, M. & De Voto, B.A. (1940). *Mark Twain in eruption: hitherto unpublished pages about men and events*. New York: Harper.

Acknowledgements

This work was supported by an NSF grant (NSF#0354453) to the LIFE Science of Learning Center. Thanks to Hiroki Oura for comments on an early draft of this material.

Designing for Democracy in Education: Participatory Design and the Learning Sciences

Betsy DiSalvo and Carl DiSalvo,
Georgia Institute of Technology
85 5th Street NE, Atlanta, GA 30332
bdisalvo@cc.gatech.edu, carl.disalvo@lmc.gatech.edu

Abstract: Within the learning sciences the concept of design is important in developing learning environments and conducting research. We propose the field of learning sciences seek better understandings of design from the field of design. In this paper we focus on Participatory Design (PD), where goals for real world impact and democratic process resonate with goals of the learning science. Our reflection on applying PD in the learning sciences suggest a more democratic and effective methods for objectives such as: defining research and learning goals, practices that consider multiple participants in design, practices for establishing participatory living labs for design research, developing infrastructure for sustainable participation, and using PD as a way to seek transferable rather than generalizable outcomes. Finally, we reflect on how PD might help us answer some of the issues with interest driven learning that motivated our work.

The effort of the last two-thirds of a century has been successful in building up the machinery of a democracy of mind. It has provided the ways and means for housing and equipping intelligence. What remains is that the thought-activity of the individual, whether teacher or student, be permitted and encouraged to take working possession of this machinery: to substitute its rightful lordship for an inherited servility. (Dewey, 1903, p 193-194)

Introduction

The goal of the learning sciences is to not only understand the phenomena of learning, but also to impact current practices to enable more effective learning. Part of this includes building things, putting them in the world, testing them, refining them and iterating. We propose that the learning sciences would benefit by reflecting on this process by exploring the discipline of design and design research. One method within the discipline of design, participatory design (PD), offers particular promise for the learning sciences because of aligned goals for engaging in the real world applications and shared democratic values.

Although the quote that started this paper is from 1903, we still, persistently, struggle to focus education on the learner, their self-motivation, identity, interest, and what they find personally meaningful, rather than simply meeting criteria of content acquisition. PD offers opportunities to uncover self-motivations, identities, and interests and to construct meaningful engagements by working together with participants to co-creating learning environments that meet the needs of the whole community that is engaged in the learning practice.

In this paper we will particularly reflect on challenges with interest driven learning that may be ameliorated with PD. While this is not the only area of the learning sciences that can benefit from PD it offers a number of concrete problems where PD methods may prove useful. Beyond interest based learning will explore some other current uses of PD in the learning sciences. Then we will introduce participatory design, as it is understood in the field of design. And finally we will reflect on some concrete contributions that PD may offer the learning sciences and how it could impact interest driven learning.

Challenges in Designing Interest Driven Learning

Currently many learning activities work to build upon the *content interest* of the youth such as leveraging young people's interest in games (Y. B. Kafai, 2006), new media (Forte & Guzdial, 2004) or focus on maker culture (Yasmin B Kafai et al., 2013). A constructionist perspective suggests that learning is most effective when the learner is constructing a *personally meaningful* product (Papert, 1993). By creating learning experiences around content interest or personal interest designers seek to increase motivation and persistence with learning (Schiefele, 1991). However designing and implementing effective interest driven learning activities poses many challenges, including: What is interesting or personally meaningful to young people?; How do we balance learning and interest in design?; and How can we create persistence with interest driven learning?

What Is Interesting or Personally Meaningful to Young People?

Identifying what content is interesting or what activities are personally meaningful young people can be more difficult than we assume. In some theories of interest it is suggested that interest is content-specific, and it is directive to how a student is motivated to learn. This type of content interest can be short lived or enduring (Schiefele, 1991). In contrast, use of the phrase *personally meaningful* is frequently used in constructionist theory to speak to interest that are more value driven. The ways that personally meaningful is used suggests interest that are not necessarily content specific, but have a value to young people as individuals, rather than simply as cultural signifiers. Some ways that are suggested to create more personally meaningful experiences are to offer learners a chance to present their work in a public or authentic manor (Shaffer & Resnick, 1999), to have learners engage in narrative (Bruckman & Resnick, 1995), or to imagine future use of learning (Luckner & Nadler, 1997). These notions of creating motivation based upon the individual's interest, whether content interest or personally meaningful, are more complex than simply looking at what is popular in youth culture and require an ability for learning scientist to discover with youth what content or values motivate learning.

How Do We Balance Learning and Interest In Design?

Even after we uncover the interest of youth it may be difficult to leverage that into a learning experience. Many have lamented the *edutainment* (Resnick, 2004) or gamification (Bogost, 2011) of learning that take popular or “fun” activities like current media or games and use them to candy coat learning. Edelson and Joseph (2004) outline a number of challenges in the logistics of basing learning responsive to individuals' interest:

First, it requires rare levels of internal motivation on the part of children. Second, it provides no mechanism for promoting learning objectives even marginally outside of learners' interest. Finally, it requires impractical levels of resources and flexibility in serving the divergent interest of individuals. (Edelson & Joseph, 2004, p 168)

How Can We Create Persistence with Interest Driven Learning?

While there are many promising interest driven learning experiences, it is challenging to engage in persistence with learning based upon interest alone. Dewey suggest, in his book *Interest and Effort in Education*, that external attempts may catch someone's interest, but holding it and engaging learners in persistent effort and identification with a topic area are much more difficult (Dewey, 1913). Edelson and Joseph have addressed this in their Interest-Driven Learning Design Framework (IDLDF) by focusing on what the learner will perceive of having a future usefulness outlining five methods for sources of usefulness for designers to use in generating interest; Pleasure, Concern, Identity formation, Life goals, and Curiosity (Edelson & Joseph, 2004). However, this framework assumes that the designer or researcher can identify what is pleasurable or of concern to a learner, what activities do they attach to identity or life goals, or simply what makes a young learner curious.

Design and Participation in the Learning Sciences

The learning sciences use the word *design* to describe many of the methods used to plan, develop, test, and iterate learning programs. In addition, there has recently been an emphasis on *participatory* culture in scholarship related to youth culture (Jenkins, 2009). While the use of these two terms has little direct relationship to participatory design (PD) as it is defined in the field of design, PD still closely aligns with a number of efforts in the learning sciences.

The Learning Sciences' emphasis on approaches such as Learner Centered Design (LCD), Design-based Research (DbR) and Community Based Design Research suggest that the field is receptive to thinking about design critically. In LCD the learner is considered as central to the design process in terms of consideration rather than the teacher or facilitator (Soloway, Guzdial, & Hay, 1994). In DbR the emphasis is on using design to test and building theories in the ever changing and confusing real world rather than a lab setting. To do so it is imperative to engage participants in real world settings and to gather their input in the process (Barab & Squire, 2004). In both of these approaches engage in aspects of prototyping, testing and iterating with participants makes them participatory by nature however neither would be engage in PD as it is understood in the field of design.

Community based design research as it has been realized by Bang and colleagues with indigenous communities grows from a tradition of Participatory Action Research (PAR) (Bang, Medin, Washinawatok, & Chapman, 2010). PAR and PD share a commitment to working with communities to articulate issues of concern and develop plans of action toward those concerns. But whereas PAR is grounded in inquiry, PD is grounded in design itself. Community based design research could be seen as a bridge between these traditions and practices, but this line of thinking has yet to be developed in the learning sciences.

There are few examples of participatory design in learning, for instance, Druin has lead a number of efforts the focus on PD in research on and development of human computer interaction for young people (Druin, 2013). While the development of new technology in her work is not explicitly focused on learning,

because the participants are youth, frequently learning is a goal of the design process (Druin, 2002). Spikol, Milrad, Maldonado and Pea brought a strong Scandinavian influenced co-design approach to the development of mobile science collaborations (Spikol, Milrad, Maldonado, & Pea, 2009). DiSalvo has included elements of participatory design methods in learning science research (B. DiSalvo, 2012) and in collaborative work with the Carnegie Mellon School of Design on the Click! Urban Adventure (B. DiSalvo, Parikh, & Crowley, 2006; Hughes, 2007).

Background of Participatory Design

The origins of participatory design can be traced to Scandinavia in the 1970 and 80s (Ehn 1993, Simonson and Roberston 2012). Early participatory design projects focused on the workplace and were often conducted in conjunction with unions. Participatory design was seen as an approach to ensure democratic agendas in labor, particularly as information technology and automation was introduced into work environments (Simonson and Roberston 2012). From the earliest days, then, participatory design has been at one and the same time a set of methods, a practice of engagement, and a commitment to a particular set of political values—all enacted through design.

Though still niche within design, over the past several decades participatory design has developed into a rich domain of research and practice, extending far beyond its origins in Scandinavia. Workplace environments are still central to participatory design, and in addition to the factory floor there are substantial case studies across fields of work, most notably in healthcare Sjöberg, C., & Timpka, T. (1998) and in government (Anthopoulos, L. G., Siozos, P., & Tsoukalas, I. A. (2007). But as information and communication technologies have changed and expanded beyond the workplace, the field sites of participatory design have likewise changed and expanded. Increasingly there are examples of community-based participatory design (C. DiSalvo, Clement, and Pipek, 2012) that continue with the methods, practices, and commitments of participatory design but in new contexts, with new actors, and new purposes.

For example, in the city of Malmö, Sweden there are a series of “Living Labs” that have become sites for a new mode of participatory design that engages residents in the open exploration of inventive creative uses of information and communication technologies (Ehn 2008; Björgvinsson, E., Ehn, P., & Hillgren, P. A. 2010). Notably, many of the residents that participate in these Living Lab programs are immigrants to Sweden. What these programs provide are opportunities to conceptualize and construct systems that function as experiments in cultural expression, for example, sharing food practices or enabling DIY music production. Other examples of community-based participatory design explore engagement in technical practices such as environmental sensing and monitoring. Neighborhood Networks was one such project that brought together designers and neighborhood residents to explore the practices of environmental monitoring and prototype civic interventions and alternate forms of expressing environmental data (C. DiSalvo, et al, 2008). What the Neighborhood Networks project demonstrates is an approach to participatory design that seeks to directly engage the public in the activities of design, not simply for instrumental ends, but as a kind of co-operative study of the potential of information and communication technologies.

Pelle Ehn, one of the early practitioners of participatory design, has articulated these changes in participatory design through a series of papers that explore what he refers to as “participation in design things.” (Ehn 2008) These so-called design things are the instantiation of contemporary controversies, such as citizenship or pollution (to draw from the examples above). To participate in design things is to collaborate in both an investigation of the conditions of these controversies (who counts as a citizen? what counts as pollution?) and also to explore how we might respond to these conditions—all through design.

These new practices of participatory design are challenging to the participatory design community, and to the practices of professional design and design research more generally, because their purpose is often seen to be obscure. Certainly, it is more nebulous than the early work of the field in which the relationship between the design of technical systems and the agency of workers was (or at was seen to be) direct and an unambiguous democratic concern. Still, for Ehn, and many others this mode of participatory design is an extension of the foundational methods, practice of engagement, and commitment to democratic values. What is different is that experiences and events take the place of the technical systems. That is, rather than the end goal being the design of an operational system, the end goal is an experience or event that develops the agency of participants. This may involve a technical system, but more as a prop or process. The activities of participatory design become activities of *infrastructuring*, that is providing the resources necessary to prompt, support, and sustain, this collective and collaborative inquiry through design (Ehn 2008; C. DiSalvo, Clement, and Pipek, 2012)).

Arguably, learning is an implicit aspect of many of these new participatory design endeavors. Oftentimes there is an aspect of the projects that involves the development of technological literacy or fluency. In some cases, that is an explicit aspect of the project, in others, a by-product of the design activities. For example, in the Neighborhood Networks project participants learned about using simple sensors for monitoring various aspects of air quality and sound levels. In addition to learning the operation of devices, they also encountered the limitations of those devices in use. And too, they had to consider, and ultimately construct,

methods of communicating that data in a public forum to their neighbors and other interested parties. To achieve this, the designers designed a series of activities and formats — the infrastructure of the project. Likewise with the projects of the Malmö Living Labs — *infrastructuing* becomes central to these new modes of participatory design.

Reflection on What PD Offers the Learning Sciences

There are a number of ways that the methods, practices for engagement and democratic goals of PD can provide a useful framework for the use of design with in the learning sciences. In the following section we reflect upon the use of PD in the learning sciences and envision five ways that PD methods, practices and goals can enhance the research and output of the learning sciences. These ways include, democratic practices for defining research and learning goals, practices that consider multiple participants in design, practices for establishing participatory living labs for design research, developing infrastructure for sustainable participation, and using PD as a way to seek transferable rather than generalizable outcomes. Finally, we will reflect on how PD might help us answer some of the issues with interest driven learning that motivated our work.

Democratic Research and Learning Goals

Frequently in the learning sciences previous findings, policy, administrative directives, or the passions and interest of the researcher drive our goals for research or learning outcomes. PD proposes a different approach to setting research and learning goals. In PD it is the community or the participants one is working with that establish the goals of design. In practice much of the work of PD is in providing the tools and establishing a dialogue with the participants so they can set the agenda for new design.

This work is often playful or provocative activities that help participants feel empowered to share their perspective no matter their level of expertise. Examples of activities are games that involve juxtaposition of images, mapping conceptual ideas on geographic maps, or getting participants to create reflective art or craft projects. At times the outcomes reinforce our understanding or inspire new lines of inquiry, but at other times they can be difficult to interpret. It is important to note that interpretation of such activities frequently is secondary to establishing a dialogue and vocabulary between the designer and the participants. We see this work of PD to be beneficial to researchers because it helps break down knowledge, language and power barriers that limit the ability of participants to share in setting goals and design directions.

Considering Multiple Participants

Many of the DbR classroom studies are designed with input from teachers; many constructionist programs are designed to give young people choices so they can direct their own learning. Both of these are reflective of PD practices, yet traditional PD methods call for designers to identify all of the participants and those impacted by the design. For example, in classroom environments teachers, students, and administrators would all be considered important participant in the design process, but other participants may also be included in the design process such as parents or policy makers. In informal learning environments, such as in designing museum exhibits the exhibit designers often consider the visitors from various backgrounds, such as experts, parents, youth, or the elderly. However there are other participants that may be overlooked in design, such as the museum educators, guards, and maintenance staff. We have outlined a few participants that may be overlooked in the design process or whom have roles that can be reconsidered with the use of PD approaches.

Teachers

Teachers are frequently already included in the design of classroom interventions because of their expertise, but also because they are necessary implementers – if teachers do not feel ownership and believe in an experimental project they will not implement it well. However, in other environments such as educational software, afterschool programs, or museums teachers are not frequently considered participants in the design process. If teachers are invited into early speculative design process about more informal learning environments they may provide new incites into connections and breakdowns in relationships to classroom learn.

Youth

Inviting youth into defining their own interest would seem like a necessary first step in designing interest driven learning, but they are often not part of the design process for learning scientist in the formative stages of design. The use of more playful or provocative PD activities may give youth the self-efficacy, skills and vocabulary to contribute to the design of interest driven learning. The structure outlined by the IDLDF identifies one way to structure a learning environment for direction of seeking relevance from the learner rather than selecting relevant topics based upon the designer or the teachers perspective (Edelson & Joseph, 2004). We suggest that methods and practices from PD will help learners contribute in more effective ways to the design process

Administrators, Policy Makers and Maintainers

With in a PD model administrators, policy makers, and those maintaining learning environments would be included in the design process. While these participants may not be considered users or learners in a LCD design approach their successful engagement with a learning activity or environment is as necessary as other participants. By inviting them in, and giving them tools to dialogue with teachers, youth, and other community members we can create a better understand between all participants and designs to bridge different goals.

Parents

Among youth one of the most influential factors in their life and their learning is parents, yet we frequently do not have their participation in the design of learning environments. Parents are resource providers, brokers and partners in learning (Barron, Martin, Takeuchi, & Fithian, 2009) and also strongly influence value systems for young people that impact their learning (B. DiSalvo, Bruckman, Guzdial, & Mcklin, in press).

Living Lab

Research with some participant who are resistant to learning or who resent changes imposed from outsider can be difficult. There is a distrust of outsiders, seeking to do “good” with in a community by imposing their own values (Freire, 1970; Illich, 1971). Successful community based research requires a long-term commitment, with goal setting and co-designing with communities (Stoecker, 2005). One approach to developing a long-term relationship with participants in education is the laboratory school model, where practice and research on learning theories have meet (University of Chicago 2013). Other models in learning sciences, while not necessarily build for this intention, have served as informal learning laboratory spaces such as the Computer Clubhouses (M. Resnick & Rusk, 1999). Existing living labs, such as a the Malmö Living Labs, suggest they allow for a more democratic approach to research with diverse audiences, allowing goals and innovations to emerge from bottom-up, rather than imposed from on high (Bjrgvinsson, Ehn, & Hillgren, 2010).

Infrastructuring

With a similar goal to living labs, infrastruturing is developing tools, such as Wikipaida, where the infrastructure is designed for participation in shaping the knowledge output and the ongoing design of the environment (Ehn, 2008). Within the field of learning sciences one can see that the online Scratch 2.0 programing community as infrastructure to encourage ongoing participation and, while perhaps not at formalized participatory infrastructure as Wikipedia, it does have strong participatory elements in the way the program is shaped (Resnick et al., 2009). We would encourage further reflection on infrastructure for participatory methods in the learning sciences as a way to design more sustainable participation in the design of the learning experience and therefore retaining the interest of learners.

Transferability Rather than Generalizalbity

A concern with DbR is the generalizability of findings because research is conducted outside of the laboratory in the highly variable classrooms or informal learning environments. By focusing on developing design methods or principles, rather than research interventions, the learning sciences gains an approach to offers transferable design findings that can be used new contexts. For example, the issues of designing participatory learning infrastructures addressed with a DbR approach could yield new theory and findings that can transfer to the design of other participatory learning infrastructures.

PD and Designing Interest Driven Learning

The interest driven learning approach is often about creating a “hook” that will get kids interested in an activity where learning is embedded. Frequently these efforts result in candy coated stale and not valued learning activities that might get immediate attention, but with little persistence. We outlined three questions in designing interest driven learning that addressed this challenge and will wrap up our reflection with exploring how PD can help us answer these questions.

What Is Interesting or Personally Meaningful to Young People?

We see PD as a way of identifying what is an interesting context to the participants, but also as a way to uncover values of the whole community involved in learning. The engagement of the whole community involved with young peoples learning will help in shaping fundamental goals for the environment, the research and the learning that will move a project beyond what is “cool” or “fun” and address issues of value to the learners and the people that matter to them.

How Do We Balance Learning and Interest in Design?

By inviting all of the participants to the design table we can engage in design that helps bridge conflicting goals. This means not only talking youth about what is interesting to them but also engaging them in design activities

about what is not interesting and what they resist. By allowing the young people and the others in their learning community to better understand motives we can design better interventions and address issues of content coverage that call into question what content needs to be covered and why.

How Can We Create Persistence with Interest Driven Learning?

By designing for participatory infrastructure learners and other community members will be engaged in ongoing design, with potential for learning environments that are flexible, democratic, and remain interesting.

Conclusion

The learning science relies heavily on the concept of design in establishing methods and practices. It is time that the field looks outside of the common use of design terms to understand the methods and practices of the field of design. In this paper we found that reflection on the methods, practices and goals of participatory design was fruitful in addressing issues in interest driven learning and in constructing new ways of looking at design in the learning science. In future work the use of PD can be leveraged to develop design processes for interest driven design, moving past what is “cool” and making learning valuable to learners and their community.

References

- Anthopoulos, L. G., Siozos, P., & Tsoukalas, I. A. (2007). Applying participatory design and collaboration in digital public services for discovering and re-designing e-Government services. *Government Information Quarterly*, 24(2), 353-376.
- Bang, Megan, Medin, Douglas, Washinawatok, Karen, & Chapman, Shannon. (2010). Innovations in culturally based science education through partnerships and community *New Science of Learning* (pp. 569-592): Springer.
- Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *The Journal of The Learning Sciences*, 13(1), 1-14.
- Barron, B., Martin, C.K., Takeuchi, L., & Fithian, R. (2009). Parents as learning partners in the development of technological fluency. *International Journal of Learning and Media*, 1(2), 55-77.
- Björgvinsson, E., Ehn, P., & Hillgren, P. A. (2010, November). Participatory design and democratizing innovation. In *Proceedings of the 11th Biennial Participatory Design Conference* (pp. 41-50). ACM.
- Bogost, I. (2011). Gamification is bullshit. *Ian Bogost*.
- Bruckman, Amy, & Resnick, Mitchel. (1995). The MediaMOO Project Constructionism and Professional Community. *Convergence: The International Journal of Research into New Media Technologies*, 1(1), 94-109.
- Dewey, John. (1903). Democracy in education. *The elementary school teacher*, 4(4), 193-204.
- Dewey, John. (1913). *Interest and effort in education*: Houghton Mifflin Company.
- DiSalvo, B., Bruckman, A., Guzdial, M., & Mcklin, T. (in press). Saving face while geeking out: Navigating motivations of non-learners. *Journal of Learning Sciences*.
- DiSalvo, B. (2012). *The Glitch Game Testers: The Design and Study of a Learning Environment for Computational Production with Young African American Males*. (Ph.D.), Georgia Institute of Technology, Atlanta, GA.
- DiSalvo, B., Parikh, A., & Crowley, K. (2006). Developing the ultimate urban adventure game for middle school girls. *Women in Games*.
- DiSalvo, C., Nourbakhsh, I., Holstius, D., Akin, A., & Louw, M. (2008, October). The Neighborhood Networks project: a case study of critical engagement and creative expression through participatory design. In *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008* (pp. 41-50). Indiana University.
- DiSalvo, C., Clement, A., & Pipek, V. (2012). Participatory Design For, With, and By Communities. *International Handbook of Participatory Design*. Simonsen, Jesper and Toni Robertson (Eds). Oxford: Routledge. (2012), 182-209.
- Druin, Allison. (2002). The role of children in the design of new technology. *Behaviour and information technology*, 21(1), 1-25.
- Druin, Allison. (2013). Kid Design. Retrieved 11/8/2013, 2013, from <http://www.cs.umd.edu/hcil/kiddesign/>
- Edelson, Daniel C, & Joseph, Diana M. (2004). *The interest-driven learning design framework: motivating learning through usefulness*. Paper presented at the Proceedings of the 6th international conference on Learning sciences.
- Ehn, P. (1993). Scandinavian design: On participation and skill. *Participatory design: Principles and practices*, 41-77.

- Ehn, P. (2008, October). Participation in design things. In *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008* (pp. 92-101). Indiana University.
- Forte, A., & Guzdial, M. (2004). Computers for communication, not calculation: media as a motivation and context for learning. *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*, 96-105.
- Freire, P. (1970). *Pedagogy of the Oppressed*. New York, NY: Herder and Herder.
- Hughes, K. (2007). Design to promote girls' agency through educational games: The Click! Urban adventure. In Y. B. Kafai, C. Heeter, J. Denner & J. Sun (Eds.), *Beyond Barbie and Mortal Kombat*. Cambridge, MA: MIT Press.
- Illich, Ivan. (1971). *Deschooling society*. New York.
- Jenkins, H. (2009). *Confronting the challenges of participatory culture: Media education for the 21st century*. The MIT Press.
- Kafai, Y. B. (2006). Playing and Making Games for Learning: Instructionist and Constructionist Perspectives for Game Studies. *Games and Culture*, 1(1), 36.
- Kafai, Yasmin B, Searle, Kristin, Kaplan, Eliot, Fields, Deborah, Lee, Eunkyong, & Lui, Debora. (2013). *Cupcake cushions, scooby doo shirts, and soft boomboxes: e-textiles in high school to promote computational concepts, practices, and perceptions*. Paper presented at the Proceeding of the 44th ACM technical symposium on Computer science education.
- Luckner, John L, & Nadler, Reldan S. (1997). *Processing the experience: Strategies to enhance and generalize learning*. ERIC.
- Papert, S. (1993). *Mindstorms: Children, computers, and powerful ideas*. New York, NY: Basic Books.
- Resnick, M., Maloney, J., Monroy-Hernandez, A., Rusk, N., Eastmond, E., Brennan, K., . . . Silverman, B. (2009). Scratch: programming for all. *Communications of the ACM*, 52(11), 60-67.
- Resnick, M., & Rusk, N. (1999). 11. The Computer Clubhouse: Technological Fluency in the Inner City. *High technology and low-income communities: prospects for the positive use of advanced information technology*.
- Resnick, Mitchel. (2004). Edutainment? No thanks. I prefer playful learning. *Associazione Civita Report on Edutainment*, 14.
- Schiefele, Ulrich. (1991). Interest, learning, and motivation. *Educational Psychologist*, 26(3-4), 299-323.
- Shaffer, David Williamson, & Resnick, Mitchel. (1999). "Thick" authenticity: New media and authentic learning. *Journal of Interactive Learning Research*, 10(2), 195-215.
- Simonsen, J., & Robertson, T. (Eds.). (2012). *Routledge handbook of participatory design*. London, UK: Routledge.
- Sjöberg, C., & Timpka, T. (1998). Participatory design of information systems in health care. *Journal of the American Medical Informatics Association*, 5(2), 177-183.
- Soloway, Elliot, Guzdial, Mark, & Hay, Kenneth E. (1994). Learner-centered design: The challenge for HCI in the 21st century. *interactions*, 1(2), 36-48.
- Spikol, Daniel, Milrad, Marcelo, Maldonado, Heidy, & Pea, Roy. (2009). Integrating co-design practices into the development of mobile science laboratories. Paper presented at the *Advanced Learning Technologies, 2009. ICALT 2009*. Ninth IEEE International Conference on Advanced Learning Technologies.
- Stoecker, R. (2005). *Research methods for community change: A project-based approach*. Thousand Oaks, CA: Sage Publications, Inc.

Leveraging the Cultural Practices of Science for Making Classroom Discourse Accessible to Emerging Bilingual Students

Enrique Suárez and Valerie Otero, University of Colorado – Boulder, 249 UCB, Boulder, Colorado, 80309
enrique.suarez@colorado.edu, valerie.otero@colorado.edu

Abstract: Despite extensive research and national reports that call for students' engagement in scientific practices, these strategies continue to be virtually absent in classrooms throughout the U.S. Thus, the question of how students who are learning English as a second language fare in these environments is virtually unanswered. This study presents results from an investigation on 3rd grade emerging bilingual students' participation in a physics lesson on sound production. We focus on the changes in participation as such changes pertain to scientific practices of argumentation, modeling, and experiential, imaginative, and mechanistic reasoning strategies. Drawing from students' discourse and gestures, conjectures are made about scientific practices being particularly well suited for fostering productive disciplinary engagement for emerging bilingual students.

Problem

There is a large number of students in the educational system who are learning English as a second language, with a small number of teachers trained to meet their language needs. Because learning environments are predominantly designed to accommodate English as the sole language of communication and instruction, emerging bilingual students are often not fully included in classroom activities. Forecasted shifts in the U.S. population predict an increase in the number of emerging bilingual students in schools (Kindler, 2002), which means an even greater number of students run the risk of being excluded from classroom activities. / Because of the emphasis on acquiring English as a second language, this group is often referred to as English Language Learners (ELLs), a label defines students solely based on language acquisition (Escamilla and Hopewell, 2010). Since we want to highlight the range of resources they bring to the classroom, throughout this study we use the term *Emerging Bilingual Students* to refer to these students.

With most school efforts focusing on increasing emerging bilingual students' fluency in English, science has lost priority in the elementary curriculum, which we see as a missed learning opportunity. Pressures to increase the language fluency of emerging bilingual students has led to a nationwide trend of teachers and administrators increasing the amount of time spent on regularly tested subjects such as mathematics and language arts. Consequently, there has been a significant decrease in classroom time devoted to science (McMurrer, 2008). However, science could provide opportunities for students to use language as a tool for communicating shared observations about the natural world and making evidence-supported claims. Thus, decreasing classroom science time may be counterproductive to students' development of language. We argue that a science classroom centered on scientific practices is particularly well suited for including emerging bilingual students. If the learning environment has certain conditions, we can tap into this certain thing students have, which can lead them to engage and participate in activities. By capitalizing on students' natural curiosity and desire to understand phenomena, learning environments can support students in engaging with peers in discussions about shared observations and questions that are meaningful to them. By participating in these kinds of conversation, students can simultaneously improve their English fluency and understanding of the physical world.

The creation and adoption of the Next Generation Science Standards (NGSS) in K-12 schools (Committee on Conceptual Framework for the New K-12 Science Education Standards, 2012) promises changes in the landscape of science education, prompting more research for understanding how students engage with disciplinary practices. As more states adopt NGSS, there is a push for integrating scientific practices such as model building, mechanistic reasoning, communication, and argumentation into classroom activities. And while there have been valuable efforts to understand the role language plays when emerging bilingual students learn science (Rosebery, Ogonowski, DiSchino, Warren, 2010; Suárez and Otero, 2013; Suárez and Otero, in press; Warren, Ballenger, Ogonowski, Rosebery, Hudicourt-Barnes, 2001; Warren, Ogonowski, Pothier, 2005), it is still not yet clear how to increase the participation of emerging bilingual students in science classroom activities. In this study we explore the hypothesis that science, physics in particular, is particularly well-suited for increasing the participation of emerging bilingual students in negotiating meaning and participating in class. We analyze a group of 3rd grade emerging bilingual students as they make sense of the sound produced by a guitar-like object. First, we present an analysis of how a student's proposed terminology for describing a physical phenomenon spread across the class and was adopted as a sort of "academic language" later used to talk about mechanisms on the basis of observations. Second, we analyze the participation and behaviors of one student who increasingly became more engaged as the discussion progressed. This work is relevant both practically and

theoretically. Theoretically, it contributes to conceptualizing how various scientific practices and academic language are mediated by students' everyday language and experiences. Practically, findings lead to concrete recommendations for how various features of a learning environment could increase participation of emerging bilingual students.

Theoretical Background and Literature Review

We view learning as a social practice (Rogoff, 1994; Wenger, 1998) and assume that students' development of conceptual understanding and language skills is co-constructed through social interactions. Therefore, learning environments that favor English as the only language for teaching and communicating creates a distinct set of challenges for emerging bilingual students. Specifically, students' language skills and discourse practices are often different than the ones valued and used in school settings, an incongruence that often leads to the disfranchisement of students and hinder learning (Warren *et al.*, 2001). In thinking about designing learning environments that bridge this language barrier, we propose that physics has a clear advantage over other subjects because experimentation typically involves the presence of tangible objects and shared observations. Moreover, scientific practices involve supporting and promoting student participation through reasoning, argumentation, and sharing ideas about observations of phenomena, all of which can be accessible to students with different levels of English fluency.

Researchers have proposed Productive Disciplinary Engagement (Engle & Conant, 2002) as a construct for describing active student participation. Specifically, this type of engagement is defined as one in which students spontaneously participate, substantially contribute, and attend to each others' ideas in a way that resembles disciplinary discourse practices and furthers intellectual progress. Four measures have been suggested to evaluate if a learning environment can foster this type of engagement. First, teachers should encourage students to problematize the content through questions, proposals, and challenges. Second, it is important for "students to be authors and producers of knowledge ... rather than mere consumers" (Engle & Conant, 2002; p. 404). Third, students should be held accountable, particularly by how their work is responsive to what community insiders and outsiders have established. Finally, it is necessary for students to have relevant intellectual and/or material resources to aide sense-making (Engle & Conant, 2002). While this is not the only framework that describes disciplinary engagement, we chose it because of how well the four principles align with disciplinary practice, especially authorship and accountability to the local and disciplinary communities.

We see almost a one-to-one correspondence between the four principles proposed by Engle and Conant (2002) and what we would expect to find in a learning environment based on inductive reasoning. To make the connection apparent, here we define scientific induction. The process of scientific induction ideally consists of observing physical phenomena and collecting evidence, followed by testing and postulating generally applicable principles supported by the data. Subsequently, these evidence-based claims are presented to the scientific community and subjected to the peer review process. The community ultimately (though not usually directly) arrives at consensus regarding inclusion of these principles in the larger corpus of knowledge.

Classrooms that espouse the scientific inductive process almost effortlessly meet the criteria described above for productive disciplinary engagement and promote meaning-making during activities. In these learning environments students problematize content through their observations and are authors of the evidence-based claims; these abstractions of are submitted to, and evaluated by, other students in the community. This framing aligns with Vygotsky's theory of concept formation (Vygotsky, 1962), which proposes learners move between two types of spaces: informal spaces that are populated with everyday experiences and interactions with the physical world; and academic spaces that house formal knowledge made available and validated by disciplinary communities. During the process of concept formation learners develop conceptual understanding through grounding academic concepts in everyday experiences, while leveraging these particular instances for deriving generalizable, academic concepts (Otero and Nathan, 2008).

Even though structuring classroom science activities on principles of PDE has great potential for increasing accessibility, it is important to consider what features of the learning environment are effective in supporting the inclusion and participation of emerging bilingual students. Some argue that what is needed is the creation of hybrid spaces (Gutierrez, Baquedano-López, Turner, 1997) where students can recruit everyday language and communication practices when sharing ideas associated with formal, academic terminology and classroom norms. These emergent *third spaces* integrate everyday space (counterscript) and academic space (official script), "creating the potential for authentic interaction and learning to occur" (Gutierrez *et al.*, p. 372). These third spaces allow for the trying out of new, academic terms through the negotiation of meaning by deploying one's own, everyday language and discourse practices. Gutierrez and her collaborators (1997) suggest that literacy development is related to students' deployment of everyday language in the service of testing of formal literacy practices. Through this bi-directional process, students construct and negotiate meaning. We argue that the same process is possible in science.

Framing students as active participants in the meaning negotiation process extends to the conceptualization of language development. Based on the work by Razfar and colleagues (Razfar, Licón Khisty,

Chval, 2010), we adopt a sociocultural perspective in which language development and construction of knowledge arises from interactions within classrooms. From this viewpoint, a sociocultural model of language development redefines “language as a mediational tool for learning rather than the object of learning and instruction” (p. 214). Razfar et al., (2010) prompt us to think about language development as concrete steps students take towards “the use of any word to signal an object in a decontextualized manner” (p. 201) and abstraction, which can only be achieved through situated meaning-making. This has significant implications for how teachers structure learning environments and activities for students to engage with. Particularly, the fundamental question for designing and organizing learning environments should be “how will students use language to reach the learning goals?”, rather than “what language do we expect students to acquire?” This position differs significantly from *second language acquisition* models, which construct language as an external device that needs to be acquired by the learner, instead of a mediational tool that is socially and iteratively constructed through interactions. Finally, sociocultural models avoid the pitfall of positioning learners as having a deficit that needs to be addressed by the teacher. Instead, it positions students as capable of learners whose development depends on the participation on multiple types of meaning making activities.

Building on scientific induction, models of student engagement, and models of English language development, we expand our initial conjecture. We hypothesize that science is particularly suited for creating hybrid spaces that provide opportunities for emerging bilingual students to become the authors and evaluators of evidence-based claims generated by shared, tangible experiences, furthering conceptual understanding. Simultaneously, science activities that promote the interaction between students’ everyday language and academic language can aide the processes of English language development, as well as in the construction of scientific language and meaning. In this study we address the following general question: *what role do everyday and academic language play when emerging bilingual students engage in making sense of the physical world?* In order to address this question, we analyze a five-week unit on sound, in which third grade emerging bilingual students establish a connection between physical properties of strings (length, tension, and frequency of vibration) and the characteristics of the sounds produced by them (pitch and volume).

Research Design and Procedure

Thirteen third grade students participated in this study, all enrolled in a large K-8 urban public school; the names of the participants are pseudonyms. The school ran a “Mainstream” program for monolingual English speakers and students who were considered “proficient,” and a Sheltered English Immersion Program (SEIP) for students who were classified as English language learners by the district. Each grade level was divided into two mainstream classrooms and two SEIP classrooms. The two SEIP classrooms were organized according to the school’s determination of students’ level of English proficiency: *beginner*, *intermediate*, and *transitioning* (to mainstream classrooms). Within the school, 66% of students were classified as ELLs, 76% qualified for Free and Reduced Lunch, 45% of students identified as Hispanic, 31% as White, 13% as Asian, and 9% as African American.

Data were collected from a third grade, SEIP classroom with a mixture of students identified as having beginner and intermediate proficiency. Students represented nine different first languages and students and their families came from eleven different countries including Haiti, Albania, Brazil, El Salvador, and Nepal. Students’ length of residence in the country ranged from US-born, to arriving up to three months before recording the lesson we present below. The episode we analyzed for this paper was part of a five-session unit on Sound. Based on the state’s curricular and language requirements, the teacher and first author planned the unit with the intent of highlighting the vibratory nature of sound and the relation between vibrations and sound characteristics (e.g., pitch and volume). We focused our analysis on the third session, “Making sense of a guitar,” since it seemed to have the richest discussion by students investigating the sounds produced by a guitar-like instrument. Specifically, we selected excerpts that showcased a shift in students’ engagement, as well as sophisticated reasoning.

In an attempt to capture the complexities of ideas and interactions in the classroom, video recordings on four of the sessions described above were collected. In addition to the video, the first author wrote journal entries describing what had taken place during each session, and took notes during the debriefing meetings with the teacher. Videos were transcribed and coded according to the four principles for productive disciplinary engagement proposed by Engle and Conant (2002): (1) problematizing observed phenomenon (further broken down below), (2) display of authorship of ideas, (3) accountability to their local community, and (4) use of material resources. We identified two categories of problematizing observed phenomena that pertained to the sound production mechanisms students referred to when exploring the guitar: *physical features of the producer of sound* (size or length of the string, tension on the string, frequency at which strings were perceived to vibrate) and the *characteristics of the sound produced* (differentiate between low and high pitches). These codes for problematizing observed phenomena, which attend to mechanistic reasoning (Russ, Scherr, Hammer, Mikeska, 2008), were further investigated in terms of how, and for what purposes, students used everyday and academic

language in the service of constructing explanations. Using an inductive coding approach we identified a wide range of uses of language by coding both language and gestures.

This study explores the following research question *what role do everyday and academic language play when emerging bilingual students are making sense of the physical world?* Specifically: (i) what features of a classroom based on scientific induction foster engagement of emerging bilingual students? and (ii) how can this engagement support students' reasoning about mechanisms that drive observable phenomena?

Findings and Analysis

We focus on student participation throughout a discussion on sound to exemplify how attending to physical phenomena can provide emerging bilingual students access to classroom discourse. In this particular episode, students presented their ideas on how the physical properties of the strings of a guitar-like object affected the type of sound produced. As suggested below, students' use of familiar language to describe the different mechanisms that drive differences in pitch seemed to have been instrumental in making discussions accessible to students. Through leveraging everyday terminology, which was sanctioned by students through its usage, students contributed to the group's efforts of making sense of the underlying mechanisms connecting the physical properties of strings and the pitch of the sounds produced.

For this lesson, the students all sat on the carpet in a circle. The session began when the researcher (first author) held the guitar-like object and asked students what they made of it. Gustavo, a Brazilian-American student who was always ready to offer his thinking, stood up and approached the instrument, claiming he knew what it would sound like. The researcher handed the instrument to Gustavo and, while plucking each of the four strings, he aurally imitated the sound he heard: "tick tick, tack tack, tock tock, tuck tuck." Gustavo handed the guitar back to the researcher, and the following exchange ensued:

22. Researcher: Can you - can you tell us again, Gustavo? The sound, the sound. Just make the sound again.
23. Researcher: With your mouth. With your mouth.
24. Gustavo: tick tick, tack tack, tock tock.
25. Researcher: OK, and which ones makes ting ting? This one makes ting ting?
26. Gustavo: No, the first one.
27. Researcher: The first one makes ting ting? And the second one makes...
28. Gustavo: tang tang.
29. Researcher: And the third one?
30. Gustavo: tong tong.

When asked to describe what he heard when plucking the strings, Gustavo offered that the first and shortest guitar string sounded ting ting; the second guitar string sounded tang tang; the third guitar string sounded tong tong; and the fourth and longest guitar string "doesn't make any noise." The group first received this contribution with laughter. This was the first time that any of the students had characterized sound in these terms. In preceding sessions, the teacher used academic terminology when describing sounds produced by vibrations and students were familiar with the use of academic terminology. Soon after Gustavo's remarks, other students appropriated the labels (ting, tang, tong) when talking about the sounds produced by the strings, and the use of "ting tang tong" increased as the discussion progressed. We see the increased and continued use of "ting tang tong" in reference to the sounds made by the instrument as evidence that this everyday language was becoming sanctioned by students, and increasingly became the "formal" class vocabulary for referring to the various strings on the instrument.

Students took up the terms "ting tang tong" and appropriated this language to expand on and communicate scientific ideas. In a sense, these labels became *linguistic footholds* due to their onomatopoeic nature, rooted in students' observations, in conjunction with their connection to the tangible object that could be held, touched, and played by any student in the group. We claim that this invented and/or informal terminology, together with the presence of a tangible object, and the opportunity to negotiate meaning about this observed phenomenon increased the discussion's accessibility. We argue that this connection to students' everyday lives and language fostered students' practice of, and confidence in, authoring explanatory ideas, given students' familiarity with and expertise on sound and musical instruments. Often, students' innate ability and desire of making sense of the world around them is overlooked, missing an opportunity for creating a connection between their everyday lives and science.

The intersection between the formal science (script) and the informal everyday language and experiences (counterscript) is where we see the potential for the emergence of third spaces. In the particular case of the discussion on the guitar, we observed three critical elements leading to the emergence of a third space: (1) a participant, Gustavo, invented these labels, (2) the group appropriated and used these terms frequently, and (3) the teacher allowed this type of (relatively informal) talk to continue. We argue that a third space was generated

and it created opportunities for students to participate. The teacher did not correct Gustavo's language or override these labels with the formal academic terminology. Therefore, space was provided for students to express themselves by integrating everyday, onomatopoeic terms, their experiences, the instrument, and the space to embark on a discussion using the scientific practice of mechanistic reasoning to negotiate meaning of the causes of sound. To further support the explanatory power of the notion of linguistic footholds and third space, we focus the rest of this section on Gyorgy, a Hungarian student who had been in the U.S. for a little over six months by the time this episode was recorded.

Before the start of the discussion, Heidi, the classroom teacher, showed students the guitar she had built. While she was talking, we saw Gyorgy, a typically shy, reserved child, in the background moving his hands and arms as if he were strumming a guitar, and then playing a piano, then tooting a trumpet, and playing the drums. Although his gestures were obvious after watching the video recording several times, they were relatively disguised in the background with respect to the circle of students in the classroom discussion. Gyorgy sat back in the circle, with his back hunched and his shoulders curled in and forward. While Gustavo was presenting his ideas about how each string sounded, Gyorgy sat facing Gustavo and was completely quiet. When Gustavo took the guitar from the center of the circle back with him to his seat on the rug, Gyorgy was within reaching distance and made a couple of quiet attempts to touch and strum the strings. He still did not contribute to the conversation, although his gesturing suggests he must have had previous experience with music. Despite demonstrating knowledge and enthusiasm about musical instruments, Gyorgy did not voice his experience-based ideas. But his silence was about to be broken.

Minutes after "ting tang tong" was introduced, Gyorgy watched Gustavo scratching his chin and raising his hand to ask for a turn to speak. Gyorgy then appeared to mimic Gustavo's behaviors by scratching his chin and almost raising his hand (in a non-exaggerated, somewhat timid way). It seemed that he also wanted to contribute to the conversation and was impressed by the ease with which Gustavo did so. When called on by the teacher, Gyorgy's facial expression showed what could be perceived as anxiety over having to speak up, and his body slouched forward and kind of curled up as he began shrinking backward away from the circle. When the researcher asked Gyorgy if he would like to try explaining his thinking, Gustavo immediately shot his hand up, to which the teachers replied by asking him to let Gyorgy try. Once again, Gyorgy's face appeared anxious, but this time he looked down, stared at the floor, and shook his head and timidly gestured, "no." Heidi, intervened:

- 70. Heidi: Gyorgy, do you want him to repeat the question? Or do you understand the question?
- 71. Heidi: Use your English. You're doing great. You've been doing really great.
- 72. Gyorgy: No, no.
- 73. Heidi: Yes, yes, you have.

Once again, Gyorgy looked down, avoiding eye contact with Heidi, and remained in the far edge of the circle and away from the guitar for about five seconds. The teacher continued to wait. After that wait time Gyorgy took a deep breath and said quietly at first, "I have a question," and approached the center of the circle where the guitar sat on the floor. Gyorgy moved forward, kneeled and looked at the guitar from above, paused for a five seconds, and said:

- 76. Gyorgy: This [first string], ting ting. And this, tang tang [second string]. And this one [third string], Gustavo?
- 77. Gustavo: tong tong.
- 78. Gyorgy: Because this [first string] not really – (pointing with to the first and third strings)
- 79. Gustavo: Short.
- 80. Gyorgy: And this [first string] small, and big (pointing to second string), and bigger (pointing to third string).
- 81. Researcher: Small, big, and bigger? So, small is -
- 82. (Gustavo: Medium.)
- 83. Researcher: ting ting (Gyorgy shakes head in approval)? And medium is tang tang (Gyorgy shakes head in approval)? And bigger is tong tong (Gyorgy shakes head in approval)?
- 84. Gustavo: Can I say why?
- 85. Heidi: And why?
- 86. Researcher: Yeah. Why do you think is that, Gyorgy?
- (20 seconds of silence)
- 90. Researcher: Are you done, Gyorgy? Do you want to say anything else?
- 91. Gyorgy: I don't know what to say, because - I'm done

92. Researcher: OK. Thank you, Gyorgy. That was really good. Thank you.
93. Heidi: Nice job, Gyorgy
- ...
104. Researcher: Gisela, do you wanna try? What do you think?
105. Gisela: Because, when you like "eeeeeeeee," and it goes high, and that (first) string goes high. And when we go "aaaaaaa," that (third) string goes low.
106. Researcher: OK. So, that's why we say it's high pitched and low pitched? It sounds very similar? OK.
107. Gustavo: I know why.
(Gyorgy raises his hand)
108. Researcher: Ok. Gyorgy?
109. Gyorgy: Is because this [first string] is very fast, is very fast; and this [second string] not very fast; and this [third string] is slow.
110. Researcher: Oh.
111. Gyorgy: See? This [first string] very fast - (Gisela: Like the ruler.) and this [first string] hard. And this [second string] not hard. And this [third string], this easy.

After being encouraged by Heidi, Gyorgy's demeanor changed: he approached the instrument and presented his ideas to the group. He no longer appeared shy or anxious, and instead lunged himself forward almost completely dominating the space around the guitar. Additionally, when offering an explanation, Gyorgy appeared comfortable pointing to the specific string he was thinking about, while looking at his peers and teachers. Gyorgy seemed to have accepted the invitation of becoming part of the discussion, appearing in complete control of the situation. It is important to remark that, this exchange was the first time the onomatopoeic labels were used by a student other than Gustavo, and marked the start of a formalization process. As the discussion progressed, Gyorgy became more vocal and shared his ideas with more ease and confidence.

The excerpt above also shows Gyorgy engaged in the scientific practice of mechanistic reasoning. Particularly, he pointed out possible mechanisms that produced the different sounds, "this (first string), ting ting, and this (second string), tong tong. And (first string) small, and (second string), big and (third string) bigger" (lines 76 and 80). Engaging in mechanistic reasoning (Russ et al., 2008) is one of the fundamental practices scientists engage in when developing and evaluating explanatory models, and should be central to the teaching and learning of science. Moreover, Gyorgy used the labels Gustavo had invented and also checked with the author to see if he was using the correct invented terminology (line 76). We interpret this remark as denoting the connection Gyorgy was making between the physical features of the strings and the characteristic of the sounds produced. Specifically, here we see that he, like Gustavo, used the onomatopoeia to talk about *the relationship* between length and pitch, e.g., "the ting ting string is small." The second connection Gyorgy made was between the frequency at which the string vibrated and pitch, when trying to understand why the short string sounded "eeeeee," "[first string] is very fast" (line 109). He quickly added that the first string was also hard (line 111), which we interpret as Gyorgy relating the tension he felt on the string and the sound's pitch. We argue that implicit in this discussion was Gyorgy's experience with instruments and his innate drive to make sense of the world around him.

Gyorgy's participation throughout the discussion illustrates the emergence and maintenance of a third space that allowed students to leverage elements of their everyday language and experiences in service of exploring physical mechanistic processes. Specifically, the creation and sanctioning of linguistic footholds made the discussion about the guitar accessible to Gyorgy. We argue that for him "ting tang tong" served as the foundation through which he would externalize his thinking, and from which he would continue to expand his understanding of sound. These experiential labels had great information about the system codified into them, which in this case is the pitch of the sound produced by each of the string. Therefore, when referring to the "tong tong" string, for example, students are talking about a particular string within the guitar (position) and their expectation of sound characteristic (pitch). Moreover, we claim that the creation of onomatopoeic labels that derive from observations signaled students establishing a bidirectional connection between experiences and abstract ideas (Vygotsky, 1962). The invention of a label is a step towards exteriorizing and objectifying ideas, which plays an important role in constructing disciplinary knowledge. Moreover, the onomatopoeic nature of "ting tang tong" reified the shared experience so that, even though it was no longer present, students could now elaborate with abstract and mechanistic reasoning. In general, we recognized that students often used non-academic language when describing their ideas about sound, which eventually became sanctioned by the rest of the group through its widespread use.

Additionally [in addition to?] to the richness embedded in and symbolized by the linguistic footholds, we assert that they fostered Gyorgy's *productive disciplinary engagement* (Engle & Conant, 2002). The labels seemed to have served the important function of helping Gyorgy publicly problematize the physical processes in

the guitar, by on the causal mechanisms that connected the physical properties of the string and the pitch of the sound he heard. Rather than just stating the fact that the first string was shorter, Gyorgy tried to make an explicit connection between the length, tension, and vibration frequency of the string and the pitch of the sound he heard. And we surmise that his reasoning about mechanisms contributed to the group's understanding of the guitar, given that he was trying to establish a causal link between the cause (length) and the effect (pitch) of sound production.

Once Gyorgy's participation in authoring ideas increased, it appeared as if his confidence and position also changed. At first, despite demonstrating knowledge about musical instruments, he refrained from making contributions to the conversation. Once the teacher called on and encouraged Gyorgy, he appropriated the vocabulary created by Gustavo and used it for communicating his thinking. Subsequent remarks seemed to be related to his own thought process, rather than revoicing other students' comments. Again, we see this process as evidence that Gyorgy became an author of ideas in this space, rather than a consumer, while remaining accountable to the local community by continuing to use formalized vocabulary and supporting claims with evidence. Finally, we would like to comment on the importance the presence of the tangible object had on the conversation. It was by plucking the strings that Gustavo generated the onomatopoeic labels that served as linguistic footholds for Gyorgy and the other students. Moreover, Gyorgy's contributions appeared to be significantly tied to his experience with the physical object, e.g., through plucking the strings for measuring the tension and frequency of vibrations, and pointing to the string he was referring to. Therefore, we argue that the emergent third space supported Gyorgy's productive disciplinary engagement in the discussion about the connection between the strings' features and pitch.

There are some limitations to this study. While our claim is that learning environments that support the emergence of third spaces invite students to participate, we recognize that not all students are able to do so. Our choices for data collection only give us information on students' productive disciplinary engagement as they publically engaged each other, and it leaves out how students could have been partially engaged in a private way.

It would have been beneficial to ask students to produce individual written text on their ideas about the guitar before the discussion. This would have given us a sense of students' previous experiences with guitars and/or intuitions on the relation between the physical properties of the strings and the characteristics of the sounds produced. Additionally, we acknowledge the range of social dynamics in play during discussions that can promote or hinder certain students' willingness to engage. In this case, the written text could have served as an avenue for students to express themselves without feeling the pressure to participate in a public forum. Finally, although we have made some inferences about the aspects of the learning environment that led to productive disciplinary engagement, it is not fully clear what combinations of features were the critical ones.

Conclusions and Implications

Based on the evidence presented above, we claim that central features of the learning environment that provided access and inclusion in classroom discourse were: (i) tangible nature of class materials; (ii) shared nature of experiential observations; (iii) students' inherent desire to understand their worlds; and (iv) the coexistence and blending of everyday and academic language. We argue that these central features of the learning environment, also central features of scientific inquiry, facilitated and mediated a rich discussion between students about mechanisms, furthering conceptual understanding and belonging. Additionally, the presence of familiar language in these conversations gave access to students who may have felt unsure about their perceived level of understanding and/or language skills. Moreover, students responding to their peers' conjectures, and co-constructing knowledge, is evidence how third spaces distribute authority of knowledge and language among students themselves. Also, the discussion provided evidence of how common and invented terminology can become formalized, in a process resembling the creation of scientific discourse conventions. As it is hypothesized in third spaces, students became comfortable using the invented terminology, allowing them to express themselves and construct knowledge freely, and even to test academic language that was introduced through schooling. Moreover, these emergent third spaces provided opportunities for student to engage in scientific cultural practices, specifically by problematizing phenomena through attending to causal mechanisms, supporting their claims with evidence from observations, and sustaining a group discussion that contributed to the negotiating of meaning and conceptual understanding.

References

- Committee on Conceptual Framework for the New K-12 Science Education Standards, N. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (p. 320). Washington, DC
- 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.

- Escamilla, K. & Hopewell, S. (2010). Transitions to biliteracy: Creating positive trajectories for emerging bilinguals in the United States. In J. Petrovic (ed.). *International Perspectives on Bilingual Education: Policy, Practice and Controversy*, Information Age Publishing, 69-94.
- Gutierrez, K. D., Baquedano-López, P., and Turner, M. G. (1997). Putting Language Back into Language Arts: When the Radical Middle Meets the Third Space. *Language Arts*, 74(5), 1997.
- Kindler, 2002 A. L. Kindler. (2002). *U.S Dept of Ed. OELA*, Survey of the states' limited English proficient students and available educational programs and services: 2000–2001.
- McMurrer, 2008 McMurrer, J. (2008). *Instructional Time in Elementary Schools: A Closer Look at Changes for Specific Subjects* (p. 8). Washington, DC.
- Otero, V. K., & Nathan, M. J. (2008). Preservice elementary teachers' views of their students' prior knowledge of science. *Journal of Research in Science Teaching*, 45(4), 497-523.
- Rogoff, B. (1994). Developing understanding of the idea of communities of learners. *Mind, culture, and activity*, 1(4), 209-229.
- Rosebery, A. S., Ogonowski, M., DiSchino, M., & Warren, B. (2010). “The coat traps all your body heat”: Heterogeneity as Fundamental to Learning. *the journal of the learning sciences*, 19(3), 322-357.
- Razfar, A., Licón Khisty, L., & Chval, K. (2011). Re-mediating second language acquisition: A sociocultural perspective for language development. *Mind, Culture, and Activity*, 18(3), 195-215.
- Russ, R. S., Scherr, R. E., Hammer, D., and 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.
- Suárez, E., & Otero, V. (2013) “Because it hibernates faster”: 3rd grade English Language Learners making sense of sound. *Physics Education Research Conference Proceedings 2012*, 406-410
- Suárez, E., & Otero, V. (in press). Physics as a Mechanism for Including ELLs in Classroom Discourse. *Physics Education Research Conference Proceedings 2013*
- Vygotsky, L. *Thought and Language*, MIT Press, Cambridge, 1962.
- 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.
- Warren, B., Ogonowski, M., and Pothier, S. (2005). “Everyday” and “Scientific”: Rethinking Dichotomies in Modes of Thinking in Science Learning. In R. Nemirovsky, A. Rosebery, J. Solomon, and B. Warren (Eds.), *Everyday Matters in Science and Mathematics: Studies of Complex Classroom Events* (First., pp. 119–149). Mahwah, NJ: Lawrence Erlbaum Associates
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge university press.

Using Contrasting Video Cases of the Enactment of Cognitively Demanding Science Tasks in Professional Development

Miray Tekkumru Kisa, Mary Kay Stein, University of Pittsburgh, Learning Research and Development Center
Email: mit34@pitt.edu, mkstein@pitt.edu

Abstract: Prior research indicated the challenges of getting students to think at high levels as they work on cognitively demanding tasks. Teachers often unwittingly lower the cognitive demands of tasks during their enactment. We argue that teachers should be provided with opportunities that can support them to make sense of the ways in which teachers' actions can be consequential for the level and type of thinking in which students engage. In this paper, we examined how the use of contrasting video cases can support participants' learning to identify the ways in which teachers' actions shape students' opportunities to think and reason during their engagement with cognitively complex instructional tasks. Qualitative analysis of two sequential professional development sessions revealed that contrasting video cases supported participants' productive discussions about teacher's pedagogical actions and the ways that these actions shaped how students work on cognitively demanding tasks.

Introduction

The new rigorous standards for science education in the United States set the stage for students' thinking and reasoning about disciplinary ideas and their engagement in scientific practices (Next Generation Science Standards, 2013). For these standards to be realized, however, the nature of the classroom work needs to become more cognitively demanding; teachers need to select and effectively implement high-level instructional tasks that will position students as active agents in generating ideas, developing explanations, and designing models and investigations. However many studies in science and mathematics demonstrate the challenges of getting students to think and reason at a high level as they work on such cognitively demanding tasks (Blumenfeld et al., 1991; Blumenfeld & Meece, 1988; Doyle 1983; Stein, Grover, & Henningsen, 1996).

Prior research indicates that even when cognitively demanding tasks are selected, the level and type of student thinking that they require often declines during their enactment (Doyle 1983; Henningsen & Stein, 1997; Sanford, 1987; Stein et al., 1996). Teachers often *unwittingly* lower the cognitive demands of tasks by focusing on the correctness and completeness of procedures and answers rather than how students are thinking about the task and/or "taking over" the thinking and actually doing the work for the students (Doyle 1983; Henningsen & Stein, 1997; Stein et al., 1996). During the enactment of project-based activities in science classrooms—activities that are often complex and ambiguous (Blumenfeld et al., 1991)—teachers encounter many challenges that may lead to reducing high-level student thinking such as reluctance to provide autonomy to their students, allocating enough time for in-depth exploration of ideas, and being able to provide appropriate amount of scaffolding (Blumenfeld et al., 1991; Marx, Blumenfeld, Krajcik, & Soloway, 1997; Marx et al., 1994). All of these suggest that the way teachers facilitate students' engagement in cognitively demanding tasks shapes the way students are positioned to think and reason about the disciplinary ideas and engage in the practices of the discipline in the classroom.

Given the critical role teachers play in shaping students' opportunities for learning, we claim that teachers should be provided with opportunities that can support them to think about and make sense of the ways in which teachers' actions could be consequential for the level and type of student thinking that occurs as students engage with challenging tasks. It is particularly important to help them realize that selecting high-level tasks does not always guarantee high-level student thinking in their classrooms. This study focuses on high-school biology teachers' learning within a video-based professional development (PD) about the role of teachers' pedagogical actions in facilitating students' learning during the enactment of cognitively demanding science tasks.

Video use has increased in studies designed to support teachers' learning within the last decade (e.g., Star & Strickland, 2008; van Es & Sherin, 2006). When used as a tool for supporting teachers' learning, the structures and tasks designed around the video should be considered carefully to get the most benefit from what the video can offer (Le Fevre, 2003; van Es & Sherin, 2009). As part of our PD intervention, we made careful design choices about the activities in which we asked teachers, who participated in PD sessions, to engage surrounding the video cases. Drawing on prior research about contrasting cases (e.g. Bransford & Schwartz, 1999; Garner, 1974), we used *contrasting video cases* of enactment of the same high-level task to help PD-participants recognize finer distinctions in teachers' pedagogical actions between classrooms in which the cognitive demand of high-level tasks was maintained versus declined: One of the contrasting video cases represented high-level student thinking, the other one illustrated low-level student thinking during the enactment of the *same* cognitively demanding science task in two different high-school biology classrooms.

Even though video cases of instruction have been widely used to support teachers' learning, no study has yet used contrasting video cases to support science teachers' learning within a professional development setting. We argue that contrasting video cases could be a productive tool in professional development for helping participants to identify specific details in teacher's pedagogical actions and the role that these actions play in students' thinking. Thus, in this study we examined the differences in the nature of PD participants' pedagogy related comments between two sequential PD sessions: one involving participants' discussion about a *single* video case and the other one about *two contrasting* video cases. Our study findings address the following research question: How, if at all, do contrasting video cases in a PD support participants' learning to identify teacher's pedagogical actions and the role they play shaping the level and type of student thinking?

Theoretical Framework

Why Focus on the Enactment of Cognitively Demanding Tasks?

Prior research shows that interactions between the teacher and students play an important role in shaping whether the cognitive demand of tasks is maintained or declines as students work on the task (Doyle, 1988; Henningsen & Stein, 1997). As underscored by Doyle (1988), teacher's role is critical in shaping the level and kind of opportunities that students have for thinking in the classroom; he stated, "teachers affect tasks, and thus students' learning, by defining and structuring the work students do" (p.169). Teacher-related factors, which were identified in prior research as associated with declining cognitive demand, include teacher's "taking over" difficult pieces of the task and telling students how to do the task, and shifting focus from understanding to completeness or accuracy of answers. Teacher-related factors associated with maintaining high-level student thinking, on the other hand, includes scaffolding of student thinking, sufficient time for exploration, providing students with means of monitoring their own progress, and sustained press for meaning through teacher questioning and comments (Stein, Smith, Henningsen, & Silver, 2000).

Why Video Cases?

Using classroom-based artifacts in PD sessions such as copies of students' work, videotapes of classroom lessons, and curriculum materials helps to situate PD in the context of teaching (Ball & Cohen, 1999; Borko and Koelner, 2008). Prior research indicates that videos are powerful in capturing the richness and complexity of classroom instruction; they provide a more realistic picture of the learning environment by capturing the voices, body language, and interactions of classrooms (Borko, Jacobs, Eiteljorg, & Pittman, 2008; Colestock & Sherin, 2009; Koc, Peker, and Osmanoglu, 2009; Le Fevre, 2003; Miller & Zhou, 2007). Moreover, viewing videos of instruction can help teachers to notice problems in their own teaching and become more willing to accept that certain aspects of their teaching need to change (Tripp & Rich, 2012). Given these benefits of videos, we used *contrasting video cases* to support teachers' learning about the enactment of cognitively demanding tasks.

We used Task Analysis Guide in Science (TAGS) (see Figure 1) to identify the cognitively demanding tasks whose enactments were shown in the video cases as well as to determine the level and type of student thinking in these video cases. The TAGS is a two-dimensional (i.e. cognitive demand and integration) framework involving nine categories each of which represents different levels of student thinking demanded by science tasks (i.e. cognitive demand) that focus on either science content or scientific practices or the integration of the two. Science tasks require different cognitive demand levels by making students to reproduce previously provided knowledge or follow scripted procedures meaninglessly to get to a right answer, or by guiding them to make sense of the scientific ideas and/or practices, or requiring them to think like a scientist (For details, please see Tekkumru Kisa, Stein & Schunn, under review).

	Scientific Practices (e.g. argumentation and investigation)	Science Content (i.e. scientific body of knowledge)	Integration of Content and Practices
5 DOING SCIENCE TASKS			Doing Science (DS) Engaging in practices to make sense of content and recognize how scientific body of knowledge is developed.
4 TASKS INVOLVING GUIDANCE FOR UNDERSTANDING	Guided Practice (GP) Guided work with practices	Guided Content (GC) Being guided for understanding a particular content	Guided Integration (GI) Guidance for working with practices tied to a particular content
3			
2 TASKS INVOLVING SCRIPTS	Scripted Practice (SP) Following a script to work on practices	Scripted Content (SC) Following a script about a content	Scripted Integration (SI) Following a script to work on practices tied to content
1 MEMORIZATION TASKS	Memorized Practice (MP) Reproducing definitions/ explanations of practices	Memorized Content (MC) Reproducing definitions, formulas, or principles about a particular content	

Figure 1. The Task Analysis Guide in Science (Tekkumru Kisa et al., under review).

Why Contrasting Video Cases?

Prior research suggests that contrasting cases help to make particular aspects and dimensions of cases become more salient and differentiated from others. Emphasizing the role of contrasting cases, Garner (1974) claimed:

The single stimulus has no meaning except in a context of alternatives. When somebody uses the term circle, they infer that it could have been some other form, such as square or triangle. When somebody says there are two circles, or that the circle has two lines, they inferred the alternative of fewer or more lines. Each descriptive term used defines what the alternatives are, by defining what the stimulus is not (p. 185).

Others have also emphasized the importance of contrasts in allowing people to notice distinctive features that they may miss in the absence of a contrast set (Bransford et al., 1989; Bransford & Schwartz, 1999). Drawing on this body of research, we used contrasting video cases of task enactment. We examine the claim that contrasting video cases can be used to surface how different forms of teacher facilitation can be consequential for differences in the level and type of student thinking during students' work on the same high-level task.

Research Design

This study was situated within a larger NSF-funded project that focused on the development and implementation of design-based STEM units that aim to teach big ideas in biology tied to mathematics. It took place during the implementation of one of these units named, "Modeling Genetics: The Gecko Breeder Challenge". All the video clips that were discussed in the PD were selected from the video-records of classrooms during an earlier implementation of this unit. Teachers who participated in the study agreed to implement the four-week long unit, attend two project-related meetings and seven PD sessions, which were specifically designed for this study. In this paper, we focused on two of these seven PD sessions.

Intervention: Video-Based Professional Development

The PD sessions took place once or twice a week from the first week of February 2012 to the first week of March 2012 for a total of seven sessions (each was about 3 hours in duration), about half of which was allocated to viewing and discussing the video cases. In the first session, PD participants were introduced to the TAGS. In the next two sessions, they analyzed a video case that illustrated low-level student thinking (in session-3) and high-level student thinking (in session-4) during the enactment of two different cognitively demanding biology tasks. The fourth and the fifth PD sessions involved the use of contrasting video cases. Finally, in the last two sessions, participants analyzed video cases from their own classrooms.

Because the purpose of this article is to understand the role of contrasting video cases, we focused on two sequential PD sessions, which involved discussion about a single video case (session-3) and contrasting video cases (session-4) (see Table 1 for details about the sessions). **Session-3** involved the first video case in the entire PD in which participants viewed high-level student thinking during the enactment of a cognitively demanding biology task. In this task, students were provided with the PCR results that showed the variation in DNA for the same gene. PCR results showed two separate crosses of a male and a female gecko and their offspring. The task required students to generate the rules of inheritance (i.e. how the genetic information is transferred from parent to offspring) by analyzing the PCR results. In the video case, students in small groups were sharing their observations and interpretations of the PCR results with the teacher. The level of student thinking was at level-4, guided integration based on the TAGS.

Table 1: PD sessions 3 and 4: Activities and design rationales.

PD Session-3 Activities:

1. Analyzing a science task (as it appeared in print materials) based on its potential cognitive demand
2. Viewing and discussing "set-up" video showing how above task was introduced to students
3. *Viewing and discussing "enactment" video showing high-level student thinking & identifying the level and type of student thinking in the video case*

Design Rationale:

* Providing opportunity for PD participants to identify the level and type of student thinking in all three phases (print materials, set up, and enactment)

PD Session-4 Activities:

1. Analyzing print-based tasks based on their potential cognitive demand
2. Summarizing participants' analysis of print-based tasks and video cases in PD sessions 2 and 3
Presenting the change in cognitive demand across the phases of print-based materials, set up and enactment
3. Viewing the first video case showing low-level student thinking during the enactment of a high-level task
4. Viewing the second video case (contrast) showing high-level thinking during the enactment of the same task

Discussing (a) similarities and differences between the video cases and (b) identifying the level and type of student thinking in the video cases

Design Rationales:

- * Providing opportunity for PD participants to see that cognitive demand of a task can be maintained or decline
- * Using contrasting cases to learn to distinguish levels of student thinking as represented in the TAGS
- * Using contrasting cases to support judgments of maintenance or decline
- * Using contrasting cases to reveal pedagogical factors associated with maintenance or decline of the cognitive demand of the task

As seen in Table 1, in **session-4** before viewing the video cases, participants were first introduced to the key idea that tasks can change in their level of cognitive demand as they pass from print materials to how they are set up by the teacher in the classroom to how they are actually enacted or carried out by the students (Stein et al., 1996). This presentation helped to frame participants' viewing of and discussing *together* the contrasting video cases. The cognitively demanding task showed in the contrasting cases was about modeling how genetic information is inherited. In this task, students were guided through designing a paper-based simulation of gecko breeding that helped them to understand how the offspring gets one allele per gene from each parent and parental alleles are "packaged" inside eggs and sperm. In both video cases, students worked in small groups with the simulation materials and the teacher walked around and helped students. Video cases were designed to show that students were required to engage in scientific practices (e.g. modeling) in both of the classrooms but only in the second classroom they were required to make sense of the scientific ideas (e.g. how alleles are packaged into gametes) while designing the simulation. Therefore, how students were positioned to engage in this simulation task was very different in these classrooms. In the first video case, student thinking was at level-2, scripted integration according to the TAGS. Students followed the procedures of the simulation without really understanding the underlying content, and they engaged in a set of actions needed to complete the simulation because they were told to do so, mostly by the teacher. The opposite was the case in the second video in which student thinking was at level-4, guided integration based on the TAGS. Students were challenged to make sense of inheritance patterns explained in the simulation. They were asked to justify the simulation procedures by using what they learned about Mendelian Inheritance.

Participants of the Study

Five high school biology teachers from several different school districts participated in the PD. These teachers, who voluntarily participated in the study, were paid for their participation. Linda and Susan were the two most experienced teachers in the PD with 16 and 13 years of teaching experience. Linda was the only teacher in the PD who had prior experience in implementing the Genetics unit mentioned above. Barbara and Nancy were from two different schools operating under the same charter school organization, which focused on the use of project-based practices in the classroom. They both had three years of teaching experience. Lastly, Carol, with five years of teaching experience, was from a private school. Like Linda, Carol had experience in working with the project team, but on a different science unit.

Data Sources and Analysis

In our analysis, we focused on the parts of the PD transcripts during which teachers discussed the video case(s) and interpreted the level and type of student thinking. The unit of analysis was a segment of transcript in which one or more PD participants commented on a particular pedagogical action in the video. Therefore, each unit was separated from the next one in terms of the pedagogical action that was discussed. For example, while in one unit participants' comments were about the nature of questions used by the teacher in the video case, the next unit involved comments about how the teacher focused students' attention on the procedures of the task. For each unit, we 1) paraphrased the pedagogical action discussed by the participants, identified: 2) whether pedagogical action was grounded in students' thinking about the task, 3) whether any reference was made to the level and type of student thinking, and 4) whether the discussion was influenced by the facilitator's prompting.

Results

Our findings revealed interesting differences in the nature of PD participants' pedagogy-related comments between session-3 and session-4. Overall, in session-3, unless prompted by the facilitator, participants made general observations about the teacher's pedagogical actions. Their overall conclusion regarding the level of student thinking was that students were positioned to *do the thinking* in the classroom. In session-4, in contrast, participants—independent of the facilitator—began to make pedagogical remarks, which, were grounded in students' thinking about the task. Their comments involved characterization of how students were positioned to *think and reason at different levels of cognitive demand* (as defined in the TAGS) in the two different classrooms that used the same biology task.

These overall differences between the sessions were as we expected because a single video case presents a particular level of student thinking, which can be categorized based on the TAGS as Guided Integration (focusing on the integration of science content and scientific practices and at a cognitive demand level-4). Noticing the *salient* features about the teacher's pedagogical actions and their association with the level and kind of student thinking depicted in that single video case depends on the expertise of the PD participant that analyzes the video case. Contrasting video cases, on the other hand, presents *variations* in the level and kind of student thinking across two classrooms, which helps participants to notice more easily and precisely the differences between how students were positioned to think at different levels in these two classrooms as a result of the differences in the teachers' pedagogical actions. In what follows, we will discuss in more detail the nature of participants' discussions about the teacher's pedagogical actions and the role that these actions play in how students were positioned to think in the two classrooms.

The Nature of Discussion in Session-3: A Single Video Case

A little more than half of participants' pedagogy-related comments in session-3 (57%) were at a general level. With that we mean participants were talking about the teacher's actions without a clear reference to students' work on the task (e.g. what students think about the concepts covered in the task and how the teacher's pedagogical moves were influenced by what students do or say surrounding the task). For example, the following conversation illustrates describing the nature of teaching at a very general level:

Carol: And she [the teacher] didn't tell them anything.

...

Carol: Everything was a statement.

...

Nancy: Yeah. She [the teacher] didn't guide them. It was all questions.

As demonstrated in this conversation, such general comments were only about what the teacher did (or did not do) and they were independent of what students seemed to be doing at that time.

Some of the pedagogical actions that were discussed by the participants at a general level in session-3 include "teacher did not tell what was right or wrong"; "teacher did not transmit knowledge" and "teacher did not lead students to any conclusion". For example, the following exchange between Carol and Susan illustrates their observation about the teacher's pedagogical actions in the video case:

Carol: When they [students] say something, she'll [the teacher] say, "Oh, why?" like [line] 121 [in the transcript] "Oh, so why are you saying this band is thick?"

Susan: Yeah. But then ... she [the teacher] further questions them [students] about what they're saying. So like [line] 124 "Are you telling me the male has only one gene for this trait?" So she's just trying to get them [students] to clarify what they're saying.

Carol and Susan noticed that the teacher in the video asked the students clarification questions. They provided concrete examples from the transcript to show that the teacher asked questions to the students to clarify what they said. They did not, however, continue and discuss specifically *how* these clarification questions influenced or got influenced by students' thinking about the task. For example, the teacher's question, "Are you telling me the male has only one gene for this trait?" might have made the student to think further and understand that each organism has two alleles per gene.

The discussion in session-3 about the teacher's pedagogical actions was not always at a general level. Our analysis revealed that nearly half of the participants' pedagogy-related comments were grounded in students' thinking (43%). In other words, there was a clear reference in participants' comments to *how* the teacher's actions influenced or got influenced by what students said or did while working on the task. In all of these types of comments, though, the facilitator's prompting was influential for grounding participants' comments in students' thinking about the task. For example, one of the participants pointed out that the teacher in the video suggested that students use another resource (a previous task that they worked on) to scaffold their work on the task. Based on this comment, the facilitator asked why the teacher might have suggested students to use this resource at that time but not before. In response to that Linda said:

I think she [the teacher] did it because this whole group is focusing on the word dominant, dominant blizzard, dominant traits, ... and I think she wants them to maybe look back at that phenotype and see that the blizzard's not the dominant, because the females are normal color... So they're linking that thick band to the word dominant and the dominant to the phenotype blizzard. So rather than telling them that they were wrong ... Let's just see if you can support that with the phenotypes, is where I think she would have been leading them.

As illustrated in this comment, Linda made an explicit reference to students' thinking about the task and how the teacher scaffold students to think differently about the DNA data from the mating of two geckos instead of telling them that their interpretation of the data was not accurate.

In session-3, when participants made comments about the teacher's pedagogical actions, they often did not discuss the consequences of these actions for *how* students were thinking as represented in different levels of the TAGS; rather their remarks were more about that students were *positioned to think*. There was an overall agreement among the participants that students in the video case were doing the thinking. For example, Barbara said, "I felt like she [the teacher] really didn't lead them [students] to any conclusions, which I thought was nice, because they were really coming up with it on their own". Similarly, Carol said, "She made them think different and explain or use different words. It was like massaging their brains". These types of comments by the participants generally suggested that the teacher made students "do the "thinking". Participants seemed to consider thinking (in a general sense) as a key aspect of higher-level cognition. They did not, however, often talk about a *certain* type and level of student thinking even though they could accurately identify that video case represented high-level student thinking.

The Nature of Discussion in Session-4: Contrasting Video Cases

In session-4, participants' comments about the teacher's pedagogical actions became more grounded in students' thinking about the task (1). Participants were often talking about the teacher's actions in relation to what students were thinking and saying about the task (57%). In contrast to session-3, about all of such pedagogy related remarks (except one) were made without the facilitators' prompting. The following comment by Linda illustrates such pedagogy related comments grounded in students thinking:

- Linda: The very first group did fertilization and didn't have alleles in their egg and sperm. She [the teacher] didn't make them pull them apart. She made them add the alleles. ... It went against that biological concept of the alleles don't combine. So she didn't do the meiosis part. Second piece of evidence was once she started doing this and these with all those groups, she just kept saying, "How many combinations did you get? ... It was all about the combinations for her, how many combinations –
- Co-Facilitator: Combinations not attached to the meaning
- Linda: The biology
- Co-Facilitator: The biological meaning
- Linda: The meiosis that made them, the fertilization that's happening.

As seen in the excerpt, Linda's remark is targeted on how students' were being encouraged to make sense of the task (or not). Comments like this were common across other remarks made by the participants; there was a close attention to what students did or did not understand about the task as a result of the teacher's facilitation. For example, Susan commented on how the teacher in the second video case helped students grapple with their uncertainty around constructing the simulation of inheritance since they could not apply their prior knowledge easily to what they were working on. Susan said:

... in the last group that we watched, which is like line 46 there, 48 [in the transcript]: [The student says:] "I know there are two alleles for each gene, so, I mean, is it separate two?" And so she [the teacher] says, "Right, it could be black or white." But she's just helping clarify. But then she lets them [students] go through that whole process of putting two and two [alleles to the gametes]. She doesn't tell them, "Just go ahead and try it." ... [So] they're [students] still having to work through it and figure out why they were wrong, whereas I feel like the first teacher would have just said, "No"...

With this comment, Susan clarifies how the teacher in the second video case facilitated students' making sense of the inheritance idea that there should be one allele per gene in each gamete as students are working on modeling how the genetic information is inherited from parent geckos to their offspring.

Participants' comments also revealed some patterns in terms of how students were positioned to think about the task through the teacher's facilitation in each video case. For example, regarding the first video case several of them said that students were focused on the procedures of the given task (e.g., creating the combinations) that had no connection to any biological sense making (e.g. fertilization, having alleles in egg and sperm). This observation by the participants reflects an aspect of students thinking at level-2 (scripted integration) according to the TAGS. Therefore, in session-4 participants' comments had implications for the characteristics of *particular* levels and types of student thinking (often as defined in the TAGS framework). For example, Carol said, "she [the teacher] removed the biology from it [the simulation about genetics]. At the end, they [students] weren't even making geckos [the animal used in the simulation task]. They were making

combinations.” As revealed with this example, teacher-related factors that were discussed by the participants involved some reference to the consequences of the teacher’s pedagogical actions in terms of *particular* kind and level of student thinking, which is more a salient feature of student thinking than only stating that the students were thinking.

Session-4 ended with a detailed discussion about the level and type of students thinking that was going on in the video cases as represented in the TAGS. There was not an agreement among the participants in terms of which levels of TAGS each video case represented even though they all agreed that video-1 represented low-level student thinking and video-2 represented high-level student thinking. Some of this disagreement was because participants were not clear about the distinction between all different levels in the TAGS framework (e.g. what is the difference between memorized practice at level 1 and scripted integration at level 2). The discussion helped to resolve some of these confusions that teachers had.

It is important to remember that, in the contrasting video cases, the same high-level task was enacted in the classrooms. This may have helped PD participants think more deeply about how students’ interaction with the same task (i.e. how they were positioned to think) could be different in two different classrooms. Even though participants were told that the video cases depicted the enactment of the same task in two different classrooms, one of them said towards the middle of the discussion, “I think similarities-wise, it seemed like the kids got the same materials. They were working with the same stuff. And I don’t know, but it also seemed like they had the same handout of directions.” This comment was critical because it reminded them that they were focusing on how the teacher’s actions shaped students’ thinking about the ideas in the same task. Following up on this comment, the facilitator told participants, “I see like there are three variables ... it is the same *task*.... There are *students* asking the same questions and now the *teacher* responding back to them in a very different way, which leads to a different conversation among them,” thereby bringing the participants’ attention to how the interaction of the teacher with the students surrounding the same task could create different opportunities for students’ learning.

Discussion and Conclusions

Teacher learning has become a focus in the learning sciences (Fishman & Davis, 2006), and there has been a growing interest in designing new environments to support teachers’ learning. In this paper, we described findings based on one such effort to support teachers’ learning about the role they can play to maintain high cognitive demand levels in their classrooms. Specifically, our paper focused on the role of contrasting video-cases of the enactment of cognitively demanding tasks in revealing the distinctive features of teacher’s pedagogical actions and the role that these actions plays in students’ thinking. The results are promising because they indicated that in comparison to session-3 participants’ talk about the teacher’s pedagogical actions got more grounded in students’ thinking in session-4 and participants began to see more details regarding the level and kind of student thinking that students were positioned to engage in. This all suggests that contrasting video cases may be a productive design feature of video-based PD for supporting teachers’ learning about salient features of classroom interactions.

It’s important to be cautious about the generalizability of the findings regarding the use and study of contrasting video cases. Even though we used high-level tasks in both of the sessions, the tasks used in session-3 and in session-4 were different from each other. The nature of the task might have created differences in terms of what can be “visible” in the contrasting video cases, which showed the enactment that task in two separate classrooms. However, if we showed the enactment of the same task with a single video case in session-3 to keep the task constant across the PD sessions, then exposing teachers to the task in session-3 would be a confounding variable while interpreting the nature of participants’ comments in session-4; by knowing the task from the previous session, participants could attend to more details in the video cases in session-4. As a result, more research is needed to examine the potential of contrasting video cases revealing the salient features of classroom interactions that teachers should notice and make sense of.

Endnote

- (1) The discussion about the contrasting video cases started with participants’ debating whether the cognitive demand was declined in both of the video cases or not. Because their debate was staying at a general level, the facilitator suggested participants to discuss what was happening in the video cases that might have caused maintenance or decline instead of trying to make a decision about that. The facilitator’s question, then, changed the nature of the discussion.

References

- Ball, D.L., & Cohen, D.K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3-32). San Francisco: Jossey-Bass.
- Blumenfeld, P. C., & Meece, J. L. (1988). Task factors, teacher behavior, and students' involvement and use of learning strategies in science. *The Elementary School Journal*, 235-250.

- Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26(3&4), 369-398.
- Borko, H., & Koellner, K. (2008). *Situativity: a theoretical lens for designing and studying programs of professional development*. Paper presented at the International Commission on Mathematical Instruction Symposium, Rome, March 2-5. Retrieved on April 3, 2012 from www.unige.ch/math/EnsMath/Rome2008/WG2/Papers/BORKO.pdf
- Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education*, 24(2), 417-436.
- Bransford, J. D., Franks, J. J., Vye, N. J., & Shenvood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 470-497). New York: Cambridge University Press.
- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 24, 61-100.
- Colestock, A. & Sherin, M. G. (2009). Teachers' sense-making strategies while watching video of mathematics instruction. *Journal of Technology and Teacher Education*, 17(1), 7-29.
- Doyle, W. (1983). Academic work. *Review of Educational Research*, 53(2), 159.
- Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist*, 23(2), 167-180.
- Fishman, B., & Davis, E. A. (2006). Teacher learning research and the learning sciences. *The Cambridge handbook of the learning sciences* (535-550). New York: Cambridge University Press.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28(5), 524-549.
- Koc, Y., Peker, D., & Osmanoglu, A. (2009). Supporting teacher professional development through online video case study discussions: An assemblage of preservice and inservice teachers and the case teacher. *Teaching and Teacher Education*, 25(8), 1158-1168.
- Le Fevre, D. (2003). *Designing video-based multimedia curriculum for teacher learning*. Paper presented at the AARE Conference. New Zealand: Auckland. Retrieved on March 12, 2013 from <http://www.aare.edu.au/03pap/lef03185.pdf>
- Marx, R.W., Blumenfeld, P. C., Krajcik, J., Blunk, M., Crawford, B., Kelly, B., et al. (1994). Enacting project-based science: Experiences of four middle grade teachers. *Elementary School Journal*, 94, 499-516.
- Marx, R. W., Blumenfeld, P. C., Krajcik, J. S., & Soloway, E. (1997). Enacting project-based science. *The Elementary School Journal*, 97(4), 341-358.
- Miller, K. & Zhou, X (2007). Learning from classroom video: What makes it compelling and what makes it hard. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 321-334). Mahwah, NJ: Erlbaum.
- NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.
- Sanford, J. P. (1987). Management of science classroom tasks and effects on students' learning opportunities. *Journal of Research in Science Teaching*, 24(3), 249-265.
- Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education*, 11(2), 107.
- Stein, M.K., Smith, M.S., Henningsen, M.A., & Silver, E.A. (2000). *Implementing standards based mathematics instruction: A casebook for professional development* (First Edition). New York, NY: Teachers College Press.
- Stein, M.K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classroom. *American Educational Research Journal*, 33(2), 455-488.
- Tripp, T. R., & Rich, P. J. (2012). The influence of video analysis on the process of teacher change. *Teaching and Teacher Education*, 28(5), 728-739.
- van Es, E. A., & Sherin, M. G. (2006). How different video club designs support teachers in "Learning to Notice". *Journal of Computing in Teacher Education*, 22(4), 125.
- van Es, E. A., & Sherin, M. G. (2009). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education*, 13(2), 155-176.

Insights Into Teacher Reflective Practice During Planning

Michael T. Dianovsky, South Dakota State University, Department of Chemistry and Biochemistry,
Box 2202, Brookings, SD, 57007, michael.dianovsky@sdstate.edu

Donald J. Wink, University of Illinois at Chicago, 845 W. Taylor St., Chicago, IL 60607, dwink@uic.edu

Abstract: This study was designed to get an insight into the reflective nature of teacher planning. Specifically, we looked at what external and internal resources do teachers draw on in their reflections during the action of planning. Two cases of experienced urban high school chemistry teachers planning instruction for a similar set of chemistry topics were used in this study. The types of data used in this study included think-alouds, interviews, classroom observations, and artifact collection. The findings analysis of the two cases that both teachers used a variety of external and internal resources during their planning. The teachers' use of resources was associated with indications of reflection. The type of resources that a teacher used influenced the type of reflection the teacher would make. In order for reflection to occur, the resource had to have meaning to the teacher.

Introduction and Theoretical Background

Professionals of all types, including teachers, are in positions in which their prior experiences form the basis of their management of particular problems that they confront in their practice. Building on the work of Dewey (1933) and Schön (1983), a critical component of the way in which professionals respond when presented with new situations is the act of reflection. Schön (1983) argued that professional could benefit from and improve their practice by reflecting. However, Marcos and Tillema (2006) cautioned that research that analyzes teacher reflective thought out of context would reveal only part of the story. Thus, research should make a connection between reflection and specific actions of teacher practice. Research studies on teacher reflection typically focus on teachers' change in beliefs, personal identity, attitudes towards teaching, and self-efficacy (Marcos, Sanchez, & Tillema, 2011). These studies look at the internal mental dynamics of the teacher, but often stop short of linking those internal characteristics to the external actions of creating and executing a lesson plan; in other words, the actions that result from teacher reflection.

Teacher planning is one action of teacher practice that researchers can examine to better understand teacher reflection in action. Currently, teacher cognitive research has predominantly examined teacher thought processes before instruction (planning), during instruction, and after instruction (reflection) in an effort to understand the decisions teachers make and the cognitive processes they employ (Hall & Smith, 2006). Also, type of research occurs in distinct phases and not as a holistic process. Although studies examining reflection have appeared periodically, more recently it has become a growing area of inquiry (Ethell & McMeniman, 2000). While few studies try and link planning with instruction, little research has been conducted examining planning, instruction, and reflection as a holistic process (Hall & Smith, 2006). There is purpose in describing in depth the reflection that occurs during the action of teacher planning. Reflective teaching practice is viewed as an important component to enhance the development of effective teachers (Dallas, Reed, & Graves, 2010; Schön, 1983; van Manen, 1995). Currently, the literature does not include a way to document reflection during planning. Without any evidence of what teacher reflection looks like during a specific action of his or her practice, teachers are still told they need to be reflecting during his or her practice, and are even trained to do so by researchers and educators (Craig, 2010).

Teacher instructional planning is about selection and use of resources. Researchers need to evaluate the way in which science teachers use specific curriculum materials to improve instruction. Choosing curriculum content is one of the most complex planning tasks faced by teachers (Ball & Cohen, 1996). If there is a specified curriculum, teachers do not need to spend large amounts of time deciding which content they will be teaching next or how the curriculum aligns with standards (Duschl, Schweingruber, & Shouse, 2007; NRC, 1996). On the other hand, teachers do need to plan for instruction while using reform curricula in order to select appropriate resources. Reform curricula provide multiple resources to aid planning, such as pacing guides, coaching, model lessons, and teaching summaries. Thus, researchers should ascertain which resources teachers use during planning and which they do not. Therefore, in order to fully describe reflection during planning the types of resources a teachers uses as well as how teachers use these resources to promote reflection needs to be examined. This examination provides an opportunity to explain what resources afford or constrain a teacher's opportunity to reflection during planning.

In addition, experienced teachers can offer important insights about teacher reflection. A primary goal of many teacher education programs is to develop reflective teachers (Dallas, Reed, & Graves, 2010). The current study can offer insight into the types of reflection employed by experienced teachers and provide cases to consider for those in teacher preparation programs. Thus, it is important to characterize reflection during specific components of teacher practice.

For this study, the High School's Transformation (HST) initiative provided an opportunity to study the reflective practices of two science teachers in an urban high school context with ongoing reform and a specified curriculum. The specific reform used for the study was Inquiry to Build Content—An Instructional Development System (IDS). The IDS provided a community of teacher an opportunity to build science content, concepts, and skill development around the instruction approach of guided inquiry.

In this paper, we discuss a research study designed to examine case studies on the ways in which reflection can support the careful planning that is an essential part of good teacher practice (Carlo, Hinkhouse, & Isbell, 2010; van Manen, 1995). Specifically, we answer the question of what types of reflections are associated with planning in relation to the kinds resources a teacher draws on during his or her reflections. This study looked at two cases of experience high school chemistry teachers planning instruction for a similar set of chemistry topics. Providing explanations of what reflection and resource use actually looks like during teacher planning will offer educators and researchers a chance to understand reflection during planning. With this knowledge researchers can train teachers in how to develop a better reflective practice while engaged in teacher planning during professional development. Also, curriculum developers can get a better insight into how teachers use curriculum materials for the purpose of planning.

Method

The goal of this research was to describe and analyze teacher reflective practices during the instructional planning of chemistry teachers and the materials or experiences teachers used to support the reflections. To achieve this goal a qualitative multiple-case study approach was used. A case study was appropriate for this research because little is known about reflection during planning for instruction and the relationship between the boundaries and phenomenon are not evident. Two cases comprised the current study, defined as experienced urban high school chemistry teachers planning instruction for a similar set of chemistry topics.

Study Context

Two high schools participating in the IDS reform project in a large Midwest urban school district functioned as the study sites. The criteria for selecting teachers to participate in the study were their experience level in teaching, their level of participation with the curriculum, and plan for instruction on a regular basis. The two participants were urban high school chemistry teachers. The pseudonyms selected for the two teachers were Janice and Christina. The two teachers participated during the Fall 2010 semester.

The two teachers that participated in this study derived from a sample of ten teachers. Both teachers had four years of experience in working with the IDS curriculum and were committed to the IDS program. At the beginning of the school year, both teachers indicated they would be actively implementing the IDS curriculum in their classrooms and working to gain ownership in the curriculum. However, when the study was about to start Christina decided that she would be using parts of the IDS curriculum in her classroom as well as the curriculum she used prior to participating in the IDS. Both teachers taught very similar units covering almost identical chemistry content.

Data Sources

Data for this research was collected during six weeks of teacher instruction. The first data source consisted of teacher lesson plan think-alouds. Since experienced teachers are more likely to engage in mental planning instead of writing more formal lesson plans, the teachers were asked to participate in think-alouds during the act of planning. Both teachers were given Livescribe Pulse Smartpen® to complete the think-aloud planning session tasks. Livescribe® developed an application that records both audio and writing during the same moment in time. This tool allowed to the researcher to capture reflective thought during the action of planning. Both teachers planned 9-10 lesson in the Livescribe notebooks provided to them. The second data source was artifact collection. At the end of the study, each of the teachers' Livescribe notebooks, formal lesson plans, and worksheets were collected. The last data source consisted of three interviews with teachers conducted after 3-4 lesson planning sessions. The purpose of the interviews was to obtain descriptive information about teacher planning and to gain a better understanding of teacher reflections and resource use during the planning process.

Data analysis consisted of analyzing transcribed think-aloud and interview sessions for both teachers. Codes were generated around two essential categories, the types of resources used by the teachers and types of reflections. Teacher reflective statements were coded based on three types of reflections defined in the literature. The first included reflection-in-action defined by Schön (1983) as a process involving thought during action. The second type of reflection included reflection-on-action, as defined by Schön (1983) as a process that involves action then thought. The last type of reflection included reflection-for-action defined by Killion & Todnem (1991) as a process that involves thought then action. Cases for each teacher were then developed to describe teacher resource use and reflection. Lastly, a cross case analysis completed by analyzing similarities and differences between the two cases by exploring themes and categories that emerged form the analysis of the two cases.

Data Analysis and Findings

Resource Use

During the coding process, it became evident that the types of resources the teachers used came from different sources. A resource was defined as a tool or other thing a teacher can use to reflect or use in his or her planning process. Two themes emerged from the coding of both teachers' data in regards to the types of resources used during reflection. The first category was external resources. These were resources provided to the teacher by the IDS or supplement resources the teachers acquired. The second theme included internalized resources. These resources once external to the teacher were internalized through the teachers' understanding and prior experiences. Table 1 displays the types of resources both teachers used during planning.

The external resources Janice used during her planning were mainly comprised of the materials provided to her by the IDS, such as, model lessons, the science writing heuristic, pacing guides, assessments, worksheets, and the textbook. In comparison, Christina did not use any of the IDS materials provided to her, since she did not implement the IDS curriculum to its full extent. She used outside external resources from other curricula as well as the Internet during her planning. The internalized resources the teachers used included their knowledge of their students, their pedagogical content knowledge (PCK), and skills and dispositions. An example of each internalized resources is provided below.

Table 1: Themes for resources used by Janice and Christina

Janice's Resources	Christina's Resources
External	
Science Writing Heuristic (SWH)	Textbooks
Model Lesson	Worksheets
Assessments	Internet
Pacing Guides	ACT Standards
Textbooks	Lab Activities
Worksheets	Assessments
Professional Learning Community	
Internalized	
Knowledge of Students	Knowledge of Students
Pedagogical Content Knowledge (PCK)	Pedagogical Content Knowledge (PCK)
Skills and Dispositions	

Throughout the course of the study both teachers utilized their knowledge of their students. Janice and Christina made sense of what they believed their students could and could not accomplish in their classrooms. Both teachers continually were building upon their knowledge of their students. Examples included their knowledge about their students': motivation, math skills, writing skills, listening skills, transferring knowledge, working in groups, and ability to think independently. One example that demonstrated the use of knowledge of students was when Christina was planning a lesson on periodic trends. She stated,

Overall my students did a decent job on the worksheet. One problem I noticed is that they are still having a hard time reasoning with data. A skill I have been working on with them since the first week of the semester. They still struggle with it. The one question that asked them to predict the electronegativity of Iodine given the electronegativity of three other elements in the same family seems like they drew a blank. If it is not something I have told them specifically how to do, they don't even try to make sense out of what the problem is asking. Instead they leave it blank or write some random answer down. No motivation.

Here, Christina drew on the resource of her knowledge of her students in order to explain why her students had a difficult time answering one of the questions on the worksheet. Christina connects this to her students' lack of motivation.

The second internalized resource both teachers used during planning was their PCK. PCK refers to how teachers' use both pedagogical knowledge and content knowledge in forming ways of knowing about how

to teach the content to their students. In regards to planning, a teacher's PCK may contribute to their understanding of what to plan for, why they are planning this lesson, and how to plan for the lesson. For the purpose of this study, experiences are knowledge gained from what one has observed, encountered, and undergone in the past. A strong link existed between both teachers' content knowledge and pedagogical approaches they used in their classroom. An example of PCK as an internalized resource was when Janice was planning a lesson on accounting for atoms. One of Janice's pedagogical decisions was to ask students to represent a chemical reaction in various ways as a bell-ringer. These representations include using symbols, words, and at the particulate view. She stated:

I know from my own knowledge that understanding what a chemical reaction represents can be tricky at first. Before I first stated this curriculum I rarely looked at chemistry at the particulate level of representations. I was never taught that. I only learned it the symbolic way and to write it in a sentence. So that is why I need to express them [the representations] independently first to my students as a bell-ringer. I want to clarify the confusion they have right away, just because the particulate was new to me when I first started, and they will need time with the material just like I had to take time, even though I understood it right away.

When Janice used her PCK as a resource, she used it to discuss her own experience in learning the material for the first time. This allowed her to connect to the experience she had when she first learned particulate representations in working with the IDS curriculum. She then connected these two experiences to the experiences her students might have when she teaches the material. She developed a deeper understanding of her own chemistry knowledge and related to this to her students' learning.

The last type of internalized resource was the teacher's skills and dispositions. Janice was the only one who used this as a resource during her planning. Skills include a learned capacity to solve novel problems. Janice drew on the specific skills of communication, organization, classroom management, and seeking feedback. Dispositions were defined as the tendency to act in a certain manner under a given circumstance in order to create change. Janice's dispositions included motivation, curiosity, creativity, patience, open-mindedness, and self-awareness. One example of Janice using her skills as a resource was when she was planning a lesson on a lab activity called metal or nonmetal. She states:

It is nice to have students move around the room and work with each other, gives them another type of learning experience, one I think is so important, however, it can get out of control, and really fast. I have to manage my students accordingly in order for learning to occur in my classes. As my goal is to provide students with the best learning opportunity and I have to manage my classroom for this to happen. I know this too well. Therefore, I will make sure the groups the students working together are ones in which little disturbance should occur.

In this example, Janice first talks about the benefits of having her students work together in groups, which represents her skill of classroom management. This learning experience for Janice's students is important for her. On the other hand, group work can lead to a disruptive learning environment. In this instance when Janice identifies a problem that she draws on her skill of classroom management as a resource to reflect on how to provide an environment in her classroom that is conducive for learning. An explanation of the relationship between the types of resources used and the reflection types are discussed in the next section.

Teacher Reflection

The type of resource that a teacher used influenced the type of reflection the teacher would make. Since teachers reflected for the specific action of planning, reflection-for-action was the most common type of reflection. All resources, external and internalized, were a basis for reflection-for-action. An example of a reflection-for-action statement by Janice is when she is making a decision to use the Science Writing Heuristic (SWH) (external resource) as part of her pedagogy for an inquiry based activity. Janice stated, "The reason why I want to use the SWH is because I want to see what they [students] are thinking and to make sure they are making sense of what they just did." In this statement Janice reflects on the SWH to justify why she wants to add it into her lesson plan. She is reflecting-for-action here. This type of reflection allowed her to move forward in her planning because she gained an understanding of what the resources afforded her in her planning process.

A second example of a reflection-for-action statement made by Christina was when she referenced the use of the textbook for another pedagogical activity on molar mass. She stated, "Students will answer textbook problems 1 through 9 on page 163 of their textbook. There are a variety of compounds in these problems. Some basic and some more complex. Students will turn these into me to be graded." When she drew on the worksheet as a resource, she illustrated what students would accomplish as well as what the concepts the

worksheet covered. She was reflecting-for-action; in regards to the pedagogical moves she planned on taking.

Not all resources types were directly related to reflection-for-action. Some resources that teachers drew on did not immediately allow them to reflect-for-action such as internalized resources. Both teachers would draw on an internalized resource to reflect-on-action. This reflection allowed the teachers to relate their understanding of the resources to their current action of planning. Following this reflection they would then reflect-for-action. For example, Janice reflects-on-action in regards to her understanding of her PCK (internalized resource), while planning a lesson that involves introducing the mole concept. She states,

Now that I am introducing a new term associated with the periodic table, I am going to make sure I review the other numbers on the periodic table with my students again. Last year I did this and my students did not get confused about which number meant what on the periodic table like my class did during the first year of teaching this. From the first year teaching atomic mass, my students kept forgetting which number was which on the periodic table. That is why I need to add something to this lesson. Since it worked last year, I am going to do it again this time around.

In this statement, Janice drew on her PCK as a resource by reflecting on two different prior experiences. The first was about a pedagogical decision she made in teaching the lesson the previous year. When she drew on this resource, she was reflecting-for-action and reflecting-on-action. Janice needed to support her pedagogical decision, the action, because she wanted to assure it was the best approach for her students to learn the material. Janice was building on her pedagogical knowledge. Janice noted a second prior experience when she taught atomic mass during the first year working with IDS. This was Janice reflecting-on-action. She identified a problem three years ago when she first taught the lesson when students had confusion about the meaning of the numbers on the periodic table. Janice had a better understanding of how to introduce the term mole to her students because she could connect prior experiences to her current experience.

Reflection-in-action occurred when the teachers encountered a problem. This allowed both teachers to then draw on an external resource. For example, when Christina was planning a lesson on oxidation-reduction she stated,

I am stuck trying to think about a good lesson to use. I do not like teaching this content. The activities and techniques I have used before just don't seem to work. I found an activity online that I am going to try out. I am going to add this into my lesson plan and see if it works with my students this year.

In this instance Christina encountered a problem and reflects-in-action. The problem she identified is to find a new approach to teach oxidation-reduction. She uses the Internet (external resource) to find a new pedagogical approach. This then allows her to reflect-in-action by continuing on with her planning process. This reflective action seemed to drive resource use by the teachers and help them in understanding something that happened in the past. Overall, the resources used by the teachers were specific to what the teacher thought they needed to accomplish and also had meaning to the teacher. The type of resources that the teachers used influenced the type of reflection the teacher would make.

Conclusion

The type of resources that a teacher used influenced the type of reflection the teacher would make. Not all resource types were directly related to reflection-for-action. This means that there are specific resources that teachers drew on that did not immediately allow them to reflect-for-action. These resources include internalized resources. Both teachers would draw on an internalized resource to reflect-on-action first. Following this reflection they would then reflect-for-action. External resources were always associated with teacher reflection-for-action.

Both teachers had different conceptions of what resources they needed during planning. What this means is that the teachers did not rely on the same resources during planning. The way the teachers structured their model of planning influenced their resource choices. If the teachers structured her model around her own understanding of her own practice, she was more likely to draw on internalized resources like her PCK and her skills and dispositions. However, if a teacher structures her model of planning around only her students, she would be less likely to draw on internalized resources. What was important here was that it was critical for teachers to have good resource choices available to them. Resources matter in the reflective planning process. Even resources one might think are not important to teachers' planning processes might actually be important to the teacher. Resources function differently for each teacher. When developing resources for teachers to use for planning, it is important to not limit or constrain how a teacher should use that resource. A teacher may see a benefit in using that resource that was different from the way in which it was intentionally designed. Every

resource could have potential value to a teacher, so it is crucial that resources are not labeled as “good” or “bad.”

However, what is essential is that the resource has meaning to the teacher. Teachers need to understand the affordances of the resources in order for reflection during planning to occur. Teachers make meaningful decisions to use resources during their planning processes. There is a reason why a teacher draws on a specific resource during a particular time in his or her planning process, and that is to meet these goals and intentions. In order to do this, teachers must reflect accordingly by drawing upon a specific resource they know will help them move forward.

Implications for Research and Instruction

Our work presents a first attempt to understand the teacher reflective practice during the act of instructional planning. Currently there is no method that has been designed to really see reflection-in-action in teacher practice. This methodology provided an opportunity to capture these two teachers’ reflective planning processes. By using this approach I was also able to document what the teachers were thinking during a given point in time of their planning. There was no wait time needed. This methodology also allowed the teachers to plan for their lessons anywhere and anytime they wanted to. They were not restricted by any means to plan for their lessons with me there with them. This process allowed for teachers to plan in their naturalistic settings, whether it was at school, at home, or on the train. There was also no restriction on the teachers to have to plan a lesson during one sitting. The teachers could plan for a lesson over a period of multiple days if they chose to.

In order to get a better understanding of teacher practice, it is important to look at teacher practice in authentic detail. What this means is that we will be able to capture the true nature of a teachers practice. This is important because we do not know what this looks like well enough. In order to accomplish this task, we need to continue to develop methodologies that do not restrain a teacher while they are engaged in a particular component of their teacher practice. This will allow teachers to make connections between different components of their practice, as well as the freedom to draw on resources that they find valid.

By looking at teacher practice researchers will get a broader view of what this practice actually looks like. We will see the uniqueness and complexity within a teacher’s practice. Also, we will get a better understanding of what is important to the teacher, and how this allows the teacher to bring in their expertise. This study has shown that it is important to value a teacher’s expertise, PCK, and their knowledge of students. The teachers in this study relied heavily on understanding their prior experiences to plan for their lessons. These teachers are continually building upon their knowledge and understanding of their practice and their students on a daily basis. Both teachers had similar and different aspects to their expertise, PCK, and knowledge of their students that they would draw on as resources during their planning process. It is important to value a teacher’s expertise and their knowledge of their practice and students because this is essential to reflection and decision making in planning. This also affects the resource choices made by the teachers.

Specifically to science reform, research that focuses on science reform does not always take into account a teachers’ expertise (White & Frederiksen, 1998; Collopy, 2003; Forbes & Davis, 2010). This especially occurs when reform projects develop curriculum materials for teachers. Teachers in science reform participate in professional development that tells them, for example, how to teach a particular lesson, how to use new curriculum materials, and how to plan for a lesson. Reform science tries to change teachers’ practice. This study shows how important it is for curriculum and PD planners to take into account a teachers’ expertise, especially in planning, when designing a science reform curriculum. Also, science reform needs to develop materials with a notion that teachers might not use them for the intended purpose they were developed for. Science reform may need to look at the content expectations instead of practices.

References

- Ball, D. L., & Cohen, D. K. (1996). Reform by the books: What is: Our might be: The role of curriculum materials in teacher learning and instructional reform? *Educational Researcher*, 25, 6-8.
- Carlo, D. D., Hinkhouse, H., & Isbell, L. (2010). Developing a reflective practitioner through the connection between educational research and reflective practices. *Journal of Science Education and Technology*, 19, 58-68.
- Collopy, R. (2003). Curriculum materials as a professional development tool: How a mathematics textbook affects two teachers’ learning. *The Elementary School Journal*, 10, 227-331.
- Craig, C. J. (2010). Reflective practice in the professions: Teaching. In N. Lyons (Ed.), *Handbook of Reflection and Reflective Inquiry: Mapping a Way of Knowing for Professional Reflective Inquiry*. (pp. 189-214). New York: Springer.
- Dallas, D., Reed, M., & Graves, E. K. (2010). The impact of required teacher reflection on classroom practice and student learning. *Kansas English*, 94, 64-70.
- Dewey, J. (1933). *How We Think: A restatement of the relation of reflective thinking to the educative process*. Boston: D. C. Heath.

- Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.) (2007). Taking science to school: Learning and teaching science in grades K-8. Washington, DC: National Academies Press.
- Ethell, R. G., & McMeniman, M. M. (2000). Unlocking the knowledge in action of an expert practitioner. *Journal of Teacher Education*, 51(2), 87-101.
- Forbes, C. T., & Davis, E. A. (2010). Curriculum design for inquiry: Preservice elementary teachers' mobilization and adaptation of science curriculum materials. *Journal of Research in Science Teaching*, 47, 820-839.
- Hall, T. J., & Smith, M. A. (2006). Teacher Planning, Instruction and Reflection: What We Know About Teacher Cognitive Processes. *Quest*, 58(4), 424-442.
- Killion, J., & Todnem, G. A. (1991). A process for personal theory building. *Educational Leadership*, 48, 14-16.
- Marcos, J. M., Sanchez, E., & Tillema, H. H. (2011). Promoting teacher reflection: What is said to be done. *Journal of Education for Teaching: International research and pedagogy*, 37, 21-36.
- Marcos, J. M., & Tillema, H. (2006). Studying studies on teacher reflection and action: An appraisal of research contributions. *Educational Research Review*, 1, 112-132.
- National Research Council (1996). National science education standards. Washington, D.C: National Academy Press.
- Schön, D. (1983). *The Reflective Practitioner: How professionals think in action*. New York: Basic Books.
- van Manen, M. (1995). On the epistemology of reflective practice. *Teachers and Teaching*, 1, 33-50.
- White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, 16, 3-118.

Analyzing Online Knowledge-Building Discourse Using Probabilistic Topic Models

Weiye Sun, Department of Informatics, University at Albany

Jianwei Zhang, School of Education, University at Albany

Hui Jin, Department of Educational Technology, Shanghai International Studies University

Siwei Lyu, Department of Computer Science, University at Albany

1400 Washington Ave, Albany NY 12222

Email: wsun2@albany.edu, jzhang1@albany.edu, nmjinhui@163.com, lsw@cs.albany.edu

Abstract: This exploratory study tested the use of machine learning techniques, in particular, probabilistic topic models, to conduct automated analysis of the online discourse a Grade 4 knowledge-building community that investigated optics over three months using Knowledge Forum. Using the Latent Dirichlet Allocation (LDA) model, we extracted ten distinct and semantically meaningful clusters (i.e., topics) from the online discourse, which overlapped substantially with—although did not directly map onto—the inquiry themes identified by students and inquiry thread topics identified by researchers. The LDA analysis further identified discourse entries relevant to each of the topics, with acceptable agreement achieved between the automated analysis results and the manual coding of two researchers.

Introduction

A focal challenge in the learning sciences community is to create collaborative knowledge building environments in line with how real-world knowledge communities work. Supported by collaborative online environments, such as Knowledge Forum (Scardamalia & Bereiter, 2006), students engage in sustained and interactive discourse to advance their collective understanding. Working as a knowledge building community requires students to take on collective responsibility for advancing their collective knowledge beyond individual learning (Scardamalia & Bereiter, 2006; Stahl, 2006). They need to develop a reflective awareness of the major themes of inquiry emerging from the diverse input of all community members, build on important ideas of peers, identify progress and gaps, and envision directions for collaborative work and contributions (Zhang et al., 2009). In parallel with such challenges for students, the teacher needs to actively follow the online discourse to understand the landscape of collective ideas, identify and assess advances in focal areas, and foster efforts to investigate emerging and deeper issues. However, manual implementation of such analyses of online discourse is often labor-intensive and demanding even for researchers, let alone teachers or students. This calls for new assessment and analysis tools to help students and their teacher trace online discourse over time and provide feedback on collective progress as well as individual participation. The purpose of this research is to test automated analysis based on machine learning techniques to discover and trace major topics of inquiry based on online discourse data. Such automated analysis may provide learners and teachers with ongoing assessment and feedback about the evolving landscape and status of their collective understanding achieved through online discourse; it also provides researchers with new and less labor-intensive tools to analyze online discourse.

Tracing and Analyzing Online Knowledge-Building Discourse as Unfolding Conceptual Streams: Idea Threads

Diverse measures have been developed to trace and examine knowledge advancement in online discourse (Hmelo-Silver et al., 2007; van Aalst & Chan, 2007; Weinberger & Fischer, 2006), with efforts further made to automate some of the analyses (Law et al., 2007; Rosé et al., 2008; Teplov & Fujita, 2009). To trace collective idea advancement in online discourse, Zhang (2004) developed inquiry threads (idea threads) analysis in light of the multilevel emergent processes of collaborative knowledge building and the discourse flow theory that explains the dynamics giving directions to the flow of thoughts in discourse (Chafe, 1997). In interactive conversations, members dynamically maintain a shared focus of active consciousness, which constantly shifts and develops as the conversation proceeds. Interrelated focuses of conversation constitute larger focus clusters—discourse topics. It is often presented in the form of a problem or gap yet to be filled that engages members' thinking in an active and determinate psychic way. The evolving streams of the whole discourse are sustained through the joint interplay of the participants, whose successive focuses of consciousness propelled forward by their reaction to what others have just said (Chafe, 1997). The whole course of conversation unfolds in an emergent and improvisational manner. Inquiry threads analysis traces emergent, collective idea progress in extended discourse by identifying focal topics or thematic problems that have emerged. Discourse contributions that address each shared problem/theme (e.g. how are rainbows created) form into an inquiry thread, a conceptual stream of discourse extending from the first to the last discourse entry (Zhang et al., 2007). Advances in each inquiry thread are further elaborated through analyzing the content of the interactive discourse

input (e.g., questions, ideas). The scope and depth of the discourse is additionally benchmarked by comparing the themes of the inquiry threads against the curriculum expectations (Sun et al., 2010). A Web-based program—Idea Thread Mapper (ITM)—has been developed to help researchers create inquiry threads based on online discourse data as a means to examine collective knowledge progress. ITM has further been turned into a classroom tool to help teachers trace and assess collective progress in online discourse beyond individual postings and engage students in reflective monitoring of progressive ideas and deepening goals. Students and their teacher work together to identify focal knowledge themes and goals and to cluster discourse entries related to each theme as a conceptual thread of inquiry. Using the ITM tool, students can visualize a set of interrelated inquiry threads on a timeline, reflect on advances of knowledge over time, envision cross-theme connections and collaborations, and identify weak/emerging areas and deeper challenges as the focus of their further work (Chen et al., 2013; Zhang et al., 2014). The current study explores automated analysis to identify major themes of discourse as well as discourse entries related to each topic. We envision that such tools can facilitate students, teachers, and researchers to analyze and assess conceptual threads of inquiry unfolding in online discourse. Furthermore, they may also help in developing assessment of online discussions in general online educational contexts. Instead of simply evaluating forum posts as stand-alone essays, instructors can use such analysis to obtain the conceptual landscape of the entire discussion (Dringus & Ellis, 2005) and further assess whole class progress as well as individual participation and contribution.

Existing Work to Analyze Online Discourse Using Probabilistic Topic Models

In educational data mining, automated content analysis has been tested as a major method for online discourse analysis (Rose et al., 2008; Mu et al., 2012; Baker & Yacef, 2009; Romero & Ventura, 2007; Romero & Ventura, 2010). Content analysis requires the design of pre-defined coding schemes focusing on various patterns of discourse of interest to the researcher. Drawing on some of the well-tested coding schemes focusing on argumentative discourse (e.g., Weinberger & Fischer, 2006), researchers have designed automated content analysis tools using natural language processing and machine learning techniques (Mu et al., 2012). The existing automated content analysis tools primarily focus on patterns of discourse at a single post level. They usually lack the ability to address conceptual themes of ideas that evolve in this whole course of the discourse.

As an alternative to content analysis, topic models that automate the discovery of thematic topics from a corpus have gained increasing attention. One appealing property of topic models is that they are unsupervised algorithms, which obviates the requirement of manually annotating the corpus, and may help to reduce cost, and improve objectivity of the analysis results. The objective of probabilistic topic modeling is to automatically discover the topics in a corpus and the topic assignments of each document using only the documents in the corpus (more precisely, all the words, or the frequency of unique words). In the language of statistical analysis, we treat the documents as the observed variables, while the topic structures—the individual topics, the per-document topic distributions, and the per-document per-word topic assignments—are regarded as the latent variables in the model. The central computational problem in probabilistic topic modeling is to use the observed documents to infer the latent topic structure. This can be thought of as “reverse engineering” the simple generative process underlying the topic models—what is the latent topic structure that is likely to generate the observed collection. Two common choices of topic models are the latent semantic indexing (LSI) (Hofmann, 2001) and latent Dirichlet allocation (LDA) (Blei, 2012). In LDA, we assume that the conditional distribution of the topic assignment given the per-document topic proportion, the conditional distribution of the observed word given all the topics and the per-word topic assignment are multinomial distributions, while the prior distributions over the individual topics and per-document topic assignments are Dirichlet distributions. According to the Bayesian framework, the generative model reduces to compute the conditional distribution of the topics and topic assignments of each word and document given the observed corpus. In practice, we use approximation methods to evaluation of the document posterior distribution, the two main categories of which are variational methods and sampling based methods. Though both methods have been shown leading to reliable inference performances, in this work, we employ the variation-based method for its running efficiency. Compared with LDA, LSI has a significant drawback in that the “topics” recovered from LSI lacks a clear semantic meaning. On the other hand, the topics in LDA are distributions of words, and correspond to components in the probabilistic generative model. More importantly, LDA usually outperforms LSI in practice (Blei, 2012) mainly due to its better handling of synonymy and polysemy because of the probabilistic association of words to topic. Therefore in this study, we tested using LDA to discover and trace major topics of inquiry based on online discourse.

Early adoptions of topic models for educational data include the work of Ming and colleagues (2012), which applied two topic models, namely probabilistic LSI and hierarchical LDA, to predict the grades of the students. Y. Zhang and colleagues (2012) applied LDA to online discussions of four Chinese classrooms to extract topics and display the temporal profiles of the topics. Their study suggests that frames built from the top terms of the learned topics support easier human interpretation. Beyond online learning, Sherin (2012, 2013) tested using LDA and latent semantic analysis to extract fragments (categories) of ideas from student interviews

in order to code misconceptions versus scientific explanations. The results of the automated analysis aligned closely with the coding of human analysts. The above studies point to the potential of LDA to capture conceptual topics and structures in student discourse data. However, this potential needs to be further explored by applying the topic model analysis to organize online discourse of productive knowledge building communities that engage in emergent, progressive inquiry over a longer term to capture unfolding directions of collective knowledge work. Such automated analysis needs to be benchmarked by comparing the results with manual analysis of human analysts. Therefore, this exploratory study intends to use topic model analysis to examine unfolding processes of collective knowledge building in the online discourse of a Grade 4 knowledge building community and compare the results with human coding.

The Corpus of Online Discourse and Classroom Context

This research analyzed the online discourse of a class of 22 fourth-graders (9-to-10-year-olds) who studied light over a three-month period supported by Knowledge Forum: a collaborative online knowledge building environment (Scardamalia & Bereiter, 2006). The students were taught by a teacher who had strong expertise in facilitating knowledge building, as indicated through a prior study (Zhang et al., 2009). The classroom design encouraged students to take on collective responsibility for high-order decision-making related to knowledge goals, long-range planning, and progress evaluation. Students generated problems of understanding, discussed diverse ideas and theories through face-to-face knowledge-building discourse, conducted self-generated experiments and observations, and searched libraries and the Internet. They contributed problems, ideas, data, and resources into Knowledge Forum for continual discourse and idea improvement. Knowledge Forum provided the public space in which student ideas and inquiry work were recorded, in views (workspaces) corresponding to their focal goals. By writing notes (discourse entries) in these views, the students contributed their ideas, data, and related information. Supportive features for knowledge-building discourse allowed the students to co-author, build on, and annotate notes; create reference links with citations to existing notes; and create rise-above notes to summarize, distil, and advance their discussions. Content analysis of student portfolio notes that summarized their knowledge advances showed significant progress made over the three months in understanding deep issues about light (Zhang & Messina, 2010).

The corpus of online discourse contains 149 notes over a vocabulary of 824 distinct words, among which 75 words are stop words, namely, words that only assume grammatical functions or carry little meanings relevant to the analysis, such as articles, prepositions, and pronouns. After removal of the stop words, the number of meaningful distinct words is reduced to 749, with each note in the corpus containing 43 distinct words on average. Hereinafter, we use the terms ‘note’ and ‘document’ interchangeably depending on the context: we refer to discourse entries as notes in the context of Knowledge Forum, and as documents in the context of topic modeling.

Topics in online discourse can be defined at different levels of specificity, ranging from broader themes of discussion to specific topics and sub-topics. Different units of topic clustering may serve different needs: reflection of students and teachers on discourse progress tend to focus on relatively larger themes and topics while researchers often prefer more fine-grained analysis of topics and ideas. In the inquiry of light, the students and their teacher co-identified eight themes to guide their inquiry and discourse about light: Why is snow white? (reflection and absorption) How do plants adapt to light? Why is night cooler than daytime? How are shadows made? How does light travel? How are rainbows created? (colors of light) How do lenses work? And how do mirrors work? Elaborating on these eight themes, two researchers collaboratively identified 17 specific topics from student online discourse, and used this list of topics to code each Knowledge Forum note. Notes addressing the same topical issue constituted a conceptual line of inquiry—an inquiry thread—that extended from the first to the last document created (see Chen, 2013 for further details of this analysis). The result was the identification of 17 inquiry threads that map onto the eight overarching themes identified by the students.

Topics Discovered with LDA Analysis

The eight overarching themes and 17 specific inquiry threads provided the contextual information needed for us to test topic discovery with LDA analysis. We tested a range of total number of topics to be discovered ranging from 5 to 17 topics, and setting the number to 10 topics generated the most interpretable result.

Table 1 shows the 10 topics discovered through the LDA analysis, with a list of top 10 keywords for each topic. The ‘Keywords’ column lists the vocabulary that has the largest β value under a certain topic, that is, the words that are mostly likely to belong to that topic. In the ‘Interpretation’ column, we present a summarization of each topic obtained by analyzing the keywords used in the documents that the algorithm assigned to the topic. Some of the topics (e.g. Topic 9) are harder to interpret than others. There are substantial overlaps (shared keywords) between topics 1 (Light travels through materials), 5 (Reflection) and 9 (Materials that reflect); and between topics 3 (Shadows, including colored shadows) and 8 (Shadows and light sources). As we navigated through the results from our test with $M = 5, 6 \dots 17$ topics, we found that some topics are interpretable at certain M s but lost their interpretability as the parameter increases or decreases.

Table 1: Ten topics extracted by lda, each with the top keywords and an interpretation

Topic	Keywords	Interpretation
Topic 0	'colour' 'r' 'green' 'yellow' 'make' 'blue' 'object' 'cone' 'primary'	Colors of light
Topic 1	'tin' 'foil' 'solid' 'glass' 'travel' 'through' 'material' 'solstice' 'can'	Light travels (through materials)
Topic 2	'mirror' 'convex' 'when' 'concave' 'reflection' 'side' 'lens' 'telescope' 'two' 'see'	Mirrors and lenses
Topic 3	'rainbow' 'when' 'shadow' 'color' 'made' 'glass' 'through' 'colour' 'can' 'think'	Shadows (including colored shadows)
Topic 4	'glass' 'what' 'see' 'eye' 'solid' 'when' 'people' 'through' 'very'	See
Topic 5	'mirror' 'shine' 'reflect' 'direction' 'will' 'line' 'plant' 'this' 'work'	Mirrors and reflection
Topic 6	'sun' 'when' 'earth' 'moon' 'eclipse' 'shadow' 'other' 'world'	Eclipses and seasons
Topic 7	'white' 'snow' 'colour' 'prism' 'black' 'melt' 'when' 'see' 'fast'	Snow and white light
Topic 8	'shadow' 'object' 'made' 'opaque' 'energy' 'part' 'call' 'umbra'	Shadows and light sources
Topic 9	'through' 'go' 'can' 'reflect' 'tin foil' 't' 'think' 'was' 'angle' 'when'	Materials

Table 2 displays example documents for the first three topics. Aligned with the interpretation, these documents discuss colors, light traveling through materials, and mirrors and reflection, respectively. The documents in Table 2 are structured as the following: the first line of the documents lists the title, author initials and document creation date information in italic font separated by ‘||’; the contents of the documents are shown in the remaining lines. The different font color and superscripts represent the topic assignment of each word. For example, a word in green font with superscript 0 means that the topic assignment of this word is Topic 0. We refer the readers to the online supplementary tables for more examples.

Table 2: Example documents for the first three topics

Topic 0 Colors of light	<p><i>How you see colour</i> [redacted] 2009, Aug 21 10185</p> <p>When⁰ you stare⁰ at⁰ a blue⁰ object⁰ for a long⁰ time⁰, the cones in your eyes sensitive⁰ to blue⁰ becomes tired. There are green⁰ and red cones as well.⁰ If after that, you look⁰ at⁰ a white⁰ surface⁰, the sensitive⁰ to blue⁰ cones do not react,⁰ resulting to you seeing yellow⁰ instead.⁰ Yellow⁰ is the complimentary⁰ colour⁰ to white.⁰ Staring at⁰ a green⁰ object⁰ will⁰ produce⁰ a magenta⁰ after-image. (Blue⁰ +red =magenta)⁰ Staring at⁰ a red object⁰ will⁰ give⁰ a cyan⁰ after image.⁰ These colours make⁰ it happen.⁰ All⁰ the colours make⁰ white⁰ light.</p>
Topic 1 Light travels	<p><i>opaque and transparent</i> [redacted] 2003, May 06 10118</p> <p>"I agree¹ with [redacted] light bounces off shiny¹ materials like tin¹ foil.¹ Tin¹ foil¹ acts like a mirror.¹ Tin¹ foil¹ is solid¹ and so that means light can't⁹ travel¹ through¹ it."</p> <p>I¹ Tin¹ foil¹ is a solid¹ material¹</p> <p>Light travels in a straight¹ line¹ until it hits an object¹ if the object¹ is opaque¹ like wood¹, the light is interrupted. If the object¹ is transparent¹ like a piece¹ of glass¹ then it passes through¹</p> <p>is there 2¹ different¹ kinds of opaque¹ one¹ that light bounces off and one¹ that it just¹ stops.</p>
Topic 2 Mirrors and lenses	<p><i>what will happen when you shine a light through different lenses?</i> [redacted] 2003, Jun 10 9927</p> <p>concave² is when² an object² like a spoon² is indented so it looks like a cave²</p> <p>convex² is when² the object² is dented the opposite² direction.²</p> <p>there are many² different² kinds of lenses there is: one² side² convex² with a flat² side², double² sided convex,² double² sided concave², one² side² concave², two² concave² lenses facing the same² direction.²</p> <p>What² the difference² is when² you shine² a light through² the different² lenses?</p> <p>Our² experiment² we are going to shine² a light through² the different² lenses and see² were or² what² happens to the light. the double² concave² made² the light smaller² except² for when² I put² the lense² farther² from the magnifying glass² the light light sort² of got² closer² to the paper² and vis-versa.</p>

LDA-Based Analysis of Topic-Document Relationships

The LDA algorithm further outputs the assignment of topics to documents. Figure 1 shows how likely each of the 149 documents belongs to Topic 0 that was interpreted as colors of light. The x-axis lists all the 149

documents, and the y-axis shows the score of how much each document pertains to the topic ‘color’, with higher numerical score indicating greater relevance. We show the content of three documents of different likelihood scores in Figure 1. The first two documents have higher likelihood scores, and their contents are more pertaining to colors, while the last document has lower likelihood score and its content is not related to colors.

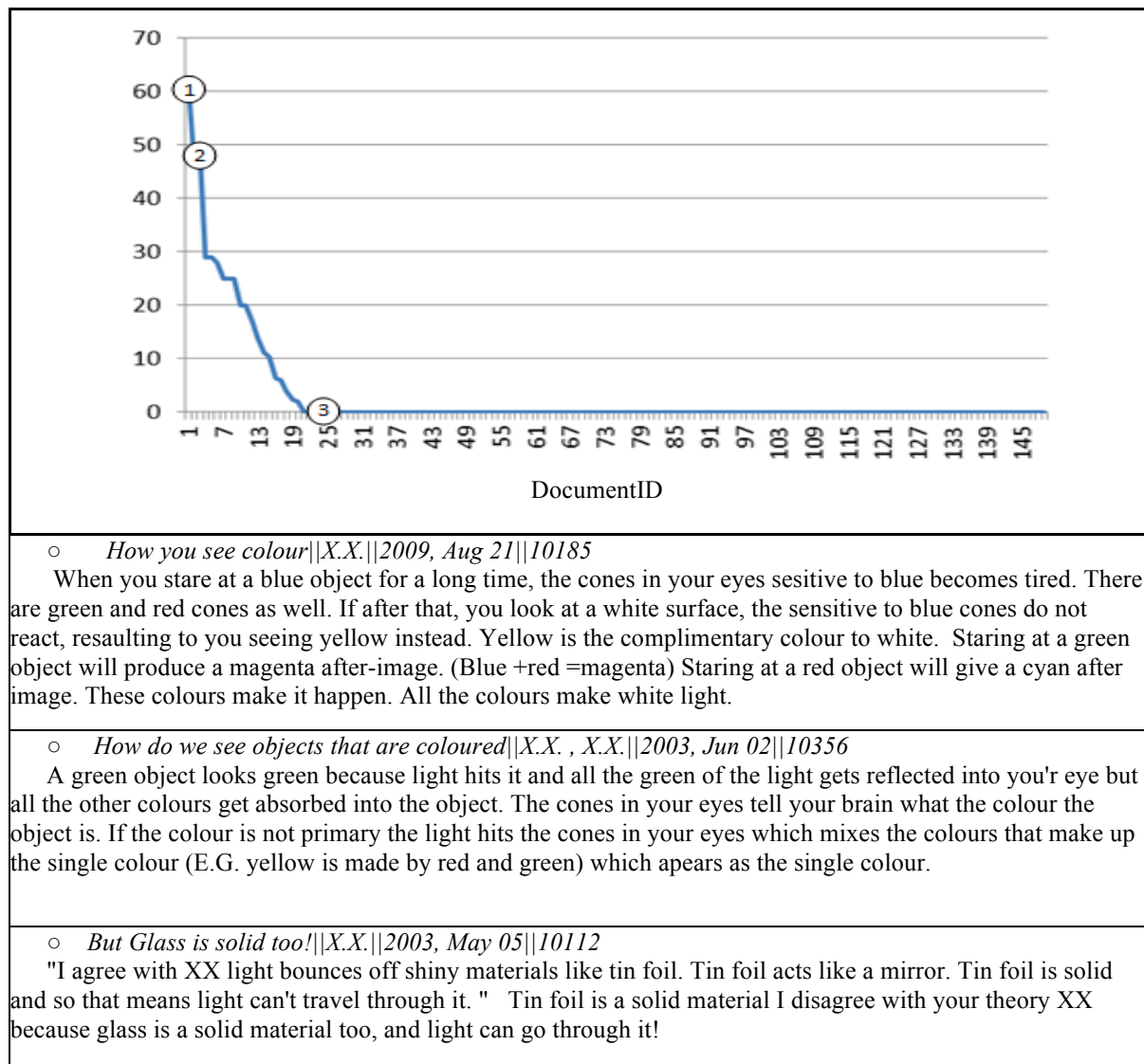


Figure 1. Assignment of Topic 0 over all documents

This diagram plots how likely each document belongs to Topic 0. The circled numbers, 1, 2 and 25 mark the position of the likelihood of the three documents shown under the graph.

Comparing Document-Topic Assignment with LDA against Human Coding

The results reported above show how the LDA algorithm can extract topics from online discourse and assign interpretable topics to each discourse entry. To further gauge the accuracies of these topic assignments, we compared the LDA assignments with those obtained with manual coding.

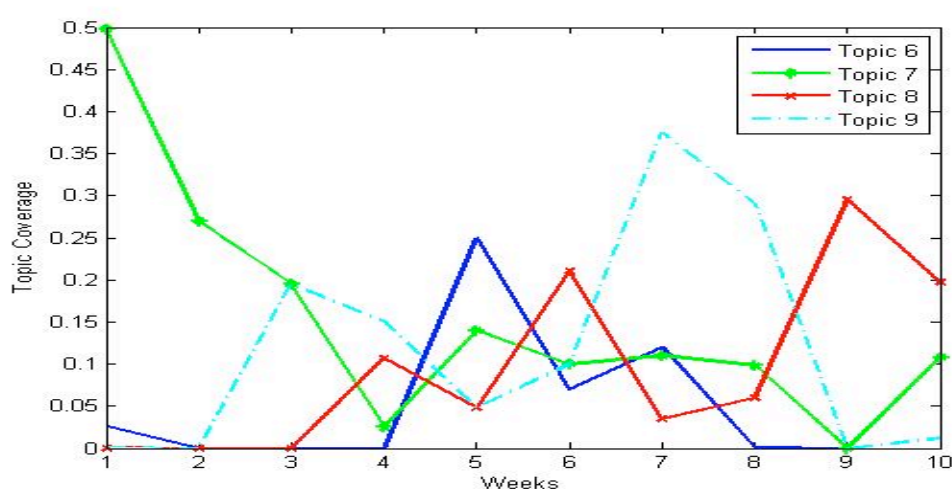
The evaluation process was as follows: we randomly selected five of the ten topics and pool the top six documents from each topic. The order of the documents was then randomized. Two human raters (two of the authors) independently read each of the thirty documents and rated the relevance of each documents to the five topics using a 7-point Likert scale (from 0-definitely not related to 6-definitely related). We then compared the algorithm's topic assignments against the average of the human raters' results.

We used two evaluation metrics: Fleiss Kappa (Fleiss, Levin & Paik, 2013) and normalized Discounted Cumulative Gain (nDCG) (Järvelin & Kekäläinen, 2002). The Fleiss Kappa score describes how much the rater/algorithm's answers differ from random assignment of topics. Hence, the Fleiss kappa tells us the percentage of agreement gained beyond chance. We used it to evaluate the agreements on the assignment of the most relevant topics to the documents. Kapper for inter-rater agreement is 1, and for system-human consistency

is 0.87. The nDCG measure compares the similarity between two lists, range from 0 to 1. The nDCG evaluation metric is suitable for our testing as it captures both the importance of the ranking and the numerical gains that our human annotators assigned. Considering that our system outputs at most 2 topics for each document (in fact, for only 4 documents, our system output 2 topics, and it assigned only 1 topic for the rest of the documents), we only calculated the result for the selecting the most relevant 1 topic, and the most relevant 2 topics. For the most relevant topic, nDCG (averaged over all 30 documents) for inter-rater agreement is 1, and for system-human consistency is 0.90. For the two most relevant topics, the inter-rater agreement in terms of nCDG (averaged over all 30 documents) is 0.99, and the system-human consistency is 0.86. These results indicate that the topic assignment generated by the LDA algorithm achieved an acceptable agreement with human judgment, even though the agreement is lower than that between the two human coders.

Analyzing Temporal Evolution of Different Topics in the Online Discourse

We further tested how LDA may be used to generate useful analysis and feedback data for educators and researchers by examining the progressive changes in student online discourse. Figure 2 shows the evolution of four topics over the 10-week period of inquiry. The x-axis represents time in term of weeks (week 1 – 10), and the y-axis shows how prominent the topic is in that week's discussion (accumulated γ scores for all the posts within given week). For the sake of clarity, we only plotted the scores for four topics in Figure 2.



Topic 6	'sun' 'when' 'earth' 'moon' 'eclipse' 'shadow' 'other' 'world' 'around' 'line' Eclipses and seasons	Eclipses and seasons
Topic 7	'white' 'snow' 'colour' 'prism' 'black' 'melt' 'when' 'see' 'fast' 'why'	White light and snow
Topic 8	'shadow' 'object' 'made' 'opaque' 'energy' 'part' 'call' 'umbra' 'what' 'go' Shadows and light sources	Shadows and light sources
Topic 9	'through' 'go' 'can' 'reflect' 'tin foil' 't' 'think' 'was' 'angle' 'when'	Reflective materials

Figure 2. Temporal projections of topics

The temporal progress of the topics indicates many interesting aspects of the learning process. For instance, Topic 7 (snow and white light) has a dominant score during the first week, and decreases over the next few weeks, then rises again in week 5. The intensive discourse about this topic in the first week as detected by LDA coincides with what actually happened in the classroom: at the beginning of the light inquiry, an early spring snow triggered students' interest in why snow is white and what would happen if it were black. These issues became the primary focus in the first week in the online and face-to-face activities, with 9 notes written in the first week. These issues related to snow became less central in the following three weeks as the knowledge building community formulated other, deeper themes of inquiry to address a wide range of optical issues. These deeper issues were represented by the increase of Topics 8 and 9 in Figure 2 as well as the inquiry threads on primary and secondary colors, light travels and interacts with different materials, mirrors and reflection, shadows, and how we see things. Interestingly, the inquiry on primary and secondary colors and how we see colors led students to understanding the nature of white light and how it can be split into different colors using a prism, with a number of new notes written in week 5-10, as detected by the LDA analysis.

Discussion

This study of applying LDA to the online discourse data showed promising results. The algorithm was able to detect semantically meaningful topics even when the target corpus has a relatively limited vocabulary and short notes and addresses coherently connected issues in one content area (optics) instead of different areas (e.g. optics vs. chemistry). The ten topics discovered by LDA reflect core optical issues investigated by the knowledge building community. These topics overlap substantially with —although do not map directly onto— the inquiry themes identified by students and inquiry thread topics identified by researchers.

An important aspect in applying LDA is to determine the possible number of topics before running LDA. In the current study, we identified the appropriate number of topics by experimenting with varying number of topics based on a possible range informed by the reflection of students on their collective knowledge themes as well as the manual coding of researchers. The researchers in this study implemented detailed coding of notes using inquiry threads to understand the process of knowledge building and benchmark automated analysis results. However, in future application of LDA analysis, such manual review can be handled more quickly to identify major topics of discussion without note-by-note coding. With an overall sense of the possible topics discussed, the analyst—researcher or educator—can vary the total number of topics to used in the LDA algorithm and interpret the obtained candidate topics by examining the keywords involved, with keywords shared between different topics showing the conceptual connections. It is important to point out that most of the topics detected in this study are hard to interpret accurately based on stand-alone keywords without looking at how these words are used in the context of the corresponding discourse entries. Future design of LDA-base assessment tools should show the keywords of each topic in context (e.g. sentences) to aid user interpretation.

Using the LDA algorithm, we further identified documents that are likely associated with each of the ten topics detected. Comparing the document-topic assignment by LDA against the judgment of two human coders resulted in high-level consistency measured using nCDG and Kappa. Such automated document-topic assignment may help students and teachers to identify important discourse contributions addressing particular focal themes as a way to review collective progress and individual contribution. It may also aid researchers in coding discourse entries based on conceptual topics to understand the evolving scope and depth of inquiry in the online discourse. Meanwhile, it is important to mention that our evaluation of automated document-topic assignment was based on a small sample of most relevant documents for five topics, with a high level of accuracy. The accuracy may decrease when assigning other documents that have lower relevance (i.e. likelihood) scores; in such cases, LDA may provide a list of candidate documents for the human analyst to choose from, so the analyst can select relevant documents for each topic more easily.

We also noticed a few challenges, which suggest headroom for further improvement of the technology. The unsupervised-learning nature of the LDA algorithm provides researchers little control over the granularity and structure of the detected topics. The algorithm may generate un-interpretable topics and it is difficult to estimate the interpretability of the generated topics without human interference. Sometimes the analyst may want to see a finer-grain analysis of certain topics but it is challenging to control the granularity of the topics detected by LDA. To address these drawbacks, future research may look into developing topic models that can incorporate structure information of the online discourse (e.g. build-on links) and information from manually-coded archived discourse on the same topic. Also, we will also explore the use of hierarchical topic models that can discover topics with intrinsic hierarchical orders to solve the granularity issue. We believe this further development will give the educator and analyst more control over the discovered topics for the purpose of more effectively and efficiently analyzing large volumes of online discourse data, and pave the road for creating auto-assessment and feedback tools to leverage sustained and productive knowledge-building discourse.

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 Data Mining*, 1(1), 3-17.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Chen, M.-H., Zhang, J. & Lee, J. (2013). Making Collective Progress Visible for Sustained Knowledge Building. 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* (pp.81-88). International Society of the Learning Sciences.
- Chafe, W. (1997). Polyphonic topic development. In T. Givón (Ed.), *Conversation: Cognitive, Communicative and Social Perspectives* (pp. 41-53). John Benjamins Publishing.
- Dringus, L. P. and Ellis, T. (2005). Using data mining as a strategy for assessing asynchronous discussion forums. *Computers & Education*, 45(1):141-160.
- Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical methods for rates and proportions*. Hoboken, NJ: John Wiley & Sons.
- Hmelo-Silver, C.E. (2003). Analyzing collaborative knowledge construction: Multiple methods for integrated understanding. *Computers & Education*, 41, 397-420.

- Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine learning*, 42(1-2), 177-196.
- Järvelin, K., & Kekäläinen, J. (2002). Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems*, 20(4), 422-446.
- Law, N., J. Yuen, Huang, R., Li, Y. & Pan, N. (2007). A Learnable Content and Participation Analysis Toolkit for Assessing CSCL Learning Outcomes and Processes. In C. A. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Proceedings of the 8th International Conference on Computer Supported Collaborative Learning* (pp. 411-420). International Society of the Learning Sciences.
- Meier, A., Spada, H., Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration process. *International Journal of Computer-Supported Collaborative Learning*, 2, 63-86.
- Ming, N., & Ming, V. (2012). Predicting student outcomes from unstructured data. *Proceedings of the 20th Conference on User Modeling, Adaptation, and Personalization*. Montreal, Canada
- Mu, J. Stegmann, K., Mayfield, E., Rose, C., & Fischer, F. (2012). The ACODEA framework: Developing segmentation and classification schemes for fully automated analysis of online discussions. *International Journal of Computer-Supported Collaborative Learning*, 7, 285-305.
- 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 CSCL. *International Journal of Computer-Supported Collaborative Learning*, 3(3), 237-271.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Eds.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York, NY: Cambridge University Press.
- Sherin, B. (2012). Computing student science conceptions with Latent Dirichlet Allocation. *Paper presented at the 10th International Conference of the Learning Sciences*, Sydney, Australia.
- Sherin, B. (2013). A Computational Study of Commonsense Science: An Exploration in the Automated Analysis of Clinical Interview Data. *Journal of the Learning Sciences*, 22, 600-638.
- Stahl, G. (2006). Group cognition: Computer support for building collaborative knowledge. Cambridge, MA: MIT Press.
- Sun, Y., Zhang, J., & Scardamalia, M. (2010). Knowledge building and vocabulary growth over two years, Grades 3 and 4. *Instructional Science*, 38(2), 247-271.
- Teplov, C., & Fujita, N. (2009). Determining curricular coverage of student contributions to an online discourse environment through the use of latent semantic analysis and term clouds. *Proceedings of the 9th international conference on Computer supported collaborative learning* (Vol. 2, pp. 165-167). International Society of the Learning Sciences.
- van Aalst, J., & Chan, C.K.K. (2007). Student-directed assessment of knowledge building using electronic portfolios in Knowledge Forum. *Journal of the Learning Sciences*, 16, 175-220.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, 46, 71-95.
- Zhang, J. (2004). The growing networks of inquiry threads in a knowledge building environment. Paper presented at the Knowledge Building Summer Institute. Ontario Institute for Studies in Education, University of Toronto.
- Zhang, J., Lee, J., & Chen, J. (2014). Deepening Inquiry about Human Body Systems through Computer-Supported Metadiscourse. Paper presented at the Annual Meeting of American Educational Research Association. Philadelphia, PA.
- Zhang, J., & Messina, R. (2010). Collaborative productivity as self-sustaining processes in a Grade 4 knowledge building community. In K. Gomez, J. Radinsky, & L. Lyons (Eds.), *Proceedings of the 9th International Conference of the Learning Sciences* (pp. 49-56). Chicago, IL: International Society of the Learning Sciences.
- Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of nine- and ten-year-olds. *Educational Technology Research and Development*, 55(2), 117-145.
- 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.
- Zhang, Y., Law, N., Li, Y., and Huang, R. (2012). Automatic extraction of interpretable topics from online discourse. In *The International Conference of the Learning Sciences (ICLS) 2012, Volume 1*, (p.443-450). International Society of the Learning Sciences.

Developing an Orchestrational Framework for Collective Inquiry in Smart Classrooms: SAIL Smart Space (S3)

Mike Tissenbaum and James D. Slotta, University of Toronto, Toronto, ON, miketissenbaum@gmail.com

Abstract: This paper describes PLACE - a 12-week cross-contexts curriculum for grade 11 physics that engaged students at home, in class, in their neighbourhoods, and in a smart classroom setting. Using a design-based research approach we introduce a smart classroom infrastructure (SAIL Smart Space; S3) and investigate its role in supporting students in the curriculum as a knowledge community. The present paper focuses on the culminating smart classroom activity, where students use the community knowledge base to scaffold their solving of ill-structured physics problems involving popular Hollywood movies. We examine the efficacy of the tools and the environment, including software agents and data mining approaches that serve to define S3, and help orchestrate the flow of activities, materials, and students during the activity's enactment. We conclude with a set of design principles that support collaborative inquiry in smart classrooms and across learning contexts.

Supporting Knowledge Communities and Inquiry Learning

Within the field of learning science there is a growing call to think of classrooms as knowledge communities where students approach learning as a collective endeavor (rather than an individual one), towards solving authentic and personally relevant real-world problems (Slotta & Najafi, 2010). The knowledge community approach, with its focus on user-contributed content and student driven inquiry, is particularly well suited for investigating learning that goes beyond the traditional confines of classroom settings. For example students may visit a local stream or waterway to investigate issues around water quality (Hsi, Collins & Staudt, 2000), or a playground to investigate geometry principles (Milrad et al, 2013). Within this work there is a corresponding challenge of connecting this learning across contexts back into the classroom in meaningful ways.

In response, a promising new approach to supporting such designs is the digital embedding of aspects of the community's inquiry in the walls, ceilings, and floors of the physical learning environment. Examples include a simulated ecology of bugs living in a classroom's walls (Moher, 2006), and evolutionary simulations of rainforest fauna and flora over millions of years (Lui & Slotta, 2013). By connecting students to these inquiry environments with personal, portable and connected computing devices (such as tablets), we can unchain them from traditional classroom configurations, instead fostering a dynamic "smart classroom" model where students can move throughout the room engaging in spatially index real time collaborative activities (Slotta, 2010).

The enactment of complex real-time inquiry activities places a high load on teachers, requiring them to simultaneously manage changing student roles and groups, assign activities, and organize materials – including potentially large and diverse community-generated content from the knowledge base (Tissenbaum & Slotta, in press). The supporting of teachers and students in the enactment of such activities is often termed orchestration (Dillenbourg, Jarvella, Fischer, 2009), and has been highlighted as a major design challenge in the learning sciences (STELLAR, 2011). In response, technology enhanced learning spaces like smart classrooms, may offer a means for supporting the orchestration of such curricula, including tracking student movement within the room, providing procedural scaffolds, and making real-time decisions including student group formation and the delivery of timely materials (Tissenbaum & Slotta, 2013).

In order to investigate the role that a smart classroom could play in supporting a knowledge community curriculum, we needed both a pedagogical model and an underlying technology infrastructure. The pedagogical approach we used for this study is the Knowledge Community and Inquiry (KCI) model, which engages students to work collectively, contributing, tagging and improving content in a shared knowledge base that serves as a resource for subsequent inquiry. In KCI, inquiry activities are carefully designed so that they engage students with targeted content and provide assessable outcomes, allowing students some level of freedom and flexibility but ensuring progress on the relevant learning goals (Slotta & Najafi, 2013). The teacher also plays a critical role within a KCI curriculum, as they must be able to adapt (orchestrate) the designed script "on-the-fly," based on emergent themes and community voices (Tissenbaum & Slotta, 2013).

SAIL Smart Space (S3) – a Technology Framework for Smart Classrooms

In order to successfully enact the kinds of complex designs required for KCI, we needed a flexible and adaptive infrastructure to support the design and orchestration of collaborative activities that include spatial, social, and semantic dependencies. To this end, we developed SAIL Smart Space (S3), an open source framework that coordinates complex pedagogical sequences, including dynamic sorting and grouping of students, and delivery of materials based on emergent semantic connections. S3 allows the physical space of classrooms or other learning environments to play a meaningful role within the learning design – either through locational mapping

of pedagogical elements (e.g., where different locations are scripted to focus student interactions on different topics) or through orchestrational support (e.g., where physical elements of the space, like projected displays), help to guide or coordinate student movements, collaborations or activities. S3 was developed to add a level of intelligence to classrooms or other learning environments, including real-time data mining and computation performed by intelligent agents to support the orchestration of inquiry scripts. A key component to these agents is that although their roles are well defined within the activity (e.g., grouping students with peers they have not previously worked with), who (or what) will satisfy these conditions cannot be known *a priori*, rather they emerge during enactment (requiring agents to process information in real-time; Tissenbaum & Slotta, in press).

In order to investigate S3 as a smart classroom infrastructure to support collaborative inquiry, three questions have driven our research: (1) How can S3 support students' inquiry activities? (2) How effective are intelligent software agents and data mining in providing students with needed materials and enacting specific pedagogical moves? (3) How does S3 support the teacher in orchestrating class activities? Below, we discuss an implementation of S3 within a grade 11 physics classroom, highlighting specific orchestrational supports (including the design of intelligent agents), and evaluating their effectiveness for supporting classroom inquiry.

Methods

We employed a co-design methodology (Roschelle, Penuel & Shechtman, 2006) working closely with a high school physics teacher to ensure that he was an active voice in the design and that the designed intervention fit his goals for the students and expectations for student learning. As the study took place in a "real class" (rather than a canned lab) a design based-research approach was employed (Wang & Hannafin, 2006). Generally, design-based research does not attempt to validate a particular curriculum, rather it strives to advance a set of theories on learning that transcend the particulars in which they were enacted (Barab & Squire, 2004). As such, a major outcome of this research was the design of the curriculum and supporting technologies themselves. To evaluate the enacted design we used a mixed methods approach to triangulate the data and get a more complete picture of the intervention (Mason, 2006). Sources included pre- and post-interviews with teacher and students, server logs, the user-contributed artifacts, and video and audio recordings.

Design of the KCI curriculum

In order for the smart classroom to be more than a supplemental activity, we needed to develop a complete curriculum in which the smart classroom was one of several learning contexts, integrated with activities across classroom and home settings. We also wanted to investigate how the smart classroom could leverage student-contributed content for authentic learning activities, and we therefore also needed the curriculum to produce artifacts that could be reused in meaningful ways. To this end, we worked closely with the high school Physics teacher to develop a curriculum that implemented KCI, including collaborative and collective forms of inquiry and critical reflection. In initial co-design meetings with the teacher he identified two main goals for the curriculum: First he wanted to help students recognize "physics in their everyday lives" and bring this view back into the traditional class setting; Second, he wanted a way for students to develop a coherent understanding of the underlying physics principles of the course, and connections amongst them (i.e., to "see that all the principles are tied together"). In response to the second goal we began by generating set of fourteen "core" principles (Table 1) that the teacher felt were of core relevance to the content of the course.

Table 1: Grade 11 fundamental principles for kinematics, force and motion, & work, energy, and power

• Vectors	• Acceleration	• $F_{net} = 0$	• Kinetic Energy
• Newton's First Law	• Uniform Motion	• $F_{net} = \text{Constant (non-zero)}$	• Potential Energy
• Newton's Second Law	• Kinetic Friction	• $F_{net} = \text{non-constant}$	• Conservation of Energy
• Newton's Third Law	• Static Friction		

We then co-designed a 12-week grade 11 physics curriculum, named PLACE (Physics Learning Across Contexts and Environments), which guided students to capture examples of physics in the world around them (through pictures, videos, or open narratives). PLACE was organized into three major units covering the first three units of the course: (1) Kinematics, (2) Forces and Motion, and (3) Work, Energy, and Power. For each unit, we developed a script that allowed students to upload, tag, and interact with physics examples.

For each unit in the curriculum students were required to upload their own found examples of physics principles (i.e., from their everyday life experiences) to a shared social space, including tags and explanations according to the relevant physics principles. The wider class community was encouraged to respond to these user-contributed artifacts: debating tags or explanations, voting, and adding new tags – with the stated aim of developing consensus about each item. Students were required to upload at least one example and to complete the *idea refinement script* on at least two of their peers' examples: (1) voting on existing tags and/or adding a new tag, (2) voting on the contributions of their peers, and (3) adding a reflection or rationale of their own. This ensured that students focused on the principles, reflected on the work of peers, and added their own thinking.

Students were also encouraged to upload their class lab reports, to gain feedback on methodologies and results and to connect them to the contributed examples. The teacher also developed homework multiple-choice problems, which the students solved using a script similar to the *idea refinement script*, where they tagged, answered and reflected on the problem. In-class, students were given further inquiry tasks, and asked to develop “challenge problems” for their peers, which built upon examples from the community’s knowledge base.

The products of these various inquiry activities became a dynamic knowledge base used within the community from one part of the overall unit to the next. This knowledge base, called *PLACE.web*, served as a resource for a culminating smart classroom activity where students applied what they had learned in the three units to solve ill-structured physics problems relating to selected Hollywood movie clips. The smart classroom environment, detailed below, was called *PLACE.neo*, and provided an important technology enhanced environment in which we could develop and test the orchestrational framework that is at the heart of this paper.

Developing S3: Orchestrating Cross-Context Inquiry in PLACE

To help students work as a knowledge community at home, in their neighbourhoods, and in class, we needed to develop a technology environment that facilitated the various interactions and knowledge contributions required by our curriculum. Consistent with Scardmalia and Bereiter’s (2006) argument that the technology environment (in their case, the Knowledge Forum) should be developed to support the particular epistemic and pedagogical forms within the research (in their case, knowledge building), we needed to develop specific software tools to support our own designs around KCI. S3 was developed to support student interactions, data structures, and pedagogical flow, including two specific software systems: *PLACE.web*, a collaborative social network, which supported students in contributing and engaging with the products of their teacher and peers; and *PLACE.neo*, a smart classroom environment in which the culminating Hollywood physics activity was orchestrated. Below we describe both learning environments, their role in supporting student interactions across context and as a knowledge community, and the role of intelligent data mining and software agents in facilitating these actions.

PLACE.web

The *PLACE.web* learning environment consisted of five distinct interaction spaces for students in support of the scripted activities described above: (1) The student status page, which showed aggregated newsfeeds that provided students with updates on the community knowledge and their own contributions; (2) The contribution upload page, where students uploaded their examples; (3) The user contribution discussion pages; (4) The assigned homework pages; (5) And the “Associative Web”- a visualization of the community knowledge, showing a complex interconnected web of social and semantic relations, where students could filter the resources to match their interests or needs.



Figure 2. The smart classroom with (1) An interactive whiteboard that orients students towards a specific Hollywood scenario, aggregates scenario specific student contributions and facilitates idea negotiation; (2) A second board with a different scenario; (3) Individual tablets provide students instructions, allow them to access the knowledge base, and contribute ideas to the shared display; and (4) An ambient display showing where students are in the room, completed tasks and time left in the current task.

PLACE.neo

An important goal of this project, in terms of developing S3, was to create a culminating activity (i.e., the Hollywood physics activity) where students needed to use materials from earlier co-constructed knowledge as a resource for an inquiry project. This would support our design of S3 intelligent agents and other features that could gather and distribute materials, assign student groups, and coordinate the pedagogical flow of activities. The overall script for the Hollywood physics activity involved three micro-scripts that spanned home, classroom, and smart classroom settings, At home, students were tasked with reviewing a collection of problems drawn from the proceeding 12 weeks (including student generated challenge problems), verifying the principle tags that had been applied by their peers, and adding equations that might be used in solving the problems. In

class, students worked in small groups to achieve consensus on a “final set” of the tags and equations for each problem. The resulting set of problems, tagged with principles and equations, were then used as raw materials for the final smart classroom script, where students would be orchestrated in a complex sequence of rotating small group assignments to set up and solve the Hollywood physics problems.

When students entered the smart classroom they were presented with four scenarios drawn from popular Hollywood movies (e.g., from the movie Ironman, a short scene where he survives a ballistic fall to earth). Each video was shown in one quadrant of the room (Figure 2) on a large display, together with other information added by students progressively throughout the activity. The activity was broken into four different steps: (1) Principle Tagging; (2) Principle Negotiation and Problem Assignment; (3) Equation Assignment, and Assumption and Variable Development; and (4) Solving and Recording. In each of these steps, intelligent agents played a role in assigning students to a new group (i.e., “move from video A to video D”) and coordinating the inquiry activities in the next step. This included distribution of materials (e.g., presenting the group with a set of equations that were datamined according to the tags the previous group had added), communications with the teacher, and coordination of group activities (e.g., consensus or voting).

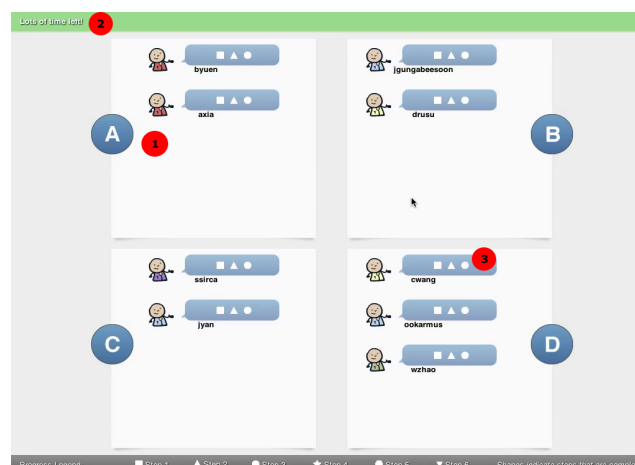


Figure 3: The ambient display (1) tracked each student within the room - when students moved location (or were sorted by an agent) their avatars moved on the display; (2) the timing of activities was tracked using a colored bar at the top of the display, which moved from solid green to flickering red time ran out; and (3) updated student progress by displaying an icon next to their avatar on the completion of each task.

The physical design of the room was a major consideration in supporting classroom orchestration. Similar to successful Problem-Based Learning (PBL) approaches, we wanted to support the teacher as a “wandering facilitator” (Hmelo-Silver, 2004). To this end, in a manner similar to the use of post-it boards in PBL, we placed the products of inquiry (the videos and collaboratively constructed student work) on the wall-mounted displays (Figure 2). This wandering facilitator model, which includes the demands of managing student groupings, monitoring the timing of tasks and supporting their inquiry with resources from the knowledge base, produces high levels of “orchestrational load” (Dillenbourg, 2012). To reduce this load we developed several technical supports. First, following a recent method shown to be effective (Alavi et al., 2009), we designed an ambient display for the front of the smart room (Figure 3). A teacher tablet was also developed, iterating on observations and feedback from previous designs (Tissenbaum & Slotta, 2012), from a device that showed student work *post hoc* (which the teacher could not act upon), to a regulatory orchestrational tool. This version of the teacher tablet showed him where each student was within the script and alerted him when he was needed to approve a group’s work (at the end of Step 3). The goal of these devices (the display and the tablet) was for them to sit at the periphery of the teacher’s awareness so that they only needed to be attended when necessary (Ishii et al., 1998).

Another critical support in managing the orchestrational load of the teacher was the use of intelligent software agents and real-time datamining, which were able to respond to the emergent class patterns in order to make on-the-fly decisions. During the smart classroom activity, four agents were implemented: (1) *The Sorting Agent* sorted students into groups and assigned room locations first (between Step 1 and Step 2) based on the frequency of their tags at each board during Step 1; and the second (between Steps 2 & 3) based on placing them students they had not previously worked with; (2) *The Consensus Agent* monitored groups requiring consensus to be achieved among members before progression to the next step; (3) *The Bucket Agent* coordinated the distribution of materials to ensure all members of a group received an equal but unique set of materials (i.e., problems and equations in Steps 2 & 3); and a *Student Progress Agent* which tracked individual, small group, and whole class progress to send status updates to other devices (e.g., the teacher orchestration tablet).

To support student participation the smart classroom activities, we developed tablet applications and interactive displays. Students' personal tablets facilitated their login at each zone, provided task specific materials and instructions, and collaboration support. Any student work done on his/her tablet while in a zone (e.g., adding tags, suggesting variables) instantly appeared on the zone's interactive display, which also held a persistent representation of all of the student-negotiated contributions. These displays were also used during group negotiation phases (e.g., deciding which equations would help in solving the video), where students physically dragged the individually suggested items into "Yes" or "No" boxes until consensus was reached.

Results and Discussion

In evaluating PLACE, two forms of analysis are important: 1) the inspectable artifact of PLACE, including the tools and the environment, including software agents and data mining approaches, that serve to define S3, 2) the student produced artifacts and collaborations that characterize learning in a KCI model. The present paper focuses on the first analysis, confirming that PLACE was enacted according to its design, specifying the technology elements that allowed such an enactment. One outcome of this work will be a set of design principles that support collaborative inquiry in smart classrooms and across learning contexts.

Student Contributed Content

In total, 169 student examples were created, with 635 total discussion notes contributed around those examples. Students also attached 1066 principle tags to the examples and cast 2641 votes on those tags. On average, students uploaded 3.84 examples, which is greater than the requirement of 3 (one per unit), suggesting active community engagement. Figure 4 shows the times of day at which students contributed to the knowledge base, with uploads or comments made at nearly every time except 3am to 6am. This highlights the ability of PLACE to seamlessly connect students within their overall community whenever they felt the desire to take part, with 53% of the contributions taking place outside of school hours (4pm-9am).

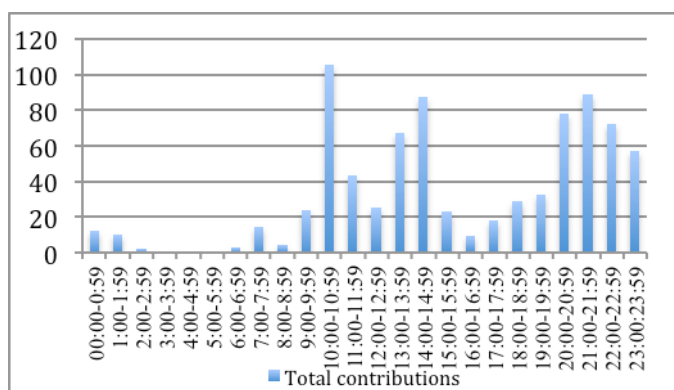


Figure 4. Student contributions to PLACE.web by time of day

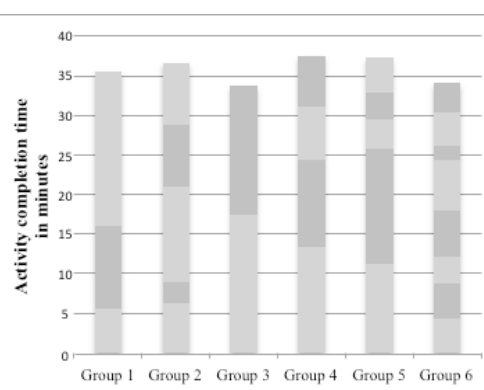


Figure 5. How many problems Bucket Agent sent to each group, and time spent on each.

During the "challenge problem" script, student groups (n=3 or 4) were tasked with developing problems for the class to use as review, drawing from the wider knowledge base of peer-contributed examples. The script successfully engaged students in leveraging the collective knowledge base towards developing new objects for peer engagement and investigation. In total, 13 challenge problem were developed by the groups, each referring to on average 2.23 examples from the knowledge base. The Associative Web employed the underlying S3 data mining structures to support students in the activity, helping them find examples that matched their expertise groups and supported their creation of challenge problems. Post-interviews indicated that students found the Associative Web very useful for filtering the overall knowledge base and finding artifacts that matched their search criteria, noting that it *"made it clear what examples are related to our concepts, because you could see what example was related to more than one of the concepts, and it's easy to browse through multiple areas."*

S3 Orchestration Across Contexts: The Hollywood Physics Culminating Activity

Within the culminating activity, students were able to access, contribute to, and use the knowledge base at home, in class, and in the smart classroom. The activity was an important test of S3's capabilities to support seamless orchestration of learning activities, and the intelligent software agents and data mining were central to our success. The following sections address how S3 supported the learning within and across these contexts.

Orchestration of the at-home activity: At home students were scaffolded in answering a subset of problems from the knowledge base depending on their “expertise groups” from the previous units. S3 agents were used to ensure that students representing all fourteen principles saw each problem.

Orchestration of the in-class activity: In-class, the S3 agents successfully grouped students and facilitated consensus building on homework problems. Two items were of particular interest to this study. First, was the ability of S3 to capture the individual student work (i.e., the assignment of principles and equations), and to re-visualize them as aggregates for the in-class activity. This highlighted S3’s ability to effectively contextualize artifacts from the knowledge base based on the scripted activity. Second was the effectiveness of the *Bucket Agent* in orchestrating the real-time distribution of problems to individual groups. As soon as a group finished a problem, the *Bucket Agent* dipped into the “bucket” of problems and sent the group a new one, until the bucket was empty. This allowed a large collection of resources to be distributed to groups as they worked in parallel, such that all the resources were attended to in a single 60-minute class and every group finished within 3 minutes of each other (Figure 5). In this way, agents were able to accommodate variations in resource difficulty and group skill levels (i.e., some resources were more challenging and some groups were quicker).

		Teacher Orchestration Move						
		Check Status	Explain Task	Start Task	Check Timing	Check/Assign Location	Clarify Task	Approval Task
Orchestration Technology	Teacher Tablet	11	4	17	0	0	0	6
	Ambient Display	2	0	1	12	10	0	0
	Student tablet	8	6	0	0	0	1	0
	Large Format Display	2	0	0	0	0	11	0
	None	4	6	0	2	4	3	0

0	1-5	6-10	11-15	16-20

Figure 6. Frequency of teacher orchestration moves

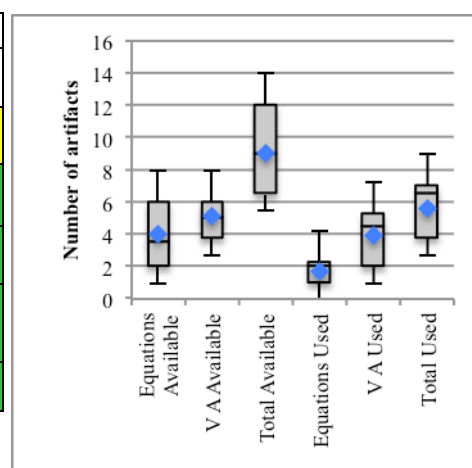


Figure 7. Negotiated Equations, Variables and Assumptions (VA) and their use in solution.

Orchestration of the smart classroom activity: The S3 agents successfully responded to emergent student actions (i.e., actions that were not known or knowable beforehand), drawing on semantically relevant artifacts from the community knowledge base to support follow-up tasks. During Step 2, the *Bucket Agent* supplied each group with problems whose principles matched those that had been negotiated by students for their scenario. The agents connected, on average, 23 problems to each scenario, distributed evenly across all group members, who then promoted 3.4 problems on average; these were negotiated down to 2.6 on average. During Step 3 (“Equation Assignment”), the *Bucket Agent* drew equations from the knowledge base, which had been tagged in the pre-activity, to the negotiated problems, to serve as a further resource for students. From these agent-filtered equations, students recommended an average of 4.9 equations, which were negotiated down to 4.3 on average.

The *Sorting Agents* also successfully sorted students based on the frequency of the tagging after Step 1, and the condition of working with new group members after Step 2. The ability of these agents to sort students based on pre-defined pedagogical goals where the students who would fit their conditions could not be known *a priori* was an exciting outcome. Grouping and moving students is a complicated and time intensive task in any classroom, and being able to automate it *and* include the processing of emergent patterns (an impossible task for a human in real-time) provided critical support for managing of orchestrational load. The success of the sorting agents used in PLACE holds promise for the design of more complex agents, and the scripts that employ them. “It was such a paradigm shifting kind of lesson, with the pacing and, the kinetics, and the motion in the room and kids moving around was a lot to follow. [But] I didn’t need to worry about it, it was just taken care of by the various technologies” (Teacher).

Video analysis of the teacher during the smart classroom activity provided insight into the orchestrational support provided by the various S3 technologies. All the teacher’s orchestrational “moves” were time stamped and coded for use of any orchestrational technology (e.g., looking at the ambient display – Figure 6). This revealed the potential efficacy of certain technologies to support certain tasks over others – such as the ambient display to check student locations, and the teacher tablet for checking group status. Exit interviews with teachers and students support the effectiveness in these technologies for support the enactment of complex smart classroom activities and in managing orchestrational load. “With the board it was like this is where we have to

go and that's how much time we have left so we didn't need the teacher for that any more. He could just focus on going around and talking to the groups" (Jen).

Analysis of the student-generated final products and server logs indicate that S3 was effective in supporting student inquiry. We compared the groups' final answer sheets with their co-developed evidence on the large-format displays and found, on average, groups used 54.6% of the assigned equations and 76.8% of the assigned variables and assumption (Figure 7 above). We were curious why (especially when compared to variable and equations) the percentage of equations was so low. Exit interviews indicated students preferred to keep more equations on hand (in their tool belt) until they were sure which they would use. "If [we were] not totally sure like it's a grey area we would put it in "yes" just in case" (Sarah).

In exit interviews, students indicated that the interactive large-format displays were effective in supporting their co-construction of evidence and in supporting them in approaching the final problem-solving step (Step 4). "Just looking at what other groups had left us you got a good sense, and then from there the group could take over and be like this is what we need to do to solve it" (Robert).

Comparing the interactions around the board with regular class activities, students indicated a greater sense of collaboration and involvement in discourse and idea generation. "In class if a teacher were to tell [a group] to solve a problem together then [laughs] I would say that rarely everyone participates, and there are one or two people who are just not doing anything, but in here it really engaged us to participate" (Rebecca).

Cross-Context Learning: Factors and Design Principles

An important outcome of this research is a description of how our design supported a knowledge community – not only *within* multiple contexts, but also in the transitions *between* contexts. This work has led to the following high-level design principles for cross-context learning. While these principles are not the only ones for supporting cross-context learning, they are grounded in a successful design and offer concrete guidance in support of future designs for cross-context learning that include individual and collaborative inquiry activities, at home, on field trips, in-class, or in smart classrooms. These principles are situated within three important aspects of cross-context learning:

Visualizations of Community Knowledge

In PLACE we needed to consider the transition from the *individually collected examples to collaborative online activities*. We wanted student inquiry to be the focus of the activities (rather than materials found in textbooks or other professionally curated materials). Our main challenge was finding ways for the students to effectively search a large repository of student artifacts to find materials that fit their specific needs. In response to this we built the Associative Web, allowing students to mine artifacts based on their assigned tags, and present them in a ways that were useful to their given context.

At home, we wanted students to see how *in-class activities* affected their own contributions to the knowledge base. This was the impetus for the aggregated newsfeeds, which leveraged system generated metadata about individual students (e.g., which artifacts they had worked on) providing them with contextualized updates and a macro view of the whole class' activities. These different aggregated and filterable views served as a bridge for students to orient themselves within the larger knowledge community.

Design Principle: In order to bridge different learning contexts, visualizations of the community knowledge must present the aggregated information in ways that are relevant to the present context and activity.

Data Structures and Semantic Metadata Supports

An important transition concerned the movement of materials and student roles between the *at-home stage* (i.e., on PLACE.web) and the *in-class stage* (using tablet apps in the classroom). In order for small groups to work on the assignment of principles and equations in class, we needed S3 to aggregate the individual homework responses in ways that fostered collaborative discussion and debate. Because the underlying metadata were semantically well defined (e.g. using tags such as "problems", "principles", "equations"), we were able to easily create views that supported the desired scripted interactions. Metadata also played a significant role in connecting the *in-class* artifacts to the *smart classroom*. In the smart classroom, S3 agents could easily leverage the semantic metadata generated by students during the activity around each video (e.g., their negotiated tags), to query all artifacts that had been similarly tagged in the *in-class stage*. These metadata allowed information to flow seamlessly across contexts and to be repurposed for the particular scripted goals within each context.

Design Principle: To facilitate the organization of student materials for use across contexts, data structures should be defined to support flexible query and representation by students, and access by intelligent agents

The Orchestrational Role of Intelligent Software Agents

In the *smart classroom* we knew that we wanted students to be able to use materials from the *in-class stage* as resources for their problem solving, and that we wanted these materials to be evenly distributed among group

members; however the materials would be needed by each group and the students that would make up those groups could not be known *a priori*. Thus, we needed to develop agents that could respond to emergent class patterns in to support our pedagogical designs – a task that on their own would place a prohibitive level of orchestrational load on the teacher (and students). Similarly, during the *in-class stage* we needed to distribute the aggregated problems completed during the *at-home* activity to the individual groups. Here, the goal here was not that *every* groups saw the same number of problems, rather it was to ensure that every problem was attended to only once in a 60-minute class. The ability of the *Bucket Agent* to quickly assess the state of the activity and draw the required materials from another context in pursuit of the scripted goals provides another layer of adaptive orchestrational support.

Design principle: To help orchestrate sequences of activities, intelligent software agents can be designed to retrieve materials from the knowledge base or execute pedagogical logic in response to real time conditions or emergent patterns within student data.

Conclusions and Future Directions

This study addresses the challenge of supporting a knowledge community that blends rich inquiry practices in the world around them with well-defined pedagogical and curricular goals. We show the promise of smart classroom infrastructures such as S3 for leveraging the products of a knowledge community in support of inquiry activities. Within PLACE.neo, we introduced a role for intelligent software agents to support the orchestration of these activities, by responding to emergent class patterns and enacting pedagogical moves that would be impossible to orchestrate “by hand”. This work demonstrates the potential for smart classroom designs to support both long-term student investigations through persistent portals such as PLACE.web, and in real-time enactments as in PLACE.neo. We believe that going forward, such infrastructures have the potential to create new avenues for knowledge awareness for individuals by connecting them more directly with the work of their peers, and for the broader community by producing unforeseen “rise above” themes for further investigation.

References

- Alavi, H. S., Dillenbourg, P., & Kaplan, F. (2009). Distributed awareness for class orchestration. In *Learning in the Synergy of Multiple Disciplines* (pp. 211-225). Springer Berlin Heidelberg.
- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, (October 2013), 37-41.
- Dillenbourg, P. (2012). Design for classroom orchestration, position paper. In P. Dillenbourg, Y. Dimitriadis, M. Nussbaum, J. Roschelle et al. (Eds.), *Design for classroom orchestration* Computers & Education.
- Dillenbourg, P., Jarvela, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In N. Balacheff, S. Ludvigsen, T. Jong, A. Lazonder & S. Barnes (Eds.), *Technology-enhanced learning* (pp. 3-19). Dordrecht: Springer Netherlands.
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16(3), 235-266.
- Lui, M., & Slotta, J. (2013). Immersive simulations for smart classrooms: Exploring evolutionary concepts in secondary science. *Technology, Pedagogy and Education*, 1-24.
- Milrad, M., Wong, L., Sharples, M., & Hwang, G. (2013). Seamless learning: An international perspective on next generation technology enhanced learning.
- Moher, T. (2006). Embedded phenomena. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '06*, 691.
- Roschelle, J., Penuel, W. R., & Shechtman, N. (2006). Co-design of innovations with teachers : Definition and dynamics. *Proceedings of the 7th International Conference on Learning Sciences*, pp. 606-612.
- 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.), *The classroom of the future orchestrating collaborative learning spaces* (pp. 215-242). Rotterdam: SensePublisher.
- Slotta, J. D., & Najafi, H. (2013). Emerging technologies for the classroom. In C. Mouza, & N. Lavigne (Eds.), *Emerging technologies for the classroom: A learning sciences perspective* (pp. 93-112). Springer.
- Tissenbaum & Slotta (in press) Scripting and orchestration of learning across contexts: A role for intelligent agents and data mining. In Milrad, Wong & Specht (eds.) *Seamless Learning in the Age of Connectivity*.
- Tissenbaum, M. & Slotta J. D. (2013). Scripting and orchestration of collaborative inquiry: An increasing complexity of designs. *Proceedings of the 10th International Conference on CSCL* (pp. 367-368), Madison.
- Tissenbaum, M., & Slotta, J. D. (2012). Scaffolding a Knowledge Community for High School Physics. *Proceedings of the 10th International Conference for the Learning Sciences* (pp. 436-440). Sydney.
- Wang, F., & Hannafin, M. J. (2005). Technology-enhanced learning environments. *Educational Technology Research & Development*, 53(4), 5-23.

Bidirectional Analysis of Creative Processes: A Tool for Researchers

Alecia Marie Magnifico, University of New Hampshire, alecia.magnifico@unh.edu
Erica Rosenfeld Halverson, University of Wisconsin–Madison, erhalverson@education.wisc.edu
Christopher T. Cutler, Independent Software Developer, ctcutler@gmail.com
TJ Kalaitzidis, University of Wisconsin–Madison, kalaitzidis@wisc.edu

Abstract: Young peoples’ creative production combines such activities as analysis of existing forms and genres; participation in processes including drafting, journaling, and modeling; sustained work on a particular idea; and mentor critique. In our work with a variety of settings, we have found that young artists’ progress depends on their engagement in all of these tasks, and that analysis of this work depends on looking across and through these varied practices. While typical descriptive analyses of such processes move forward, we have found that moving bidirectionally—from final product backward and from initial idea forward—better helps us to understand participants’ learning. We use examples from two creative production contexts to describe a technological tool developed for *bidirectional artifact analysis* by outlining its features, affordances, and constraints.

Why Analyze Creative Process?

In recent years, new literacies research has begun to re-focus the meaning of “literacy” and “multi-literacies” (e.g. Cope & Kalantzis, 2009; Moje, Overby, Tysvaer, & Morris, 2008). Beyond the comprehension and production of texts, “being literate” in a particular content area has come to describe fluency in specific “constellations” (Steinkuehler, 2007) of practices, modes, and ways of knowing. By producing content, learners use medium-appropriate tools to create artifacts and develop metarepresentational competence, an understanding of what tools are useful for expressing which ideas (Halverson, 2013). These understandings of literacy have particular purchase in digital environments and creative arts education, where interactions are often shaped by the production and critique of artifacts for a responsive audience.

Reframing literacy and learning in terms of how students use tools to communicate ideas requires an emphasis on measuring the process *and* the products of learning (Halverson, 2012). As researchers in the learning sciences, we seek to understand and design learning environments that support these complex practices. In order to do so, we must construct methods that help researchers describe the processes of creative production and the creation of successful literacy artifacts. We need analytical lenses that take the complexities of creative works themselves—from online video to school essay to multimedia collage—as units of analysis and trace these works through creation, drafting, and revision. Thus, we need tools that enable such analyses.

Because these social, artifactual, and often multimodal processes extend over varying periods of time, it is difficult to capture them using existing methods and tools for qualitative analysis. While our prior work has focused on the development of tools for analyzing youth-made films (Halverson, Bass, & Woods, 2012), these tools do not generalize to creative processes more broadly, nor do they capture the *process* of creation. Rather, they afford final product analysis. Altogether, creative production combines such activities as engagement with existing genres, problems, or frameworks; participation in interlaced processes (drafting, journaling, modeling, etc); sustained work on a particular idea; and audience-based critique (Barrett, 2000; Soep, 2006). In our work with young artists, media, and settings, we have found that their progress depends on their participation in all of these tasks—as well as ongoing engagement with their audience(s) (Magnifico, 2010, 2012; Halverson, 2012). In other words, the analysis of creative process depends on not focusing on particular types or genres of work, as in discourse analysis or artifact analysis, but looking across and through varied practices.

We first developed methods for understanding creative production processes (Halverson & Magnifico, 2013). We now have developed a tool to look across and through full trajectories of art development. Using Javascript, HTML 5, SVG, and CSS technologies, we present an alpha version and examine its affordances for data analysis and sense-making. We need such a tool to document creative practices, examine the origins of particular ideas and critiques, and trace how artists take up these concepts (in future drafts or more broadly). This year’s ICLS theme, “Learning and Becoming in Practice” highlights learning trajectories over time: ways in which learners adopt certain kinds of practices and become certain kinds of people (e.g. Duncan & Hmelo-Silver, 2009). Our tool helps streamline the analysis, visualization, and representation of the learning involved in creative production. As a field, the learning sciences value theory-building and technological design around such questions, and we are eager to participate in such ongoing conversations and inquiry.

Here, we use examples from two creative production contexts—creative writing and mobile game design—to describe our tool for *bidirectional artifact analysis*. We outline its features and offer suggestions for using the tool in other research contexts that value the role of process in understanding learning.

Audience as Fundamental to Learning Through Production

Much of the original literature that developed and documented rhetorical (e.g. Ong, 1975) and cognitive (e.g. Bereiter & Scardamalia, 1987; Flower & Hayes, 1981) viewpoints on audience implicitly assumes that when writers write with an audience in mind, these imagined readers have a measurable effect on the writing. This assumption is rooted in ethnographic and expert-novice research on how writers think and how writers write, and has been documented repeatedly with different kinds of writers. Understandings of audience and collaboration have broadened, however, as new media have pushed writers and artists into spaces where they may create, read, and view “among the audience” (Lunsford & Ede, 2009). This shift can transform writers into speakers, readers into critics, and both into collaborators. Within online and face-to-face artistic communities, it is increasingly plausible to think of readers and viewers in terms of Ong’s (1975) “collectivity, acting... on one another and on the speaker [or writer, or artist] as members of an audience do” (p. 11).

Considering how feedback guides artistic process suggests that audiences for creative work may play both a cognitive role in artistic production, in that sharing work forces creators to consider readers, and a social one, in that creators define themselves as community members by the work that they share (Magnifico, 2010). If we extend the metaphor of literacy from producing text to creating artifacts using various multimodal tools for meaning making, we can apply the same insight about audience to creative processes in general. Audience is central to the learning of creative production, not an afterthought. As a result, analyses of creative production processes must include an understanding of audience feedback and interaction as a core part of tracing learning.

The Role of Representation in Creative Processes

diSessa (2004) argues that representation design is a creative endeavor, “a venue in which creative and artistic skills are at a higher than normal premium” (p. 300). Thus, observing creative production helps us understand the role of external representation in learning (e.g. Eisner, 2002). “Getting smart” in the context of producing art “means coming to know the potential of the materials in relation to the aims of a project or problem” (Eisner, 2002, p. 72). The capacity to build representations from different materials for multiple situations is a marker of creative expertise (Hayes, 1989), and has been called “essential in investigations of creativity” (p. 262) (Hasirci & Demirkan, 2007). Since creative production requires creative representation, young artists must learn to construct and evaluate multiple representations. We have called this process “the representational trajectory” and its outcome the development of “metarepresentational competence” (Halverson, 2013).

The representational trajectory describes how producers focus on a central idea or concept, use this idea to motivate the development of skills and tools to represent their concept, and design a product that reflects this relationship between concept and tools. To study and describe this learning process requires multiple forms of evidence at multiple points in time, often with the engagement of multiple peers and mentors (Halverson, 2013). Creative production settings often encourage novices to use several media to produce their art. While representations may shift and ideas may transform through drafting, feedback, and revision, understanding how these artifacts contribute to the final “text” is necessary to understand the processes that producers employ.

Bidirectional Artifact Analysis

To address this methodological need, we developed *bidirectional artifact analysis* (BAA), an analytic method for understanding creative production that employs ethnographic observations of participants, artifacts they create, and interviews over time (Halverson & Magnifico, 2013). A key component of BAA is this concept of time. Most education research moves forward in time because learning continually builds up from prior knowledge. Our concept of time is grounded in Enyedy’s (2005) description of bidirectionality: “go[ing] ‘backwards’ in time in an attempt to trace the origins of this intervention and ‘forwards’ in time to examine what subsequent impact it had on the way other students reasoned” (p. 437). Moving bidirectionally helps us understand the role of social, collaborative feedback and audiences in young peoples’ learning. By turning the analytical lens, we see *how* learners’ reflections on past representations become shared prior knowledge.

BAA involves three steps: 1) identifying learner-created artifacts; 2) documenting relevant data around artifacts; 3) constructing threads across the data types to trace the final product’s core ideas and tools back through their development. We have used this method to understand various forms of production in different contexts including school writing (Magnifico, 2012), online creative writing (Magnifico, 2013), extracurricular digital art-making (Halverson, 2012), and radio documentaries in a college class (Bass & Halverson, 2012).

While technology has played a large role in documenting and capturing relevant data, we have found that we need a technological solution to this complex methodological problem: How can we construct threads across data types that allow us to “see” backwards and forwards across time, as BAA requires? Since the threads may include artifact analysis, multimodal analysis, and discourse analysis, we needed a way to identify, track, and compare narrative and artifactual threads both within a process (or draft) and across processes (or drafts). To address these needs, we have built a tool that aims to solve this problem. In the following sections, we describe the key features of the tool, highlighting potential affordances, constraints, and uses.

The Development of a Bidirectional Artifact Analysis Tool

Technical Details

The tool that we are developing, currently code-named Cotswold, is a javascript program that runs in a web browser. It uses the D3.js data visualization library to render its internal model of artifacts, their attributes and their relationships. D3.js uses a functional (as opposed to imperative) approach in which data structures are bound to a document object model and the appropriate rendering of the model is described in general without requiring code that iterates over individual data elements. Cotswold uses D3.js to produce a series of HTML and SVG documents that, when layered on top of one another in the top level HTML document, render the artifacts with all of their relationships and attributes. Every time the user makes a change, Cotswold updates its internal model, serializes it to a JSON string and stores the JSON on the user's computer using HTML5's local storage feature. In this way, the program remains entirely client-side. It does not store the user's data on a server in the cloud and in fact does not require an internet connection to run—only a web browser.

Cotswold is being developed in Google Chrome. It should work in any standards-compliant web browser but, as it is a working prototype, it has not yet been thoroughly tested. Currently, it lacks many features required for ongoing usability, but is an effective tool for idea generation. We are developing a variety of features and we anticipate that a beta will be available for sharing and testing at ICLS in June, 2014.

Features and Case Examples

Having established the importance of thinking bidirectionally, we turn to the decisions that we made while designing Cotswold, as well as the key features that enable data analyses and theory development. In many ways, Cotswold replicates the tools that we have used before. Like spreadsheet or transcript analysis programs, it allows open-ended coding and annotation. Like paper and colored pencil, it enables flexible multimedia timeline building and connection drawing. At the same time, the tool offers the ability to build persistent, manipulable visualizations of the ways in which artistic concepts evolve through formal critique, social conversation, mentorship, reflection, and ongoing revision. While paper and spreadsheet constructions may require several levels of abstraction and modeling to capture, code, and understand artistic drafting and critique processes—particularly when multiple media are involved—this tool uses the natural logic of such situations and processes to map and annotate them visually.

In short, the features capture how producers look backwards and forwards across their work in order to critique and revise. As we continue to develop Cotswold and use it to analyze datasets from multiple creative spaces, we anticipate examining larger questions, as well. For instance: Do commonalities in process exist *across* these environments? Do certain kinds of questions and interactions inspire creators to revise and rethink? How might these patterns be useful in designing learning situations that better support creative development?

In this section, we describe three central features of the tool and discuss their affordances for process-centered, holistic data visualization and analysis. We then explore potential constraints of this approach and consider how ongoing tool development will contribute to our analytical techniques and theory development.

Example Data

In order to define and explain Cotswold's features, we have drawn small samples of data from two previous studies that illustrate the features, affordances, and constraints of this tool. All names are pseudonyms.

Melanie's Story

The creative writing data are drawn from Magnifico's 2010 study of an extracurricular writing camp where high school writers in an urban area of the Midwestern United States worked on creative writing, workshopped their writing with peers and local writers ("writing coaches"), and performed their works in a final "showcase reading." Data collected includes observational field notes, interviews with young writers and writing coaches, recordings of "writers' circle" critique sessions, participants' writing notebooks and their stories. Each day, the young writers wrote during "sacred writing time" and participated in writers' circles. Melanie's story, excerpted here, was typical in length (900-word final draft) and genre (realistic fiction). She wrote 4 drafts during camp, as well as a 750-word anthologized version that she read at the showcase. In writers' circle, Melanie explained that her story was inspired by her mother's reading of *The Nine* and their resulting dinner-table discussion of Justice O'Connor's retirement from the Supreme Court to care for her Alzheimer's-afflicted husband.

ARIS Design Memos

Data from the ARIS Game Design project are drawn from Halverson and Kalaitzidis's 2013 study of an undergraduate course entitled "Digital Media and Literacy." This course engaged students in digital media production processes including the design of an augmented reality mobile game using the ARIS platform (see <http://arisgames.org/> for details). Students created games over the course of 6 weeks; during this process, they wrote design memos at the end of every week regardless of the progress they had made. Kristina described her

game as a guided tour of her local city, “focused on calmness and the artistic inspiration that can come from an invested sense of place” (final design memo). In order to understand how undergraduate students developed metarepresentational competence, we collected and analyzed these weekly design memos as well as all the artifacts they created during the course of the ARIS project: paper prototypes and design cards, screenshots of back end coding, and screenshots of versions of their mobile games.

Feature 1 — New Analytical Metaphors: Timepoints and Trajectories

Feature Description

Cotswold’s organization creates a timeline metaphor for data analysis. Often, analytical frameworks rely on coding, sorting, and categorizing utterances or activities. In a classroom study, researchers might separately group and analyze all of the instances where teachers teach full-class lessons and instances where students work in small groups. Creative process work, however, must examine the multiple activities that comprise the production cycle (Halverson, 2012) of a particular piece—steps that go into the story of that piece’s creation.

In other words, for us to understand the *process* of creation and all of its expected (and unexpected) developments in settings like creative writing workshops (Magnifico, 2012) and youth media organizations (Halverson, Bass, & Woods, 2012), we must examine full narratives of creative production. We initially began this work by adapting tools such as spreadsheets, paper, and colored pencils to view these large progressions, but such models are unwieldy and not easily revise-able. Thus, we designed Cotswold to digitally group artistic activities by trajectory and process.

Affordances

Changing analytical metaphors allows researchers to organize collections of pedagogical activity, artifact production, and critique by *time* rather than by type or code. Building timeline-based collections allows us to organize how we see and analyze various points on the timeline, and to examine multiple transcripts of participants’ talk and drafts of media artifacts in the same space. For example, Melanie began her story with a short vignette. In writers’ circle, she noted that “I only got a paragraph done [in sacred writing time]... it’s not a very good paragraph.” She explained that she didn’t know where to start and read the paragraph that introduced her main character’s central dilemma: A husband who has been diagnosed with dementia and is beginning to develop severe symptoms. This draft (“Draft 1”) is the leftmost timepoint, marked in Figure 1 in light blue.

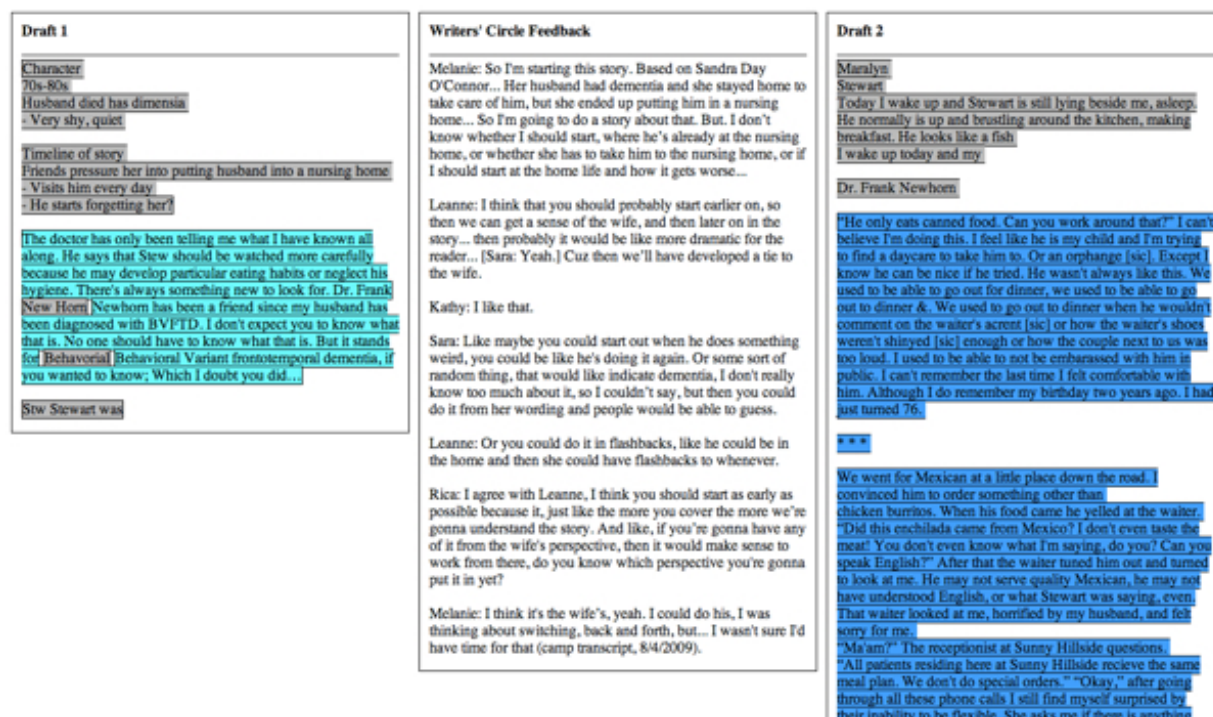


Figure 1: Timelines and timepoints place drafts in sequence and in context.

In the critique that followed (the middle timepoint, “Writers’ Circle Feedback”), the writers’ circle began to discuss what readers might need to know to understand the main character’s difficult domestic situation. Leanne initially suggested beginning the story earlier “so that we [the readers] can get a sense of the wife.” Sara furthered this suggestion by adding potential details, thinking about “start[ing] when he does something weird”

and pointing back to the text, where Melanie has discussed the husband's risk of developing symptoms like "peculiar eating habits." She considered how to pull readers in by helping them "guess" the storyline. Continuing her earlier suggestion, Leanne proposed a structure wherein Melanie might include flashbacks to "start earlier on" while maintaining a sense of the present. In the third "Draft 2" timepoint in Figure 1, marked in darker blue, Melanie reworked the opening to portray her main character speaking with an assisted living nurse, trying to place her husband in a supportive home. Later, the character flashes back to her first experiences of her husband's deteriorating condition.

At first, the two drafts look so different that it might be difficult to recognize them as drafts of the same story without the critique. The tool, however, allows researchers to mark and characterize differences between drafts; to view these artifacts in the context of creation, critique, and revision; and to visually represent changes across timepoints. Such time-based analytical structures make it possible to see the relationships among the artifacts and media that are constructed during creative production—a key to understanding these processes.

Feature 2 — Conceptual Coding: Multimedia Selection and Annotation

Feature Description

Working within the timeline metaphor, bidirectional artifact analysis relies on the in-context categorization of data such as transcripts of work sessions, critiques, and drafts of artifacts. While interpreting these data is similar to the general processes of qualitative research, the presence of multimedia artifacts complicates analyses. In response, Cotswold allows researchers to place textual and visual artifacts side by side in timelines and to annotate these elements to create codes, categories, and interpretations without separating this text from the overall context. (In future releases, we hope to incorporate video- and sound-based artifacts as well.)

Such annotation and interpretation is important because ideas, questions, and critiques often cycle through several drafts of a creative piece or across several members of a production environment. For example, in previous analyses, we have shown that asking an artist to consider the effect of a particular choice on readers or viewers—critiques that make audience involvement clear—often lead to revisions (Magnifico, 2012). Similarly, a particular motif may be introduced to a draft artifact after comments from readers or mentors. Because Cotswold can highlight words, phrases, and image fragments, as in Figure 2, it is easy to visually group such elements for analysis and write short memos about how such elements should be categorized. As codes and constructs are revised, it is possible to reorganize and rework the visual representation of these categories.

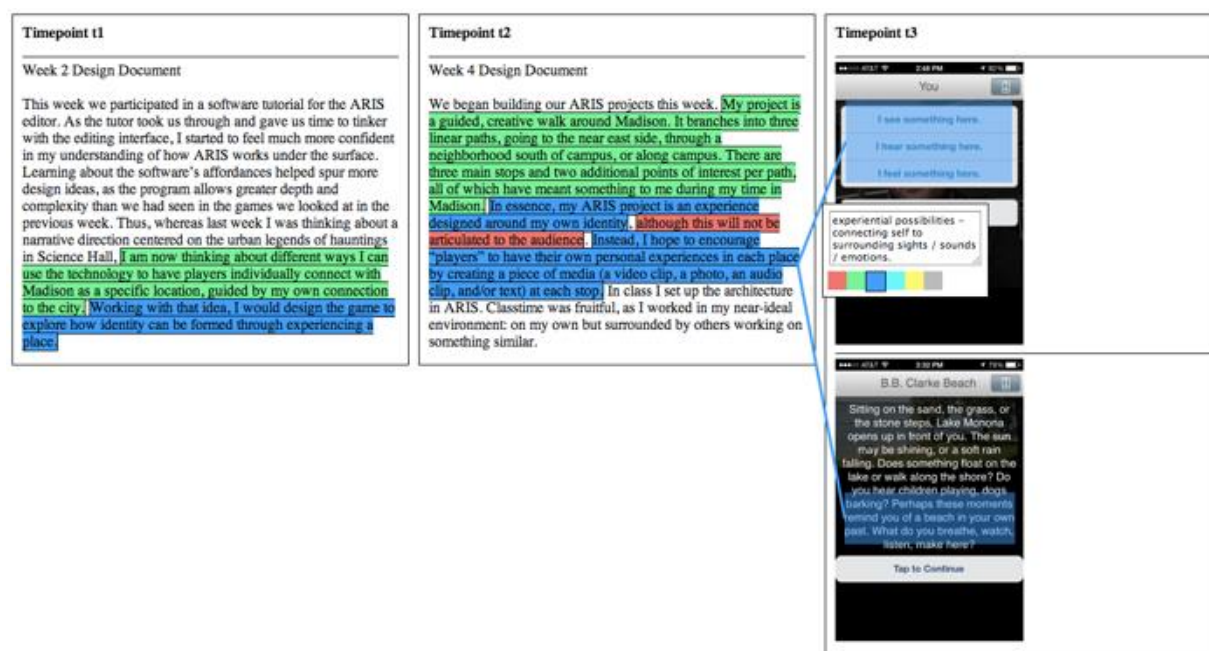


Figure 2: Highlighting and annotating talk and image side by side

Affordances

Regardless of whether such annotation captures micro or macro trends in the data, it occurs within the context of the timepoints and timeline so that researchers can perform artifact-based tracing without losing sight of broader narratives and trajectories of learning. Thus, the design of this feature affords a persistent visual, annotated instantiation of the revision and reflection that occur over a creative production cycle. Cotswold additionally allows this conceptual coding and annotation across multiple media. Researchers may place artifact

drafts next to transcripts of feedback and visually represent the threads of conversation connecting these timepoints. Such flexibility, shown in Figure 2, is useful for constructing timelines that capture creative production environments where young artists employ several media in their drafts. Such representations may reveal clear progressions of meaning even though the concepts look very different across media such as transcripts, sketches, and production prototypes.

Figure 2 shows Kristina's progress through the planning and development of an augmented-reality tour of her city, which figures prominently in the "development of her own identity." In the first two timepoints—design documents from Kristina's work on this tour—she considers her ideas about "how identity can be formed through experiencing a place" and describes wanting others to experience her city as she sees it. (Design Memo #2). In the textual Design Memos, she discusses the purpose of her game: A tool for players to capture (through photographs, video, or poetry) similar experiences and emotions. The two screenshots in timepoint three show how she translates these ideas to the screen in her ARIS game, which asks players if they "see, hear, or feel something here" and draws connections between her city's locations and similar locations "in your own past." Despite differences in narrative and image-based instances of such meanings, Cotswold allows researchers to link them through color (blue represents "identity" in Figure 2, and green represents "place") and drawn lines.

Feature 3 — Interpretation: Connecting Fragments

Feature Description

Cotswold's connection feature, shown in Figure 3, enables an interpretive phase of drawing connections and establishing threads of meaning across production activities. As previously noted, cycles in different creative environments might include such elements as sketching or drafting in multiple media, social conversation, reflection, critique, and performance. Once researchers have ordered the activities represented by the timeline and created coding categories to mark up the timepoints (using highlighting and annotation tools), they may interpret and trace artistic choices through artifacts and conversations. This feature allows researchers to connect fragments to each other (within and across timepoints) while remaining within the natural logic of creative drafting and production. It is also possible to note patterns, and annotate these connections for analysis and interpretation. Drawing interpretive connections and coding data in context enables researchers to perform detailed analyses while staying close to the process and the narrative of each creative piece.

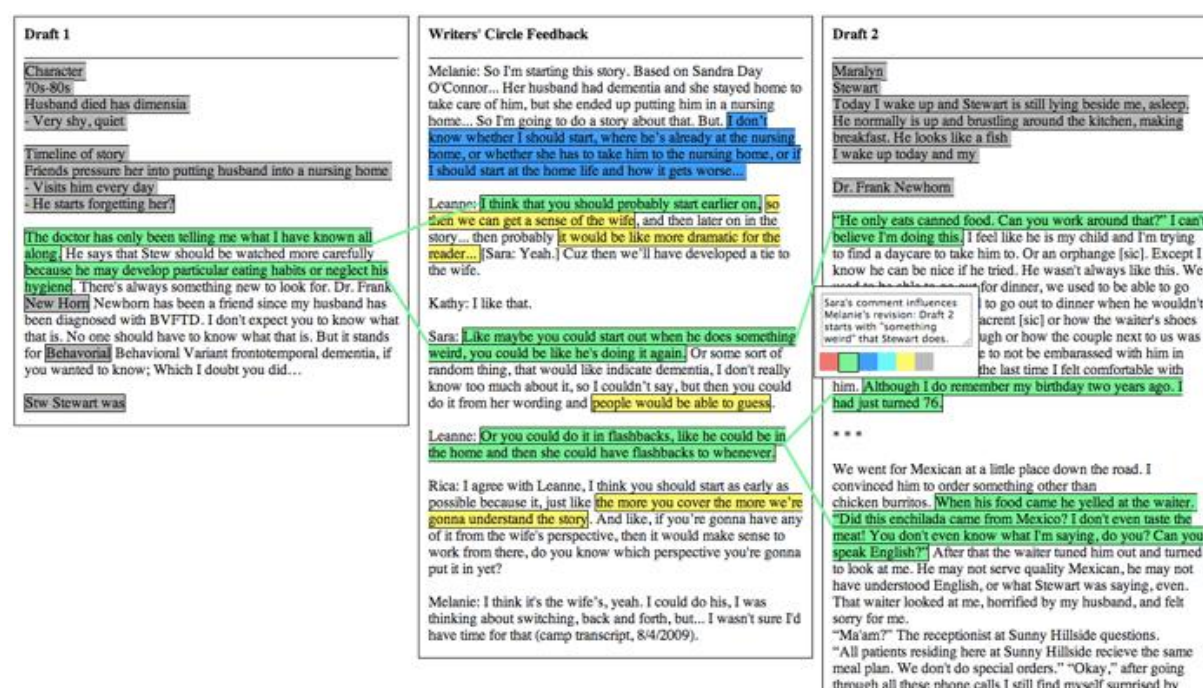


Figure 3: Tracing the development of story drafts through critique and connections

Affordances

In the course of creative production, critique often serves as ongoing, developmental conversation, so it makes sense for the tool to allow us to look backwards and forwards to annotate and connect key moments of insight without losing a sense of the overall context and activity. We return to the example from Melanie's story to show a different, more complex analysis of the relationships between her first draft, writers' circle critique, and second draft. As discussed above in Figure 1, Melanie revised her story almost completely between Draft 1

(timepoint 1) and Draft 2 (timepoint 3). In Figure 3, we examine how the feedback that Melanie received from her writers' circle centered on two concepts—structure and audience—that directed this significant revision.

The connection tool is useful here because it reveals how Melanie's revisions and additions are linked to the critique that she received in writers' circle. Without examining the critique transcript, however, it is difficult to understand why she revised as she did because her revision is so complete. As previously noted, Leanne, Sara, and Rica agreed that Melanie should revise her story's beginning scene to place her main character at the center of the story's action. Using the highlighting and connection tools, we traced the progression of two major concepts raised in this critique. In the middle timepoint of Figure 3, "Writers' Circle Feedback," Melanie first directs the critique (highlighted in blue) by requesting feedback on where the story's action should begin. Leanne suggests a structural revision to begin the action of the story "earlier on." This structural comment is highlighted green and connected to the Draft 1 text to show that "earlier on" means before the Alzheimer's diagnosis. Leanne looks back to Melanie's Draft 1 and discusses why a change is useful: If "we" "the readers" get to know the main character first, it will be "more dramatic" as her husband's behavior deteriorates and leads to this diagnosis.

Other members of the writers' circle reiterate this idea about structure and suggest possible revisions, such as adding flashbacks or clarifying the Alzheimer's diagnosis (all highlighted in green and annotated as structural suggestions in Figure 3). These critiques employ similar justifications for revision: Readers must make sense of the story and characters. These comments are yellow-highlighted and annotated as invocations of audience. In Draft 2, the rightmost timepoint, green highlights and connection lines interpret Melanie's adoption of these suggestions. Instead of beginning by "telling" readers the diagnosis, Draft 2 begins with a conversation between the main character and a nursing home aide and then propels the story back in time with a flashback.

Using Cotswold, researchers may highlight a range of related fragments and add connections between them. This interpretive analytical phase thus builds evidence for understanding revision motivations, as well as whether revisions are tied to particular critiques or pedagogical elements. Two fragments can only be connected to each other once, but any fragment can be connected to multiple other fragments. One piece of feedback may be the root of several revisions, as with Leanne's suggestion, and each of these can be linked to its root. As with selection and highlighting, researchers can use colored connections to code particular kinds of interpretations.

Constraints of this Approach

We have discussed the many affordances of using Cotswold for bidirectional artifact analysis—which represent, of course, reasons why we attempted the challenge of designing this tool. At the same time, this timeline-based and qualitative-coding-inspired view of creative process analysis constrains our sense of the data.

When we conceptualized BAA, we wanted to articulate "relationships between processes and products" and to examine "drafts as representations that grow and change as a result of individual cognitions, mentor responses, and social reflections" (Halverson & Magnifico, 2013, p. 406). By setting up timelines and marking artifacts and conversations as production timepoints, the tool forces researchers to examine data within its broad context and narrative arc. Writing about the patterns of a full production cycle, though, requires micro-examination of young artists' draft artifacts, critiques, reflections, and revisions. Cotswold is particularly useful in these fine-grained analyses, enabling us to see textual and imagistic changes and to trace them.

Constraints emerge because this analytical tool begins to dissect process in order to impose coded structures; it returns researchers to breaking creative production into pieces in order to document patterns. It is possible that looking for general patterns necessarily requires us to take such actions, but we cannot help but note that such decontextualization pushed us to develop a holistic view of creative process in the first place.

Conclusions and Implications

Understanding creative production processes is a complex endeavor that requires researchers to parse macro-level and micro-level data and, often, to combine data drawn from multiple media and creative activities. BAA offers a way forward by articulating a framework that combines methodologies for conducting inquiry with such qualitative datasets (Magnifico & Halverson, 2013), and Cotswold helps us track and document this inquiry without losing sight of the broader narratives and contexts in which learning occurs. While the tool currently enables the analysis of relatively small-scale textual and image-based datasets, we plan to expand these capabilities both in terms of scope and in terms of media range. As we continue to develop these methods and use the tool to analyze creative spaces, we anticipate examining larger questions, as well. We are particularly interested in learning whether creative processes seem to exist *across* these environments and how these patterns could potentially be useful in designing learning situations to support creative development.

We believe that BAA, and Cotswold as a tool for such analysis, has particular relevance to the learning sciences because of the field's interest in considering learners' individual content knowledge, their adoption of certain kinds of practices, and their becoming certain kinds of people as a result of their engagement in designed environments or problems. Many studies have established that using tools and making artifacts are key components of knowledge (Halverson, 2012; diSessa, 2004; Eisner, 2002). This framework and tool, in effect,

helps researchers to map and document the learning progressions in which young people participate in situ (cf. Duncan & Hmelo-Silver, 2009). We focus on artistic and creative production in our own work, but there are strong cross-content implications for this analytical approach. When we conceptualize learning as the development of representational competence that may be measured through individual cognitions, mentor responses, and social reflections (Halverson, 2013), bidirectional analysis becomes a possibility for understanding students' learning through experience and artifact, as well as contributing to ongoing curriculum alignment, assessments, and standards.

References

- Barrett, T. (2000). Studio critiques of student art: As they are, as they could be with mentoring. *Theory Into Practice*, 39, 29–35.
- Bass, M.B. & Halverson, E.R. (2012). Representation radio: Digital art-making as transformative practice in the college classroom. *Pedagogies: An International Journal*, 7(4), 347-363.
- Bereiter, C. & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.
- Cope, B., & Kalantzis, M. (2009). 'Multiliteracies': New literacies, new learning. *Pedagogies: An International Journal* 4, 164-195.
- diSessa, A. (2004). Metarepresentational competence: Native competence and targets for instruction. *Cognition & Instruction*, 22(3), 293-331.
- Duncan, R.G. & Hmelo-Silver, C. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, 46(6), 606-609.
- Eisner, E. (2002). *The arts and the creation of mind*. New Haven, CT: Yale University Press.
- Enyedy, N. (2005). Inventing mapping: Creating cultural forms to solve collective problems. *Cognition & Instruction*, 23(4), 427-466.
- Flower, L. & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition & Communication*, 32(4), 365-387.
- Halverson, E.R. (2013). Digital art making as a representational process. *Journal of the Learning Sciences*, 22(1), 121-162.
- Halverson, E.R. (2012). Participatory media spaces: A design perspective on learning with media and technology in the 21st century. In C. Steinkuehler & K. Squire (Eds). *Games, Learning, & Society: Learning and Meaning in the Digital Age*. New York: Cambridge University Press.
- Halverson, E.R., Bass, M. & Woods, D. (2012). The process of creation: A novel methodology for analyzing multimodal data. *The Qualitative Report*, 17, 1-27.
- Halverson, E.R. & Magnifico, A.M. (2013). Bidirectional artifact analysis: A method for analyzing digitally-mediated creative processes. In R. Luckin, S. Puntambekar, P. Goodyear, B. Grabowski, J. Underwood, & N. Winters (Eds.), *Handbook of Design in Educational Technology* (p. 406-416). New York: Routledge.
- Hasirci, D. & Demirkan, H. (2007). Understanding the effects of cognition in creative decision making: A creativity model for enhancing the design studio process. *Creativity Research Journal*, 19(2-3), 259-271.
- Hayes, J. R. (1989). Cognitive processes in creativity. In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *Handbook of creativity*. New York: Plenum.
- Lunsford, A.A. & Ede, L. (2009). Among the audience: On audience in an age of new literacies. In M.E. Weiser, B.M. Fehler, & A.M. Gonzales (Eds.) *Engaging audience: Writing in an age of new literacies*. (pp. 42-69).
- Magnifico, A.M. (2012). Audience effects: A bidirectional artifact analysis of adolescents' creative writing. 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. 411-415). Sydney, NSW, Australia: International Society of the Learning Sciences.
- Magnifico, A.M. (2010). Writing for whom: Cognition, motivation, and a writer's audience. *Educational Psychologist*, 45(3), 167-184.
- Moje, E. B., Overby, M., Tysvaer, N. & Morris, K. (2008). The complex world of adolescent literacy: Myths, motivations, and mysteries. *Harvard Educational Review* 78(1), 107-154.
- Ong, W. (1975). The writer's audience is always a fiction. *PMLA*, 90(1), 9-21.
- Soep, E. (2006). Critique: Assessment and the production of learning. *Teachers College Record*, 108(4), 748-777.
- Steinkuehler, C. (2007). Massively multiplayer online gaming as a constellation of literacy practices. *E-learning and Digital Media*, 4(3), 297-318.

Learning from Self-Explaining Emergent Phenomena

Kasia Muldner¹, Winslow Burleson¹, Michelene T. H. Chi²
Arizona State University

¹Computing, Informatics & Decision Systems Engineering

²Learning Sciences Institute, Mary Lou Fulton Teachers College
{katarzyna.muldner, winslow.burleson, mtchi}@asu.edu

Abstract: To date, relatively little work has explored how students learn about a particular class of processes, namely emergent ones. The research that has investigated these processes has primarily employed a case-study methodology. Here, we report on a controlled experiment comparing how students learn about the emergent topic of diffusion from self-explaining vs. from reading. In contrast to a prior study that found self-explanation was not associated with learning about emergence, students learned significantly more in the self-explaining condition. To shed light on how different types of self-explanations are related to learning, we analyze the content of students' explanations and their association to learning outcomes; we also present qualitative analysis of students' misconceptions and how these relate to existing theories of emergent attributes.

Introduction

Many science processes can be classified as being *emergent*, including diffusion (Chi, Roscoe, Slotta, Roy, & Chase, 2012), crystal growth (Blikstein & Wilensky, 2009), and natural selection (Dickes & Sengupta, 2013). Briefly, in an emergent process many micro-level agents behave according to simple rules to produce a more complex, macro-level outcome or pattern (Levy & Wilensky, 2008). In contrast to other “sequential” processes like the circulatory system, emergent processes correspond to a decentralized system, in that there is no controlling agent directing the behavior of the micro agents (Resnick, 1996). Moreover, in an emergent process, visible patterns at a given level (e.g., micro) do not necessarily correspond to patterns at another level (e.g., macro) – for a detailed comparison of emergent and sequential processes, see (Chi et al., 2012). To illustrate emergence using diffusion, suppose some blue dye is added to water. The micro-level agents correspond to the dye and water molecules, and these follow the simple rule of continuous, random motion. The macro-level outcome that arises from the molecular motion is a flow of the dye into water prior to equilibrium or a stable, unchanging light blue solution at equilibrium.

Given its prevalence in various phenomena, it is paramount that students understand emergent processes (Jacobson & Wilensky, 2006). Unfortunately, emergent concepts are very challenging (Asterhan & Schwarz, 2009; Chi et al., 2012; Meir, Perry, Stal, & Klopfer, 2005). For instance, some students, even at the university level, believe that during diffusion, molecules stop moving at equilibrium because the solution appears to be a uniform, unchanging color (Meir et al., 2005). This misconception suggests that students think the pattern at the micro-level must correspond to the pattern at the macro level, when in fact it can be disjoint for emergent processes. Reasoning connecting the micro and macro levels is also called *inter-level* reasoning (Blikstein & Wilensky, 2009; Chi et al., 2012), and some propose it is especially challenging for students in the context of emergence (Chi et al., 2012).

As we describe below, the majority of work exploring learning of emergence has focused on the utility of computer simulations depicting emergent phenomena, and a methodology corresponding to case studies (Blikstein & Wilensky, 2009; Levy & Wilensky, 2009b). While our work also involves simulations, our focus is instead on the role of student *self-explanations* about diffusion and their association with learning outcomes.

In general, self-explanation, i.e., the process of explaining and clarifying instructional material to oneself, is highly beneficial for learning (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Fonseca & Chi, 2011; Muldner & Conati, 2010; Nokes, Hausmann, VanLehn, & Gershman, 2011; Renkl, 1997). Since the seminal Chi et al. (1989) study, there has been a great deal of research on self-explanation, for instance to investigate individual differences in self-explanation behaviors (Renkl, 1997), instructional manipulations such as the effect of introducing gaps into examples (Atkinson, Renkl, & Merrill, 2003), and the relationship of various types of self-explanation prompts (Nokes et al., 2011). Moreover, the self-explanation effect has been shown to exist in various kinds of domains, including procedural ones like probability (Atkinson et al., 2003) and physics (Chi et al., 1989) as well as conceptual domains like biology (Gadgil, Nokes-Malach, & Chi, 2012).

To date, however, very little work has explored how students learn from self explaining about emergent processes, which as we mention above correspond to a specific class of phenomena that students may find especially challenging. To the best of our knowledge, there is only one study that investigated whether self-explanation helped learning of emergent topics. Specifically, Asterhan and Schwarz (2009) analyzed explanations produced by pairs of students collaboratively learning about natural selection. The results from a correlational analysis showed that self-explanations were not associated with learning in this domain. This finding is

somewhat surprising, since one would expect at least some benefit of self-explanation, and so we believe further investigation and replication are needed.

Thus, the goals of the present study are to investigate the impact of self-explanation for an emergent topic, as well as to explore (1) the relationship of different *types* of self-explanations and learning, and (2) the particular misconceptions students have about emergence as evidenced by their self-explanations.

Related Work

The majority of research exploring student understanding of emergent phenomena has involved qualitative case studies. Levy & Wilensky (2008) present analysis from interviews with 10 students that probe their understanding of emergent features of everyday phenomena, such as a crowd scattering. The focus of the analysis was on student reasoning related to the inter-level processes and how different types of reasoning (e.g., bottom up vs. top down) fostered understanding of the mid-level attributes.

Penner (2000) performed a series of case studies to investigate students' emergent reasoning. These case studies involved groups of students as they engaged in model building activities centered around representing and investigating emergent processes. The analysis includes excerpts of students' micro- and macro-level descriptions. Others have also explored the use of model building exercises, i.e., through simulations, for emergent processes. Typically, these exercises involve having students model the micro-agents using a set of rules (e.g., atoms move randomly), for instance in the NetLogo environment, which allows them to observe the resulting macro-level pattern emerge. Wilensky and Reisman (2006) describe two simulations of biological systems involving the impact of predator-prey on population dynamics and the coordinated behavior of fireflies and how they can be used to learn about emergence. Blikstein and Wilenski (2009) present case studies of students' experience with several simulations showing emergent phenomena related to materials science. The paper presents qualitative description of students' models and their high level reasoning related to these.

Other work on supporting learning about emergent processes with simulated environments has focused on the design of activities and guided prompts to guide students (Levy & Wilensky, 2009a, 2009b). These prompts targeted the macro, micro and inter-level aspects of emergence. The strongest effects in terms of learning outcomes were found for prompts that targeted the micro and inter-level concepts. While informative, this work did not gather data on students' self-explanations.

Chi et al. (2012) provided students with micro and macro simulations of emergent processes and prompts for emergent concepts during related activities. No correlation was found between learning and responses to prompts related to the macro simulation, but there was a high correlation between learning and responses to prompts related to the micro simulation – thus, the conjecture was made that students already knew the macro concepts. As was the case with the Levy and Wilenski work cited above, in this study students were not prompted to provide verbal explanations, instead being asked to type answers to questions in the learning materials at fixed intervals. Written explanations generated in response to text-embedded questions can be considered different from ones verbally produced, since the focus for self-explanations is to try and understand the instructional materials rather than to generate a coherent answer to a question (e.g., Chi 2000). Moreover, this study did not analyze how self-explanations in general, and specific explanations in particular, did (or did not) enhance learning of emergent concepts.

Study: Student Self-Explanations in Emergent Domains

The goals of the present study were to investigate the following three key aspects:

- (1) the utility of self-explanation for learning about an emergent domain, compared to reading
- (2) the nature of students' self-explanations and which types of explanations are beneficial for learning about emergence
- (3) student misconceptions as highlighted by their self-explanations

Materials

The study involved the following materials related to diffusion:

- two diffusion simulations
- a diffusion text
- a diffusion pre-test and post-test

To help students understand inter-level concepts, the two simulations showed diffusion occurring on the visible level (*macro* simulation) and at the molecular level (*micro* simulation); see the left and right panel of Figure 1 for the macro and micro simulation, respectively. The simulations were interactive (for instance, clicking the “start” button in the micro simulation resulted in molecules bouncing and colliding).

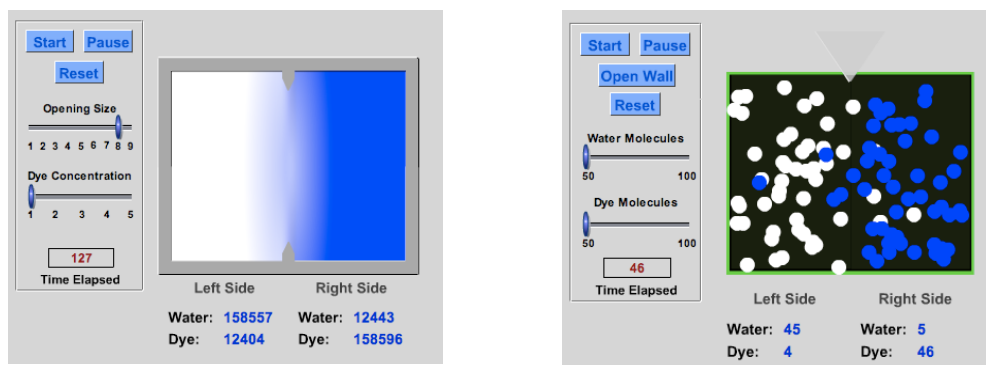


Figure 1: Macro level (left) and micro-level (right) simulations of diffusion

The diffusion text was printed on 8 by 11 sheets and was based on text used in earlier studies (Chi et al., 2012). The text described the diffusion process and also included information on emergent features and attributes, but without directly referring to emergence. The first two pages of the text were intended as a warm up, and so discussed non-emergent aspects of diffusion related to the properties of gases, liquids and solids. We followed the procedure used in (Chi, de Leeuw, Chiu, & LaVancher, 1994) to design two versions of the text: one with prompts reminding students to self-explain (*prompted text*) and one without prompts (*unprompted text*); otherwise, the content of the text was identical between the two conditions. The prompted text included a total of 52 generic prompts to self-explain (corresponding to “EXPLAIN” prompts, see Tables 2 and 3, top, for examples; 9 of these prompts were in the warm-up text). These prompts were embedded throughout the text (typically after each sentence or several sentences; whitespace was inserted between a given prompt and the next batch of text to clearly highlight what needed to be explained). Both versions of the text (prompted, unprompted) concluded with a description of the diffusion simulations and suggestions to use the simulations. In the prompted text these suggestions corresponded to 11 specific self-explanation prompts (e.g., “*What are the molecules doing during equilibrium?*”), while in the unprompted text, the suggestion was an invitation to use the simulation in free exploration mode.

The pre- and post-tests assessed students’ diffusion knowledge and were based on tests used in prior studies (Chi et al., 2012; Muldner, Lam, & Chi, 2013). The tests included questions that probed understanding of emergent aspects of diffusion at the micro, macro, and inter levels, but without explicitly mentioning emergence. For instance, to assess the inter-level *disjoint* attribute, one question asked “*As the dye diffuses away from where it was originally dropped into the water, can some dye molecules bounce back towards this original place?*” The pre-test included 25 multiple-choice questions, while the post-test included the same 25 questions and six additional questions for a total of 29 questions (10 *inter-level*, 8 *micro*, 2 *macro*, and 9 *other*).

Participants

The participants were 42 college students, who participated in the study to fulfill a psychology credit.

Design

The study included one independent variable with two conditions: *reading* and *self-explaining*. We used a stratified random sampling procedure based on pre-test performance to equalize prior knowledge between the two conditions.

Procedure

Each session was conducted individually in a private room that included a table and a desktop computer. Students signed the consent form (5 min.), filled in the pre-test (15 min.), and then read the diffusion text (*reading* condition) or read and self-explained the diffusion text (*self-explaining* condition). Students in both conditions were told to read out loud at their normal pace. Students in the self-explanation condition were also told to

follow the “explain” prompts to explain what the information means to you. For instance, you can explain what new information does each line provide for you, how does it relate to what you’ve already read, does it give you a new insight into your understanding of how diffusion works, or does it raise a question in your mind. Tell us whatever is going through your mind - even if it seems unimportant.

Table 1: Pre-test and post-test mean student performance (N = 42)

	pre-test %	ANCOVA-adjusted post-test %
<i>self-explaining condition (n = 23)</i>	47.0%	76.4%
<i>reading condition (n = 19)</i>	47.4%	68.9%

To ensure students in the self-explaining condition followed instructions to explain, they were prompted if needed by an experimenter, who sat behind them to avoid interfering with the task. It turned out that as in prior work (Chi et al., 1994), students rarely needed reminders to self-explain because they spontaneously followed the prompts that appeared in the text. To control time on task, students in the reading condition were asked to read the text twice, following the procedure in (Chi et al., 1994).

For both conditions, the diffusion text was provided on paper, while the diffusion simulations that the latter part of the text referred to were presented on the desktop computer; students could interact with these simulations (e.g., pressing the “start” button began the simulation of the diffusion process on the macro or micro scale in the macro and micro simulations, respectively). All sessions were video recorded, and student utterances following the warm-up session in the self-explaining condition were transcribed.

Does Self-Explanation Foster Learning of Emergent Topics?

To determine if self-explanation fostered learning of emergence over reading, we used an ANCOVA with the *pre-test %* as the covariate and *post-test %* as the dependent variable. This analysis showed that overall, students performed significantly better on the post-test in the self-explaining condition as compared to the reading condition ($F(1, 39) = 4.98, p = .03, \eta^2 = .086$; see Table 1 for details). Thus, in contrast to Asterhan & Schwarz (2009), in our study self-explanation was beneficial for learning of emergent topics.

Since one of our primary goals was to gain insight into the relationship between the content of students’ self-explanations and learning of emergent topics, for the remainder of the paper we focus on the self-explaining group.

What Types of Self-Explanations are Associated with Learning?

To investigate whether *what* students explained was associated with learning, we labeled each student self-explanation in the transcribed protocols as follows (see Table 2 for examples):

- *macro-level* self explanations expressed ideas about the “big picture” visible level of diffusion that one could see with the naked eye
- *micro-level* self-explanations referred to molecules and/or their interactions
- *inter-level* self-explanations connected micro-level concepts to the visible macro level
- *other* self-explanations related to concentration and miscellaneous topics

Table 2: Excerpt of the diffusion text (top) and examples of corresponding self-explanations (left) and self-explanation type (right)

[Text]: Eventually, the ink molecules are spread evenly throughout the beaker, and the diffusion process appears to have come to an end. [EXPLAIN]	
<i>Um eventually the whole thing of water will be one color</i>	macro level
<i>It appears that it's .. it's stopped because it the entire beaker is like the same color there is no more mixing I guess</i>	macro level
<i>So this is when the dye molecules of eventually collided in every area possible in the container which causes the ink molecules to disperse</i>	micro level
<i>the molecules are always moving and they are always gonna be sliding past each other and colliding so I don't know if diffusion ever comes to an end per se</i>	micro level
<i>So it's all done equilibrium is reached and then they [the molecules] are still moving but you cant really tell cause it's all the same color</i>	inter level
<i>Pretty much spread out the whole thing and once you see the beaker a certain color because the molecules from the ink has evenly divided with the water</i>	inter level

Table 3: Excerpt of the diffusion text (top) and examples of corresponding of paraphrase (left) and paraphrase type (right)

The macro level pattern of diffusion is fairly clear: The initial dark concentration of ink in one area of the beaker gradually spreads out until the whole beaker becomes evenly – but not as darkly – colored.

[EXPLAIN]

<i>Um that when you drop the beaker in that's where the flow of dye starts to go</i>	Macro Paraphrase
<i>So from the direction where the dye was dropped or where they dropped the dye into the water that is the starting point from where it moves on and then it swirls around</i>	Macro Paraphrase

As far as the grain size of the coding, following an explain prompt in the text, students tended to express an explanation that encapsulated a single, high-level idea. Thus, for the vast majority of cases, the grain size for the coding was the entire explanation a student provided following such a prompt. If the content of a student's explanation involved a given level (e.g., micro), then it was coded as such even if the text included a macro-level idea (e.g., see *micro level* rows in Table 2). The rationale behind this choice was that unless students made explicit the link between levels, there was no way to objectively identify when they were reasoning in an inter-level fashion.

As a starting point, we relied on the definition in (Chi, 2000) and did not differentiate between self-explanations that were paraphrases vs. ones that included inferences beyond information in the text snippet students read prior to generating the explanation (the latter are referred to *self-explanation inferences*, *SEI* (Chi, 2000)). Instances not considered self-explanations corresponded to cases where students merely re-read the passage verbatim, or expressed a fragment that did not include sufficient information to determine a code.

If we consider all four categories of explanations (including the *other* category), there were a total of 1087 explanations, or 47.2 on average per student. Since there were a total of 54 prompts after the warm up (43 generic and 11 specific prompts), this analysis shows that on average, students responded to the prompts with some type of self-explanation in the majority of cases.

For the remainder of the paper, we focus on the macro/micro/inter-level explanations. Note that these three levels are not unique to emergent processes – in fact, Chi et al. (2012) describes how other types of processes also can be characterized by these levels. However, the characterization of emergent processes at each level as well as across levels is unique, as compared to, for instance, sequential processes (for a full description, see Chi et al. (2012)).

Students expressed a total of 993 self-explanations related to macro, micro and inter-level concepts, with each student producing on average 43.1 total such explanations. On average per student, inter-level explanations were least common ($M = 8.5$, $SD = 4.5$), followed by macro-level explanations ($M = 12.9$, $SD = 4.4$), and micro-level explanations ($M = 21.7$, $SD = 3.5$). Students generated significantly fewer inter-level explanations than macro-level explanations (paired test $t(22) = 2.6$, $p = 0.01$, $d = .6$) or micro-level explanations (paired test $t(22) = 12.9$, $p < 0.01$, $d = 2.7$); students also generated fewer macro-level explanations than micro (paired test $t(22) = 5.9$, $p < 0.01$, $d = 1.2$).

How were different types of self-explanation (macro, micro, inter-level) associated with learning outcomes? To address this question, we conducted a correlational analysis between each type of explanation and pre- to post-test gains (calculated using the difference between pre-test and post-test, i.e., *post-test % - pre-test %*). Macro-level explanations were negatively associated with pre- to post-test gains ($r = -.46$, $p = .029$), while both micro- and inter-level explanations were positively associated with pre- to post-test gains ($r = .43$, $p = .04$ and $r = .44$, $p = .04$, respectively).

As we reported above, the post-test did not include many purely macro level questions, and so one possibility is that we are simply not measuring learning related to these concepts. This conjecture doesn't explain, however, why macro-explanations may have interfered with learning as suggested by the negative correlation. We propose an alternative interpretation, namely that macro-level self-explanations are straightforward for students to generate (as was also found, for instance, by Levy and Wilensky (2009b)). This is because macro-level concepts relate to visible phenomena that students are exposed to in daily activities (e.g., dye swirling in a liquid after it is dropped in), and so likely know about already. This conjecture can be strengthened by investigating the number of correct vs. incorrect explanations generated. This coding confirmed that macro-level explanations are clear-cut compared to the other types of explanations, in that students expressed significantly fewer incorrect macro-level explanations, as compared to the number of incorrect micro-level explanations ($t(22) = 2.6$, $p = 0.02$, $d = .6$) or incorrect inter-level incorrect explanations ($t(22) = 2.8$, $p = 0.01$, $d = .6$). Thus, by generating explanations about macro-level concepts that they for the most part likely already knew, students

were missing opportunities to explain about aspects that they did not know related to micro and inter-level phenomena.

Self-Explanation Inferences (SEIs) versus Paraphrase Explanations

While in general there is evidence that self-explanations foster learning, this should be particularly true for explanations that include content that goes over and beyond the text the student read previously to generating the explanation, i.e., that $SEI > SE$ as far as learning of diffusion is concerned. To address this question in the context of an emergent domain, we labeled each macro/micro/inter-level explanation as follows:

- *SEI* if the explanation contained domain-relevant material over and beyond what was provided in the text associated with the explanation, and
- *paraphrase* otherwise

Table 2 provides examples of SEIs (content over and beyond what was provided in the text underlined), while Table 3 provides examples of paraphrase explanations. On average, students expressed the following number of each type of SEI explanations: macro ($M = 10.2$; $SD = 4.1$); micro ($M = 17.1$; $SD = 4.6$), inter-level ($M = 8.3$; $SD = 4.6$); for paraphrases explanations, students expressed the following number of each type: macro ($M = 2.7$; $SD = 2.7$), micro ($M = 4.6$; $SD = 2.8$), inter-level ($M = .2$; $SD = .4$). Overall per student, a large proportion of explanations were SEIs (on average 81.5%), highlighting that students were constructive in the self-explaining condition. To see how SEIs vs. paraphrases influenced learning outcomes, we re-ran the correlational analysis from the previous section for each type of explanation. As far as the SEIs are concerned, this analysis confirmed our above results related to micro and inter-level explanations: micro and inter-level SEIs were positively associated with pre- to post-test gains ($r = .46$, $p = .03$ and $r = .42$, $p = .049$, respectively). The macro SEIs were not significantly associated with learning ($p = .23$). As far as the paraphrase explanations are concerned, none were reliably associated with learning, although there was a trend for the macro-level paraphrases being negatively associated with pre- to post-test gains ($r = -.35$, $p = .105$).

Thus, as reported in the above section, while in general micro and inter-level self-explanations were associated with learning, only when students provide additional inferences in their explanation (i.e., an SEI) was this association reliable.

Do Low and High Gainers Differ in their Self-explanation Patterns?

Another way to examine the association between self-explanation and learning is to divide students into low and high learners based on pre- to post-test gains, and examine their explanation behaviors. To do so, we took the low and high tails of the gain distribution, ignoring the middle group of students. The latter was done since the gains were normally distributed and so doing a straight median split did not correspond to a true low/high gain population, i.e., the students around the middle were not truly low or high gainers.

When we analyzed self-explanations as a whole, without distinguishing SEIs from paraphrases, we found that the high-gainers ($n = 7$) had significantly more inter-level self-explanations than the low gainers ($n = 6$; 12.0 vs. 5.2, respectively, $t(11) = 3.1$, $p = 0.02$, $d = 1.9$) and marginally more micro-level explanations (22.3 vs. 18.6, respectively, $t(11) = 1.8$, $p = 0.098$, $d = 1.1$). In contrast, the high gainers had significantly fewer macro-level explanations (10.4 vs. 15.8, $t(11) = 2.5$, $p = .03$, $d = 1.5$). As far as the SEIs, this pattern for the most part held, in that the high gainers had significantly more inter-level SEIs (11.7 vs. 5.0, respectively, $t(11) = 2.8$, $p = 0.02$, $d = 1.7$) and more micro-level explanations (18.7 vs. 12.6, respectively, $t(11) = 2.2$, $p = 0.048$, $d = 1.3$, respectively); there was a trend that high gainers also had fewer macro SEIs but this did not reach significance. Thus, these results support the correlational analysis above that micro and inter-level explanations are associated with learning but macro-level explanations may interfere with it.

What Kinds of Misconceptions Are Present in Students' Self Explanations?

Prior work has indicated that emergence is a challenging domain. Thus, to gain insight into student misconceptions as expressed by their self-explanations, we checked students' macro/ micro/ inter-level explanations for the presence of misconceptions. Students generated a total of 96 incorrect ideas in their explanations; on average per student, 0.68 incorrect macro-level ideas, 1.95 incorrect micro-level ideas and 1.59 incorrect inter-level ideas. However, if we look at the proportion of misconceptions, on average per student, there was a similar proportion of micro and inter-level misconceptions (on average, 39.6% vs. 40.1%), while the proportion of macro misconceptions was lower (on average, 20.3%).

Some of misconceptions pertained to students' incorrect belief that molecular motion is not random (randomness is one of the characterizing features of the agents corresponding to an emergent process (Chi et al., 2012)). For instance, after reading that a water molecule goes to the right side, a student expressed that "I'm guessing a water molecule goes to the other side", i.e., that the molecules exchange places to maintain equilibrium. Along a similar vein, another student expressed that "as one is crossing over, the other ones start shooting

over too because there is not enough space for them in that same spot” - this was said despite the fact he was looking at the micro simulation, which clearly portrayed random motion of the molecules.

Another feature of emergent processes is that the agents at the micro level do not intend to cause the macro level pattern (Chi et al., 2012), and that in general, none of the agents embody “intention”. As expected, some students expressed explanations contradicting this fact, related to all three levels. For instance, one student explained that “at micro level again molecules are just - water molecules are just *trying* to grab the ink molecules and ink molecules are *trying* to grab the water molecules”. Other students explained that the “dye molecules just go to the other side *trying to find space*” or that a molecule “crosses to the other side because there is a lower concentration of dye molecules on that side”. As another example, a student also generated an inter-level explanation that “once equilibrium has been reached the visible changes at the macro level don’t seem to change but suggesting that um at the micro level the molecules are constantly in motion *which I’m guessing is to maintain that equilibrium*”.

Yet another feature of emergence is that the micro and macro patterns can be disjoint (Chi et al., 2012). However, some students explained that the molecules followed the macro pattern, e.g., “it shows how the dye molecules are slowly moving to the left side where the water was an the water molecules are slowly moving to the right side where the dye was”, and that at equilibrium, “they are still moving but they don’t really move that much any more cause the color is equally distributed”. Another student echoed this sentiment explaining that “after a while it makes equilibrium and they don’t move that much”.

Other students thought that only one type of agent was responsible for the macro pattern when in fact all agents are (Chi et al., 2012), e.g., “it seems that the dye molecules are responsible for the flow pattern because they are the ones diffusion to the left side which contains the water molecules”. We also saw evidence of so-called inter-level slippage (Blikstein & Wilensky, 2009), when students expressed explanations inappropriately attributing features of a given level to another. For instance, during diffusion, the molecules move randomly, but at the visible macro level, there is a predictable pattern where the dye spreads from high to low concentration. However, students would attribute the random attribute to the macro level, e.g., “it’s a liquid so it’s going in random directions”. Another student reasoned in the opposite direction, by saying that “in the micro level the elements moving from the high concentration to the low concentrated part”. Yet another explanation concluded that the molecules “are gonna diffuse as it says and it becomes one particle instead of two”, which may be an indication of attributing the uniform color of the water at equilibrium to the molecules somehow combining into one type of molecule. Along a similar vein, another student explained that “eventually ... all the molecules are the same color”.

Discussion and Future Work

In this study, we focused on students’ verbally-expressed self-explanations in an emergent domain. In contrast to prior work (Asterhan & Schwarz, 2009), we found that self-explaining was beneficial, more so than reading. One possible explanation for this difference in results between our and prior work may relate to the context, in that in our study, students worked alone, as opposed to the Asterhan and Schwarz study where students worked collaboratively (and so other constructs that were analyzed, like argumentation moves, may have overshadowed explanation effects). We do acknowledge that only the students in the self-explaining condition had specific prompts (related to the simulation), which may have had an especially beneficial effect. However, these types of prompts corresponded to about one quarter of the total prompts (i.e., the majority of the prompts were generic in the self-explaining condition), and so were unlikely to account for the majority of the explaining benefit.

Having determined that self-explanation is more beneficial than reading, we then focused on analyzing how the content of students’ explanations related to their learning outcomes. We found that only micro and inter-level explanations were associated with learning. Differences between low and high gaining students mirrored this pattern, in that the high gainers preferred to generate explanations related to micro and inter-level concepts, while low gainers preferred generating macro-level explanations. This highlights that the types of explanations students generated were driven by individual differences instead of merely by the text students were reading. In a sense, the low gainers may have preferred to invest less effort and so deferred to explanations that were easier to produce (as is suggested by the lower error rate for the macro explanations, as compared to the micro and inter-level ones). Other work in non-emergent domains has also shown that low gaining students prefer to engage in unproductive behaviors, such as guessing and checking (Baker, Corbett, Koedinger, & Wagner, 2004), since these require less cognitive effort than trying to understanding the underlying domain.

Our work suggests that a fruitful future avenue for supporting emergent learning from self-explanation corresponds to encouraging students to generate micro and inter-level explanations, for instance via prompts embedded in digital environments. An open question, however, is whether in order to fully understand emergence, students need to acquire the necessary emergent schemas, and so require schema training, as proposed in (Chi et al., 2012). Others propose that students’ exposure to emergent concepts through guided interventions may be sufficient (Jacobson & Wilensky, 2006). These avenues await future research.

References

- Asterhan, C. S. C., & Schwarz, B. B. (2009). The role of argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialogue. *Cognitive Science*, 33, 373-399.
- Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, 95(4), 774-783.
- Baker, R., Corbett, A., Koedinger, K., & Wagner, A. (2004). *Off-task behavior in the cognitive tutor classroom: when students "game the system"*. In Proc. of SIGCHI Conference on Human Factors in Computing Systems (CHI'04), 383-390.
- Blikstein, P., & Wilensky, U. (2009). An Atom is Known by the Company it Keeps: A Constructionist Learning Environment for Materials Science Using Agent-Based Modeling. *International Journal of Computers for Mathematical Learning*, 14(2), 81-119.
- Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology* (pp. 161-238). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chi, M. T. H., 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.
- Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18, 439-477.
- Chi, M. T. H., Roscoe, R., Slotta, J., Roy, M., & Chase, C. (2012). Misconceived Causal Explanations for Emergent Processes. *Cognitive science*, 36(1), 1-61.
- Dickes, A., & Sengupta, P. (2013). Learning Natural Selection in 4th Grade with Multi-Agent-Based Computational Models. *Research in Science Education*, 43(3), 921-953.
- Fonseca, B., & Chi, M. T. H. (2011). The self-explanation effect: A constructive learning activity. In R. Mayer & P. Alexander (Eds.), *The Handbook of Research on Learning and Instruction*: Routledge Press.
- Gadgil, S., Nokes-Malach, T. J., & Chi, M. T. H. (2012). Effectiveness of holistic mental model confrontation in driving conceptual change. *Learning and Instruction*, 22(1), 47-61.
- 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.
- Levy, S., & Wilensky, U. (2009a). Crossing Levels and Representations: The Connected Chemistry (CC1) Curriculum. *Journal of Science Education and Technology*, 18(3), 224-242.
- Levy, S., & Wilensky, U. (2009b). Students' Learning with the Connected Chemistry (CC1) Curriculum: Navigating the Complexities of the Particulate World. *Journal of Science Education and Technology*, 18(3), 243-254.
- Levy, S. T., & Wilensky, U. (2008). Inventing a "Mid-level" to make ends meet: Reasoning between the levels of complexity. *Cognition and Instruction*, 26(1), 1-47.
- Meir, E., Perry, J., Stal, D., & Klopfer, E. (2005). How effective are simulated molecular-level experiments for teaching diffusion and osmosis? *Cell Biology Education*, 4, 235-248.
- Muldner, K., & Conati, C. (2010). Scaffolding Meta-Cognitive Skills for Effective Analogical Problem Solving via Tailored Example Selection. *International Journal of Artificial Intelligence and Education*, 20(2), 99-136.
- Muldner, K., Lam, R., & Chi, M. T. H. (2013). Comparing Learning from Human Tutoring and from Observing. *Journal of Educational Psychology*, Advanced online publication. 18 pages.
- Nokes, T., Hausmann, R. M., VanLehn, K., & Gershman, S. (2011). Testing the instructional fit hypothesis: the case of self-explanation prompts. *Instructional Science*, 39(5), 645-666.
- Penner, D. E. (2000). Explaining systems: Investigating middle school students' understanding of emergent phenomena. *Journal of Research in Science Teaching*, 37(8), 784-806.
- Renkl, A. (1997). Learning from Worked-Out Examples: A Study on Individual Differences. *Cognitive science*, 21(1), 1-29.
- Resnick, M. (1996). Beyond the centralized mindset. *Journal of the Learning Sciences*, 5(1), 22.
- Wilensky, U., & Reisman, K. (2006). Thinking like a Wolf, a Sheep or a Firefly: Learning Biology through Constructing and Testing Computational Theories -- an Embodied Modeling Approach. *Cognition and Instruction*, 24(2), 171-209.

Acknowledgements

The authors thank the anonymous reviewers for their helpful suggestions. This research was supported by a grant from the Spencer Foundation (200800196).

The Programmers' Collective: Connecting Collaboration and Computation in a High School Scratch Mashup Coding Workshop

Deborah Fields, Utah State University, 2830 Old Main Hill, Logan, UT
Veena Vasudevan and Yasmin B. Kafai, University of Pennsylvania, 3700 Walnut Street, Philadelphia, PA
deborah.fields@usu.edu, veenav@gse.upenn.edu, kafai@upenn.edu

Abstract: One challenge in bringing collaborative programming into schools has been to develop software design activities that make computation an integral part of the collective design rather than just a work arrangement between team members. In this paper we propose a collaborative approach to programming, the programmers' collective, that builds on collaborative models found in do-it-yourself (DIY) and open source communities. We describe and analyze the work of a class of high school student collectives engaged in programming music video mashups as part of a collaborative challenge in the online Scratch community. Our multi-level analysis focused on students' learning of specific programming concepts, effects of collaborative and task design on learning, and their personal reflections. In the discussion we address how students' experiences point to the value of "nested collectives," or multiple levels of designed-for collaboration as well as a needed shift from a focus on computation to computational participation.

Introduction

In computer science education, engaging and supporting students in collaborative programming has been a central focus of research for multiple reasons: the educational benefits that collaboration has shown to offer in learning interactions, the collaborative nature of software design and related industry, and resource allocation that allows students to share a limited number of computers in classrooms. Yet many early studies have not found collaborative programming efforts in K-12 to be successful in helping all members of teams develop programming skills (e.g., Webb, Ender, & Lewis, 1986). More recently pair programming approaches, in which two students work together on programming tasks, have been favored. Research has shown that pair programming can be a very effective collaborative arrangement for both K-12 and college students in helping them learn key computational concepts and persisting in programming (Denner & Werner, 2007; Porter, Guzdial, McDowell, & Simon, 2013). Yet while pair programming may meet some of the learning goals and resource needs of schools, it fails to expose students to the collaboration so commonly reflected in professional software design and self-organized Do-It-Yourself (DIY) online creative collaborations.

Most pair programming efforts that take place in schools often assign students to teams and provide them with specific tasks to be accomplished. This is substantially different from collaborative programming in online communities which focus more on self-selected tasks and work arrangements. Unlike classrooms, in this latter collective style of programming, members might not have met in person and are often geographically distributed, working across different time zones. Most prominent examples of open-source community design, such as the design of the operating system Linux (Benkler, 2006) or computer animations in Newgrounds (Luther, Caine, Ziegler & Bruckman, 2010) are facilitated by adult members. These types of collaborative activities have also found their way into youth programming communities such as Scratch where large numbers of youth can often be found collaborating together on programming role playing stories, animations, and games (Resnick et al, 2009). Integrating this type of collaborative programming into school poses a challenge (Kafai et al., 2012) because the DIY mentality of free forming groups and flexible work arrangements does not fit well within the more structured organization of schools.

In this paper, we report on the implementation and analysis of a collaborative, computational design challenge where students worked in collectives to create music video mashups in Scratch. This pilot study focused on structuring collaboration through a programming task so we could make this type of DIY collective programming available in a school setting. Conscious of tendencies for groups to divvy up tasks and reify stereotypes by assigning harder tasks to people who are perceived to be better at computers (Dumont & Fields, 2013; Kafai et al., 2012), we structured the programming task such that each individual worked on a section of their group's song. Then each group integrated the sections together into one longer music video, situating the learning of key computational concepts of initialization, synchronization, and event-driven programming. Our research questions asked: In which ways does a collective programming task support the learning of key programming concepts, especially of initialization and synchronization? What do students have to say about their collective programming experience? In the discussion we address what we learned about students' collaborative programming experiences, the design of collective programming tasks, and the integration of this type of collective programming activity in school contexts.

Background

There is no question that learning to program is challenging. As research over the last 30 years has documented in hundreds of studies in schools, students encounter challenges with both syntax and control and data structures (e.g., Robbins, Rountree, & Rountree, 2003; Soloway & Spohrer, 1990). Much of the recent comeback of programming has taken place outside of schools in online communities in which youth engage in the production of games, animations, role-playing games, and worlds (Kafai & Burke, in press). In these interest-driven communities youth come together not only to hang out and mess around (Ito et al., 2010) but also to create, remix, and share their code. While gaining access to and participation in programming communities is not that easy (Fields, Kafai & Burke, 2009), the work arrangements in online communities are even harder to replicate in school settings that have more structured time and less flexibility, leaving students ill-prepared for managing these more open-ended collectives found online.

In designing the programmers' collective, we built on online, interest-driven collaboration observed "in the wild," but intentionally developed supports for collaborations. Using an iterative design research approach, we designed and implemented collaborative programming tasks for all members in the Scratch online community to join. In successive iterative implementations we moved from a Collab Challenge in which online groups were asked to program three objects (Kafai, et al., 2012), to a Collab Camp with a more story-oriented format that asked groups to write and program an interactive story (Roque, Kafai & Fields, 2012). Our analyses of how online groups came together (or not) to accomplish these tasks led us to formulate a need for students to develop "collaborative agency"—a concept that draws on Scardamalia's (2002) work on collective cognitive responsibility—meaning that learners have to assume agency in finding others, organizing work, and contributing to and critiquing designs (Kafai et al., 2012). By all accounts these are desirable skills for students to develop in collaborative contexts and that can be supported via design features such as providing spaces to find collaborators and share curated work in galleries, furnishing mechanisms for attribution (Monroy-Hernandez, 2012), and modeling constructive criticism. These findings pointed to the need to design for what we call *nested collectives*—multiple levels of collaboration that attend to individual learning in the context of different layers of collaborative support.

In this third iteration of collaborative projects, Collab Camp 2, we focused on supporting individual learning through participation in a collective programming task. The design of the task would ensure that students encounter two core concepts that most novice programmers find challenging: initialization and synchronization. Initialization is an important concept in computer programming because it suggests a starting state or point for all of the actions and actors in a program, important enough that many modern languages provide means for automatically enforcing or checking initial states of variables. In Scratch, initialization becomes an intentional design because initial states of where something should start need to be clearly defined. When students run their programs repeatedly, they begin to notice that the program is not behaving as they desire. For example, a programmer may command a sprite (a character or image that is directed by commands) to move across the screen. On subsequent runs of the program, the sprite will begin in the last place it ended during the prior run, demonstrating the importance of initializing a sprite's x and y coordinate location at the start of the program. Thus, programmers need to learn to set a sprite's appearance, state (e.g., showing or hidden), size, and direction. A second goal was to support learning of synchronization and event-driven programming. Synchronization and event driven programming involve coordinating processes and events within a program such that they can be executed in parallel or sequentially without bringing the program to a stand still. In Scratch, coordinating multiple sprites' actions involves learning different forms of synchronization and event driven programming (Maloney et al., 2008; Fadjo, 2012).

For this study, we situated our programming task in an authentic form of collective youth media production: making (programming) video mashups of popular music. In order for a music video to flow smoothly, all sprites' locations, appearances, and sizes from each segment needed to be initialized and each event (i.e., each segment) needed to be triggered through synchronization and event commands. While this design of the programming task focused the coordination of individual contributions by making them essential to the completion of the project, we also recognize the role that larger communities, offline and online, played in supporting and providing feedback. Each programmers' collective became part of increasingly larger collectives—smaller groups within a larger workshop group within a broad, online challenge that all worked on programming music mashups. Our analysis focuses primarily on the design task but also attends to the supports for learning and collaboration at the larger collective levels as well.

Participants, Context, and Methods

This research was conducted at a high school situated in a metropolitan city in the northeastern United States. This high school is known for their intensive focus on science, technology, engineering and math (STEM). Freshman students (ages 14-15) participated in this workshop as part of their enrichment curricula in partnership with a local museum. Twenty-nine high school students participated in the workshop (12 girls, 17 boys) and of those, 23 (9 girls, 14 boys) consented to participating in the research. This included five male "mentors" in their

sophomore year that had participated in Scratch workshops the year prior. A team of researchers from the University of Pennsylvania taught the workshop: one main instructor led the course and was supported by two additional graduate student researchers. Five undergraduate interns also periodically supported students.

The students in the workshop participated in an online design challenge hosted on Scratch.mit.edu: completing a video music mashup as part of a collaborative team. The online challenge, Collab Camp 2 (see Roque, Kafai & Fields, 2012), had minimal requirements: participating groups had to submit a draft and finished version of their mashups and groups had to be comprised of two or more people. In our workshop we added a few additional constraints to support students' collaboration and learning: 1) students worked in collectives of 3-6 (versus the minimal requirement of two members), 2) each collective made a music video (versus the more general "music mashup" online), 3) each member had to design a section of the song their group chose (versus no structure on how to collaborate). The workshop comprised of eight sessions that ran for two hours each week in Winter 2012. In the early sessions, students selected songs they wanted to mash up, and were sorted into groups based on song choice. Then, each student selected a section of the song to animate or illustrate in Scratch. In addition, during the early weeks, the collectives were given a short programming tutorial and a few lines of starter code that used a particular mechanism for synchronizing events with time in the music video. During the sixth session groups merged their segments together to submit one seamless music video draft to the online challenge. Alongside other online-based groups, the collectives received external feedback from experts (youth and adult) on the Scratch website by the seventh session, and proceeded to revise and finesse their projects for the final submission to the website and presentation in the workshop.

Data from the project include weekly interim and final Scratch programs, observational data including field notes and video logs, and post-hoc interviews with a subset (11 of 25) of the students including three student mentors. To analyze student learning of specific programming concepts we turned to their programs, analyzing their music mashups for evidence of initialization and synchronization. We developed a coding scheme to determine the frequency and range of the initialization and synchronization strategies as well as how consistently teams employed these across individual sections. This analysis helped us to see how teams collaborated to ensure flow in their music videos. To further enrich this analysis, we used the interviews to understand students' perceptions towards collaboration using a two-step open-coding scheme (Charmaz, 2000). We describe the key emergent themes—collaboration, engagement/motivation, and computation and learning—in more detail in the last sections of findings.

Findings

Our findings illustrate how the collaborative programming task was successful in getting all individual students and the collectives overall to employ a range of initialization and synchronization strategies to produce cohesive music videos. We also show how the nested nature of the collectives influenced students' designs and opportunities for learning and feedback. First, we will describe the key initialization and synchronization strategies used. Then, we present a case study to illustrate one collective's collaborative process of how the workshop design, team structure, design task and informal and formal feedback helped shape their final music video, which was emblematic of other students' experiences. Finally, we consider students' reflections on nested collectives, including their learning from high school mentors, college interns, and members of the broader online challenge.

Initializing and Synchronizing Music Segments in Programmers' Collectives

To manage the programming task, each student had to think about how to most meaningfully bring their sections of music or lyrics to life. Given that their code would eventually be merged with that of members of their collective, they also had to be cognizant of initializing appearances, actions, and variables as well as synchronizing characters' actions and movements so the music video flowed. However, being new to Scratch, this required students to navigate a significant series of challenges from being creative about how to bring their lyrics to life to knowing how to develop the code to make their intended actions happen. We identified two programming concepts as targets for learning: initialization and synchronization & events. In our analysis of the final programs, we found that all individual students eventually employed initialization systematically throughout their projects. Programming an initialization involved setting the beginning state of sprites' (Scratch characters') appearance (e.g. *hide*, *show*), location (e.g. *go to x=__, y=__*), costume (e.g. *switch to costume__*), direction (e.g. *point in direction*), graphic effects (e.g. *set size to __*), and layering (e.g. *go to front*). If variables were used in a project, students consistently initialized the variable values as well. In early versions of the projects, initialization strategies were not used consistently, yet by the time collectives submitted their first draft of projects most sprites were initialized in the ways detailed above. Observations showed that as students combined their segments together any gaps in initialization were revealed since sprites from a later segment would appear at the beginning of the song or in the wrong place on the screen (for more detail, see the case study of the Jane Doe collective). This suggests that the design of the collaborative programming task of

bringing individual segments together into a chronological, unified project was successful in supporting students' systematic learning of this computational concept.

In terms of synchronization, individual students applied a variety of synchronization strategies including those suggested in the starter code given to each group. The most basic strategies included using *when green flag clicked* and *wait* to synchronize events. Students used *when green flag clicked* to start the main sequence of the music video. The more sophisticated scripts developed by students made use of the scripts *wait until* and *broadcast* to initiate as well as to conclude a sequence. As prompted by the starter code given to each group, students most often used *wait until timer > X*, to start their sequences. All members of a group had to coordinate their start and end times with the global music video timer, thus, this command became essential in setting boundaries for individuals' sections within the music video. Demonstrating that they understood the utility of this command, many students also used it to end their sequences, for instance continuing to animate a sprite(s) until the timer reached a certain time and thus having the active sprites disappear from the stage to make way for the next sequence. The seamless transitions throughout music videos indicated that the collaborative design task was also successful in getting students to consistently use synchronization strategies.

Students employed a variety of strategies to ensure that their precisely timed video segments flowed together smoothly. Of course, some students' code was more sophisticated than others. Students brand new to Scratch had simpler code and visual effects than those who had spent one prior workshop in Scratch, especially those five sophomore mentors who had participated in more than four prior workshops in Scratch. Yet even the more experienced sophomore mentors expressed that they learned something new through this design challenge. Some had not used the "wait until" command before and those that had realized that they had not used it effectively. As sophomore mentor Jacob pointed out, in earlier projects he had synchronized everything with *wait* commands, but this strategy was "a lot more messy and took a lot more time" than using the more advanced *wait until* command. This shows an advancement in use of synchronization strategies. Thus the music mashup challenge promoted learning at multiple levels, from programming novices to more advanced students. This raises the question of *how* students learned initialization and synchronization strategies within their small collectives' group negotiations? To answer this question, we turn to a case study of one programmer collective: the Jane Doe group.

Case Study of Jane Doe Programmer Collective

The Jane Doe group's collaboration is emblematic of how the design task and collaborative team structure (e.g. the programmers' collective) together led to an intricate and well-produced final music video, with the group relying on each other as well as the nested collectives of the larger workshop group and online community to finish their program. Comprised of five girls and one boy—Marjorie, Leslie, Aris, Bridget, Natalie, and Oswald—the Jane Doe group chose to make their music video mashup to the song "Jane Doe" by *Never Shout Never*. The song describes a boy falling in love at first sight with a girl whose name he doesn't know. Joined by a common interest in the song, the students formed the group and divided the song into seven main parts with one member in charge of each segment (see Figure 1), and one additional segment (a musical interlude) unassigned at the beginning.

In their early sessions, the collective faced the challenge of creating separate roles as well as an underlying aesthetic theme for the group. They were unsure of how to effectively parse a song less than two minutes long while maintaining consistency in terms of the aesthetic design across their individual contributions. The group eventually rallied around the idea of dividing the song into six short segments based on natural breaks in the lyrics and giving each section its own unique look and feel while keeping the theme of an unidentified girl (a "Jane Doe") throughout. Early on in the process, one of the students, Oswald, asked for the shortest part of the song. In exchange he agreed to consolidate all of the sections when the time came. The students proceeded to program their own sections with regular feedback from each other until the deadline for the first submission online when Oswald led the integration of sections. This became a key moment in their developing understanding of initialization and synchronization.

The most demanding programming happened when the group first merged all of their programs into one music video mashup. In her reflection, Leslie acknowledged that this was most challenging part since they discovered that suddenly all the sprites from each segment showed up on the screen at the same time. She explained that the team resolved this by initializing each sprite to *hide* and only *show* when the timing was appropriate, explaining, "Um, so we hid all the sprites and we used a timer to make sure everything comes up at the right time." As predicted, merging all of the segmented programs brought the issue of initialization to the fore. By the end of the project, each student in the Jane Doe group consistently initialized all of their variables in every relevant aspect including location, layering, and appearance so that the final sections flowed smoothly and seamlessly.

The group also faced some organizational programming challenges, especially in regard to the synchronization of the project. In keeping track of who was doing what, early on the group employed some basic organizational strategies like allocating specific timings, maintaining a sheet that held those timings, and

giving everyone the space to be creative by working on their individual parts. Eventually those written down timings came in handy as the group centralized the main synchronization and control of the program in the “stage,” a centralized programming strategy that helped eliminate the confusion of code spread more diffusely. Within each individual program the biggest sections were synchronized with the *wait until [timer > X]* command, and within those blocks of code students used *wait X seconds* or *broadcasts* to control smaller actions. That the group worked well together, despite the challenges they faced, is evidenced in that this collective was the first to finish their coding mashup and managed to merge their program into one final music video without the support of teachers or mentors.

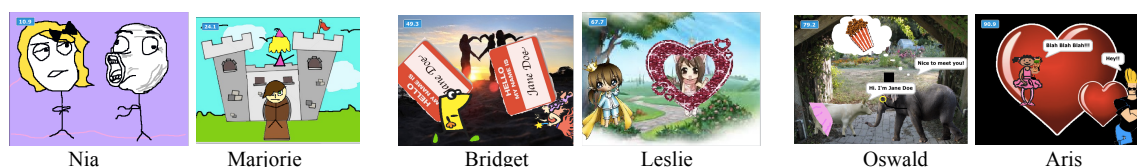


Figure 1. Screenshots from each student’s scene in Jane Doe

Interestingly, through the conversations that the group had over the course of the workshop as well as their final reflective interviews, it is evident that being in a group shaped their decision-making as well as their sense of contribution and learning. Students stated that working in a group helped them to learn more. As Marjorie expressed, “I guess I learned more from my group mates and like in the beginning I felt I didn’t know that much, like the first workshop. Then I was with my group and I sort of learned more than I did the first time.” Group members also got ideas from each other, and not only from each other but from the larger collective of the workshop group as a whole. As an illustration, two members who were also friends, Marjorie and Leslie, opted to work together on the musical interlude in the song because they had finished their individual segments early. They developed the concept and the code together, with Leslie explaining that in this collaboration her partner “thought of the ukulele switching back and forth and I thought to have the music notes like change colors and bounce.” Taking turns at the computer they together programmed the interlude with repeating patterns that were unique to other segments of the mashup. Oswald’s section of the mashup further reveals the role the larger collective of the workshop played in the group’s design. Initially Oswald was not sure what to include in his music video. In the third week of the course, he was cajoled by his group to put some action into his segments and then one of the undergrad interns helped him brainstorm, eventually giving the idea to use animals as his main sprites. In week five, one of the teachers asked Oswald to demo his section of the music mashup, and afterwards students were asked to give him feedback. During this time, students from the larger workshop and one of the teachers shared different ideas related to imagery and additional action that Oswald could include. He eventually opted to incorporate a combination of these ideas into his final segment. (Notice the popcorn thought bubble in Figure 1, an idea generated from the workshop feedback session). In Oswald’s case, his immediate group members, the undergraduate mentors, and the larger workshop gave him feedback that helped to shape his and others’ final products. This points to the role of nested collectives in the group’s success: the smaller collective of Jane Doe’s individual team embedded in the larger collective of the workshop which provided external feedback, ideas, and support. Below we turn to students’ reflections on computation and collaboration across the whole workshop.

Perspectives on Collaboration, Support, and Programming

Students’ reflections captured in interviews at the end of project provide an additional perspective on computing and collaboration that centered on three aspects: (1) value of collaboration, (2) presence of mentors, and (3) attitudes towards computing. Students articulated that collaboration was central to their learning and strongly influenced their final projects. They explained that even while there were challenges to collaborating in groups there were some general benefits as well. First, they expressed ownership in the entire project, realizing that their individual sections were part of something bigger that they could not have completed independently. For example, Ashton mentioned that while he preferred individual work in general, he could not have accomplished the big music video task without the team: “Personally, I don’t like working in groups because if I do it all by myself, I know it’s going to get done... But in that kind of project, in eight weeks... if you can’t finish [your] part, how are you going to finish a four minute music video?” A similar sentiment was also expressed by Tighe who said, “You can’t just do it all by yourself and...you got to depend on your team and work together to do it because eight weeks... eight classes... It’s just hard.” These types of comments point to the authenticity of the programming task: collaboration was necessary in order to accomplish it, and students expressed pride in the final products, which were bigger than any individual could have completed alone.

In addition to acknowledging the importance of collaboration for the complexity and size of the task, students also expressed that they learned and got ideas from their group members. For example, Aris stated that working in groups was better than working individually “because you have more ideas.” Similarly, Justus

shared that he found that his group helped keep him focused instead of “lagging off” and made his ideas better: “When I worked individually, I kind of lagged off a little bit. You know, it’s just me. Also, it [the group] helped me come up with the ideas faster. But all-in-all, I had a better idea in the group.” Students also explained that an advantage to working in groups was being able to learn from their peers. For example, Belle explained that working in a group is better than working individually because “Working in a group, you learn new things that you haven’t learned before that you can’t really learn by yourself. It’s like, ‘How do I do this?’ and then you ask the person next to you and he’ll say, ‘Well, here’s an easier way.’” And you’re like, ‘Oh, I never knew that before.’” The high school mentors further explained this collaborative advantage, with one mentor pointing to the fluid group dynamics as the reason for his group’s success. In another case, the mentor explained that the group had a student who wasn’t as comfortable in Scratch and that the other group members stepped into support her, stating, “Yeah, so we all kind of were able to work as a team to help the struggling members of the team.” He also argued that through collaboration and mutual support, the entire team was able to be successful.

The presence of mentors, interns, and teachers helped students bring their ideas to life in Scratch and in the process, broaden their understanding of the different approaches to solving problems. Mentors were seen as a meaningful source of support in both learning how to navigate programming challenges and in staying on track. Students reflected that the mentors’ knowledge of Scratch provided them insights into how to approach different challenges. Belle said of her mentor, “during our making our videos, I asked my mentor a lot of questions and he taught me to use different things [in Scratch],” indicating that the mentors could help youth work out challenges through active problem solving with them. Another student Justus, explained that in one instance, his mentor helped him strategize what synchronization commands he should use for a particular part of his project using this interactive approach. In addition to helping solve problems, several students indicated that mentors were there to keep students on track or help them out when they encountered challenges such as getting stuck. Stephen, for example, explained how the mentors’ role was “pushing everybody to do their parts.” Another student, Charlotte, explained that the college intern helped her group by checking in on them: “She would check up constantly. Like, ‘hey, where are you? You know you have to be finished by a certain time and upload,’ so that was actually very helpful.” Justus, who worked with this same college intern explained that when he got distracted, “Tara [intern] was always there to reel me back into the project. She did a good job of keeping us all working on it.” These are helpful illustrations of the kind of informal and critical support that mentors, interns and teachers provided that furthered opportunities for nested collaboration in the classroom (see also Kafai, et al, 2013).

Finally, all students expressed how their understanding of programming shifted favorably through their participation in the music mashup challenge. The students’ reflections illustrated how extended time in the workshop setting gave them opportunities to practice and overcome initial frustrations or in some cases skepticism or disinterest. For example, George said, “...when you first start Scratch, it’s kind of like this is really hard, I don’t really want to do this... But like once you start getting... knowing things it’s really fun.” George expressed the change that occurs when there is time to practice. Another student, Aris, explained that her understanding of programming changed stating, “it’s not as hard as I thought it was” and noting that as an eighth grade student the year prior she would not have seen herself enjoying a workshop like Scratch. Both George and Aris’ reflections demonstrate how participation helped them to see they could be successful and that Scratch was not as hard as they initially imagined, challenging their preconceived notions about programming. Other students pointed to the possibilities available through the Scratch platform as a compelling reason to participate. For example, Stephen explained that he thought, “Scratch is very fun, because you can actually see your creative thinking to do anything you want as long as you can... like you have to work out the bugs and problems. But if you can do that, then there’s almost endless possibilities.” These same sentiments were evident in the mentor interviews. Two high school mentors argued that they were impressed with the growth they saw amongst students they worked with during the workshop. They saw students go from not knowing or understanding how to navigate the space to being much more experienced and comfortable programmers of Scratch.

Discussion

Our goal in examining the programmers’ collective was to understand better how we can bring the DIY dimensions of dispersed online creative communities to face-to-face workshops and classrooms introducing programming to high school students. There are, of course, other models of such learning communities such as the software design studio (Kafai & Burke, in press) and Knowledge Forum (Scardamalia, 2002) that embody many of the same ideas. We chose as a starting point a software design task that not only encouraged more complex programming for beginners but also invited better coordination for collaboration. Our design parameters for collaborative programming involved choosing a relevant task such as a music video mashup that fit well within popular youth media production and drove student interest in a culturally relevant manner (Enyedy, Danish & Fields, 2011). Furthermore, what seemingly looked like a trivial task became a rich programming activity when turned it into a collective programming design. Two key computational concepts, initialization and synchronization, were needed to connect different parts of the video into a smooth production.

Thus learning and implementing these concepts were instrumental to accomplishing the collective design. As such the programming project had an interdependent design structure that required individual contributions to accomplish a task larger (and more desirable) than what one could do on one's own. Motivating learners in design activities has been a longtime challenge (Collins, Joseph, & Bielaczyc, & Josephs, 2004) and relates to the current focus on interest-driven and connected learning (Ito et al., 2013).

The need for developing and researching the design of such collaborative programming tasks is crucial because pair and group programming activities have met with varying success. Practically, there are not many ways for two people to concurrently program, which is not the case with other activities like writing which can be a concurrent and collaborative activity with applications like Google Drive. However, this study helped us to think more creatively about the design task and our goal with this work. Could we have done more to ensure higher levels of group collaboration? Or was the design task collaborative enough? On the surface, it would seem that the groups only had to collaborate at the end to coordinate their projects. This approach was one effort to ensure individual learning and ownership of a product while also promoting learning from each other through creation of a shared product within a group and through regular viewing of others' work throughout the workshop as well as videos online. As we saw in the findings, the design/structure of the physical space (e.g., kids seated with their groups) and assigning mentors and interns to groups, coupled with the opportunities for informal and formal feedback at multiple levels did in fact lead to high levels of collaboration. Throughout our observations, we found that simply sitting together allowed youth to ask questions to their teammates. In addition the structure of the class, where there was a lot of additional support available, allowed kids to request support but also receive help when mentors, interns and teachers informally circulated the room to check-in with groups. Students also had multiple sources for formal feedback, both online and offline (e.g., during group shares), actually resulting in changes to their final products. So, by designing a collaborative task and also creating a collaborative space, the workshop encouraged different kinds of collaboration. Evidence suggests that we were successful in promoting individual learning, ownership, and engagement in the programming task while also promoting a sense of learning from each other and collective ownership of a final programming project.

Beyond the classroom, the ability to learn to work with others in today's highly distributed, networked world is paramount. What we called collaborative agency in an earlier paper (Kafai et al, 2012) speaks to students' ability to self-form collectives or collabs to accomplish larger tasks than they could on their own. Yet like others before us, we found that while students can often figure out how to accomplish problems, they do not necessarily equitably divide tasks and often reinforce stereotypes about who has expertise in what. Thus in this paper we addressed a key challenge of collaboration: that of equity in individual participation and contribution. Note that we stress equity and equality in collective activities, where members contribute based on their expertise and distribute responsibilities for programming to everyone, not just relying on those perceived most capable. Expanding on our earlier definition (Kafai et al, 2012), collaborative agency means that members of a team need to reach out to each other and coordinate a task equitably, precisely what the music video task required and scaffolded through its design. Finally, we suggest the idea of *nested collectives* to acknowledge the instrumental role that not just group members but others such as mentors, teachers, and students in the larger workshop played in supporting youth programming activities. Not only this, but these programming collectives were nested within the larger online challenge and students gained much from online feedback and from seeing others' projects from around the world. Like a community of learners in Lave & Wenger's sense (1991), we brought in several layers of more advanced members, including mentors from the same school with one additional year of experience, undergrad mentors from a local university, teachers, and active members of the online Scratch community.

References

- Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. New Haven and London: Yale University Press.
- Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 509-535). Thousand Oaks, CA: Sage Publications.
- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the learning sciences*, 13(1), 15-42.
- Denner, J., and L. Werner. "Computer Programming in Middle School: How Pairs Respond to Challenges." *Journal of Education Computing Research* 37, no. 2 (2007): 131-50.
- DuMont, M. & Fields, D. A. (2013). Hybrid shmybrid: Using collaborative structure to understand the relationship between virtual and tangible elements of a computational craft. 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 2, Short Papers, Panels, Posters, Demos & Community Events*. International Society of the Learning Sciences: Madison, WI, 233-234.

- Enyedy, N., Danish, J. & Fields, D. A. (2011). Negotiating the “relevant” in culturally relevant mathematics. *Canadian Journal of Mathematics*, 11(3), 273-291.
- Fadjo, C. L. (2012). Developing Computational Thinking Through Grounded Embodied Cognition. Unpublished dissertation. Columbia University.
- Ito, M., Baumer, S., Bittanti, M., boyd, d., Cody, R., Herr, B., Horst, H. A., Lange, P. G., Mahendran, D., Martinez, K., Pascoe, C. J., Perkel, D., Robinson, L., Sims, C., & Tripp, L. (2010). *Hanging out, messing around, geeking out: Living and learning with new media*. Cambridge, MA: MIT Press.
- Kafai, Y. B. & Burke, W. Q. (in press). *Connected Code: Children as the Programmers, Designers and Makers for the 21st century*. Cambridge, MA: MIT Press.
- Kafai, Y. B., Griffin, J., Burke, W. Q. Slattery, M., Fields, D. A., Powell, R., Grab, M., Davidson, S. B. & Sun, J. S. (2013). Broadening participation in and perceptions of computing through cascading mentoring: Implementing a CS community service learning course. In *Proceeding of the 44th ACM technical symposium on Computer science education (SIGCSE '13)*. ACM, New York, NY, USA, 101-106.
- Kafai, Y. B., Fields, D. A., Roque, R., Burke, W. Q., & Monroy-Hernández, A. (2012). Collaborative agency in youth online and offline creative production in Scratch. *Research and Practice in Technology Enhanced Learning*, 7(2), 63-87.
- Luther, K., Caine, K., Ziegler, K., & Bruckman, A. (2010). Why it works (when it works): Success factors in online creative collaboration. In *GROUP '10: Proceedings of the ACM Conference on Supporting Group Work* (pp. 1–10) New York: ACM Press.
- Maloney, J., Peppler, K., Kafai, Y., Resnick, M., & Rusk, N. (2008). *Programming by Choice. Urban Youth Learning Programming with Scratch*. Paper presented at the SIGCSE 2008 Conference, Portland, Oregon.
- Monroy-Hernandez, A. (2012). Designing for remixing: Supporting an online community of amateur creators. *Unpublished dissertation*. MIT.
- Porter, L. Guzdial, M., McDowell, C. & Simon, B. (2013). Success in introductory programming: what works? *Communications of ACM*, 56(8), 34-36.
- Resnick, M., Maloney, J., Hernández, A. M., Rusk, N., Eastmond, E., Brennan, K., Millner, A. D., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. B. (2009). Scratch: Programming for everyone. *Communications of the ACM*, 52(11), 60–67.
- Robbins, A., Rountree, J., & Rountree, N. (2003). Learning and Teaching Programming: A Review and Discussion. *Computer Science Education*, 13(2), 137-72.
- Roque, R., Kafai, Y. B., & Fields, D. A. (2012). From tools to communities: Designs to support online creative collaboration in Scratch. *Proceedings of IDC 2012*. Bremen, Germany.
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Eds.), *Liberal education in a knowledge society* (pp. 67–98). Chicago: Open Court.
- Soloway, E., and Spohrer, J. (1990). *Empirical studies of novice programmers*. Norwood, NJ: Ablex Publishing.
- Webb, N. Ender, P. & Lewis, S. (1986). Problem solving strategies and group processes in small groups learning computer programming. *American Education Research Journal*, 23, 248- 262.

Acknowledgments

This material is based upon work supported by the National Science Foundation (NSF-CDI-1027736) to Mitchel Resnick, Yasmin Kafai and Yochai Benkler. The views expressed are those of the authors and do not necessarily represent the views of the National Science Foundation, Utah State University, or the University of Pennsylvania.

A Tale of Two Worlds: Using Bifocal Modeling to Find and Resolve “Discrepant Events” Between Physical Experiments and Virtual Models in Biology

Tamar Fuhrmann, Shima Salehi and Paulo Blikstein

Transformative Learning Technologies Lab, Stanford University, 520 Galvez Mall, CERAS 232, CA, USA.

tamarrf@gmail.com, salehi@stanford.edu, paulob@stanford.edu

Abstract: In this paper we demonstrate an approach to supporting students’ engagement in combined physical experimentation and virtual modeling. We present a study that utilizes a scientific inquiry framework which links students’ physical experimentation with their use of computer modeling in real time, which we call “Bifocal Modeling.” In the case of the Bifocal Modeling activities discussed here, a group of high-school students designed computer models of bacterial growth through reference to a physical experiment they were conducting, and they were able to validate the effectiveness of their model against the results from their experiment. Our findings suggests that as students compare their virtual models with physical experiments, they encounter “discrepant events” that contradict their existing conceptions and elicit a disequilibrium. This experience of conflict encourages students to further examine their ideas and hypothesis, seek more accurate explanations of the observed natural phenomena, improving the design of their computer models.

Introduction

In recent years virtual experimentation has received increasing emphasis as an alternative to conducting experiments in a physical environment. Much of the research in this area has focused on the question of the advantages of virtual experiments over physical experimentation (Jaakkola & Nurmi, 2008; Klahr et al., 2007), but more recently researchers have started to examine the effects of combining the virtual and physical modalities (Zacharia et al., 2008; Jaakkola et al., 2011; Liu, 2006). Additionally, studies report that alternation between these distinct experimental modalities in the course of individual experiments can often improve learning outcomes (Zacharia, 2012; Smith, 2010; Gire et al., 2010). Zacharia and colleagues (2008, 2012) suggested that the best way to develop a framework portraying the optimal combinations of physical and virtual manipulation is to employ the learning objectives of each experiment as the criteria for blending them. Nevertheless, there are two under researched areas in this literature. The first is that most of the virtual experiments were simulated versions of a physical experiment, often mimicking the appearance of the lab equipment, with the goal of trying to make the physical and virtual experiments as similar as possible (Blikstein, 2014). The second is that the research has mostly focused on predesigned physical and virtual experimentation. Simulation tools have been among the preferred means for providing environments for virtual experiments, but the rules and models behind these simulations often remain hidden from the students. Recent advances in inquiry learning research have sought to implement activities in which virtual experimentation is supplemented by opportunities to design computer models (Mulder et al., 2011), and the aim of our work is to examine the learning outcomes of designing these computer models that are *explicitly meant to be different* than the physical ones, in order to promote students’ critical stance towards their own models and hypothesis. The creation and critical evaluation of models are important components of scientific practice which have been increasingly recognized as a valued educational goal (Levy & Wilensky, 2008; Blikstein, 2010). Predesigned models can scaffold and direct students to attend to relevant aspects of a phenomenon, but they do not offer students opportunities to externalize and debug their models, or to evaluate their assumptions and their limitations.

Bifocal Modeling (Blikstein et. al., 2012; Blikstein, 2010, 2014) is an approach to inquiry-driven science learning that challenges students to design and compare in real time physical and virtual models in order to identify their respective differences and limitations. In these activities, students explore scientific phenomena such as heat diffusion, the properties of gases, and wave propagation by conducting physical experiments, designing virtual models, and connecting the experiments with the models in real time through iterative comparisons with empirical data. During the physical phase of the process, students design and develop their physical experiment, and they run the experiment while collecting data with embedded sensors or a time lapse camera. Concurrently, they design and develop a virtual model for the same phenomenon, and compare the behavior of the virtual model with their observations from their physical experiment (figure 1). When they identify a discrepancy, students have the opportunity to redesign their models and re-iterate the process.



Figure 1. Two examples of Bifocal Models made by students for gas laws (left) and Newton's cradle (right)

Depending on the nature of the phenomenon, for a bifocal activity, students can use different programming languages to implement their virtual models. A common implementation software is NetLogo (Wilensky, 1999), a free, open-source environment for scientific modeling. NetLogo models typically consist of a set of autonomous agents (such as gas particles or bacteria entities) moving through a virtual world and interacting to produce emergent outcomes. Students define the variables for both the agents and their worlds, and specify sets of rules for agent-level behavior, such as, "if two gas particles collide, they exchange energy and bounce off each other." The students' goal is to build a model with a behavior that matches or imitates the physical data they collect. Through the comparison of the design of the virtual model and their experimental data, the students engage in the discovery of discrepancies between the results of each modality. Piaget (1985) argued that to foster conceptual change students must be confronted with "discrepant events" that contradict their conceptions and invoke a "disequilibrium or cognitive conflict". Following the forms of equilibration in Piaget's theory, researchers (Hewson, 1988) identified two distinct types of cognitive conflicts: the conflict between the internal and external worlds of a student's conception and experiences, and the purely conceptual conflict between two different cognitive structures related to the same phenomenon. In our study, we found that when students were instructed to design a virtual model that imitates the bacterial growth curve, they used their previously acquired knowledge about the curve and the physical appearance of the bacteria in the Petri dish as reference patterns to indicate what their model should generate. When the virtual model data did not match the observed data, they were confronted by the discrepancy between the physical and conceptual worlds which led to the conceptual mismatch between two cognitive structures related to the same phenomenon.

Bifocal Modeling Framework

Given that Bifocal Modeling comprises many different tools and techniques, there are multiple possibilities for classroom implementation of each modality. To structure our studies, we divided the physical and the virtual assignments into a sequence of shorter activities (Figure 2):



Figure 2. General structure of a Bifocal Modeling activity (Blikstein et al., 2012)

- Design:** Students select research question, plan their observation, generate hypotheses, and design experiments and virtual models that will potentially confirm them. In designing the virtual model, students define variables, and conceptualize micro-rules or equations to describe the phenomenon.
- Construct:** Students structure both their physical experiment and virtual model.
- Interact:** Students interact with their experiments through direct observation or embedded sensors/cameras, and interact with their computer models by changing parameters, and recording data.

Subject Matter of the Activity: Bacterial Growth

We chose bacterial growth for their simple cellular structure, rapid reproduction, and complex ecological dynamics. The goal was for students to recognize the four distinctive patterns of the bacterial growth curve (Figure 3), understand the variables underlying them, and explain the underlying mechanisms of each stage.

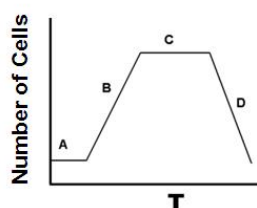


Figure 3. Bacterial growth curve over time (the number of cells is in the log scale)

- A. *Lag phase*: Population remains temporarily unchanged; in this phase the bacteria are adjusting to their new environment, repressing or inducing enzyme synthesis, and initiating chromosome replication.
- B. *Log phase*: Bacteria growth proceeds by the division of bacterium into pairs in a process known as “binary fission.” Exponential growth cannot continue indefinitely because the medium is soon depleted of nutrients, which are replaced by waste products.
- C. *Stationary phase*: The population remains constant because the bacterial growth rate is equivalent to the death rate.
- D. *Death phase*: In this final stage, the bacteria have exhausted their nutrients, lose their ability to divide, and die off. As in the rapid growth phase, the decay pattern characterizing the death phase is exponential (1)

Methods

Audience and Resources

The study was conducted with 13 students (4 females and 9 males) and lasted for a total of three days in a university laboratory setting. Students came from a 70% minority high school located in a predominantly Latino urban setting and volunteered for a 4-week, 30 hour/week digital fabrication workshop at a local university. This workshop took place during the school’s intersession, during which all students were required to enroll in a month-long extracurricular course or internship outside of school. The selection of students was governed by a complex allocation system developed by the school; consequently, since not all students ended up being able to enroll in their preferred choices, the final group was rather diverse in terms of school achievement. All students were videotaped and recorded during all activities, their computer usage was logged with screen-capture software, the researchers conducted informal interviews and kept field notes, and all student notes and sketches were saved. The entire data (15 hours of video recordings) was analyzed by the researchers following transcriptions. Episodes explicitly showing moments of comparison between the virtual model and the real experimentation were the focus of this study. In these episodes, in order to reveal the discrepancies, we analyzed the content and the context of the situations in the videos to identify iterative moments of comparison.

Instructional Sequence

The total time designated for the activity was 15 hours, divided across four smaller activities (Figure 4):

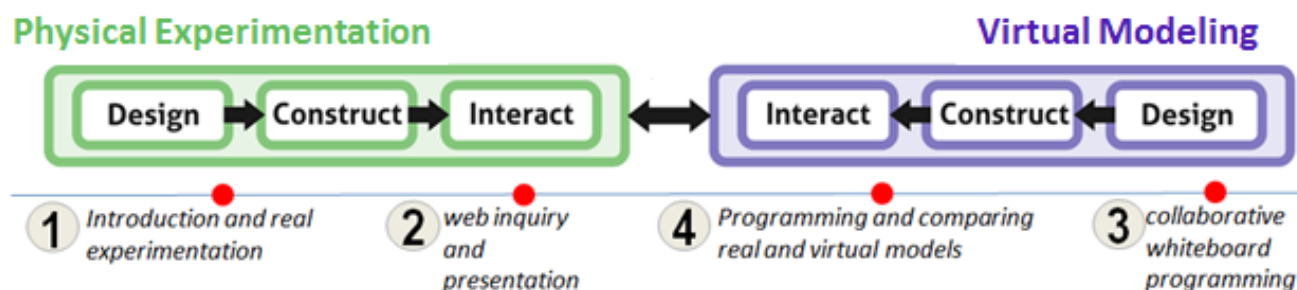


Figure 4. Bacteria Growth-Bifocal Modeling – the four activities

1. *Introduction and physical experimentation.* After the class was given an introduction to bacterial growth, the students’ had to grow real bacteria. They prepared a Petri dish with agar and collected a bacteria sample from an object likely to be contaminated (e.g., a door knob, keyboard, toilet.). They were provided a time-lapse camera to capture images at 30-minute intervals over seven days. The images were automatically compiled into a video. In response to time restrictions, we also showed the students a video of a bacterial growth experiment conducted previously by the research team in the same lab with the same toolkit.
2. *Web inquiry and presentation.* Students were grouped into pairs and asked to make a list of questions about bacteria and bacterial growth. They were asked to conduct web research to answer their questions, and presented their information to the entire class in short slideshows.
3. *Collaborative “whiteboard programming.”* Students were divided into three groups, each group with a dedicated facilitator from the research team. The task for the groups was to come up with rules that govern bacteria growth. First, students listed all variables that they thought would affect bacterial growth. Next, the facilitator proposed the iterative building of a block-based “computer” program on the whiteboard (figure 5), in which the students should generate the main stages in bacterial growth, as well as account for how each stage would develop, and how the variables would interact. After three hours of “whiteboard modeling,” the groups split, and shared their ideas with two members from the

other groups in a 45-minute discussion panel. After receiving feedback on their initial ideas from members of other groups, the groups reconvened and began programming their virtual experiment in NetLogo.

4. *Programming and comparing experiments and virtual models.* In this last phase, the facilitator would sit before the each group in front of a large television used for displaying code; the facilitator's role here was to “translate” the ideas of the students into NetLogo code. These last three hours of the study were dedicated to coding the students' virtual model, and comparing the coding results with the data collected from the experiment in the Petri dish. Students discussed the results, developed hypotheses for approaches to the validation of both models, and made changes to the virtual model in order to make it similar to the real bacterial growth observed in the Petri dish.

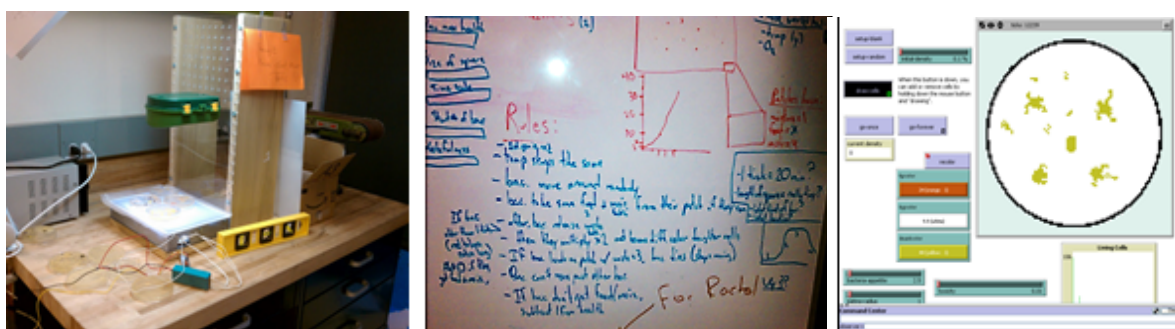


Figure 5. Physical experiment with time-lapse camera, “whiteboard modeling,” and virtual model

Data and Discussion

During the unit, students created a whiteboard model; translated the whiteboard rules into a model's specifications, “ran” the model to envision how bacteria would multiply according to the model; and compared the modeled results with both the observed real Petri dish growth patterns and with the growth curve the students generated from the physical data. Finally, students refined the virtual model by adding rules and variables to address the perceived differences between the model and the experiment. In figure 6, we present a chronological list of the additions made by the students in one of the groups (4 students). This group repeated this processes a total of four times during the 1.5 hours of the final session (activity four) and, in the process; these students increased the accuracy of the model.

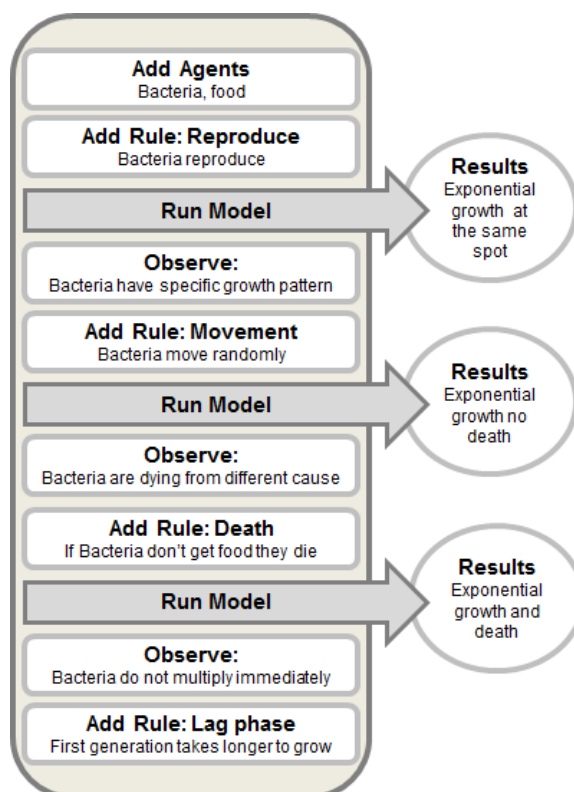


Figure 6. Chronological list of the additions made by the students to the model and instances in which they ran it

In what follows, we present three sample episodes that demonstrate this iterative process by one of the groups. This collection is presented as a representative example of the comparison moments and the discrepancy events the students encountered during their participation in the activities.

Episode # 1: How Do Bacteria Move?

Context

In the first episode, students were in the initial stages of designing their virtual model on the whiteboard. They decided to add agents such as bacteria and food, as well as a rule regarding reproduction. However, they remained uncertain about the mechanism for bacterial motion: do bacteria move at all? Do they move toward food? Do they move randomly? How do they actually move? Do they recognize one another?

Following is students' design process which we broke down into three phases for the purposes of this analysis: design, observe, and revise.

Design Process

- **Design:** as a first step, students made their virtual bacteria reproduce. 'Running' this model resulted in exponential growth. Each tick (time step in NetLogo) resulted in an increased bacteria population confined to its original location on the virtual Petri dish.
- **Observe:** while corroborating the virtual model with the experiment in the Petri dish, students observed a mismatch; they saw that in the Petri dish, real bacteria colonies did not resemble the virtual colonies of their model. In the experiment bacteria appear to have a specific and unique growth pattern. They do not grow on top of each other; rather they spread throughout the dish. At this point students asked questions and sought to explain the phenomenon. How do bacteria's unique patterns develop?
- **Revise:** students added a new rule to their virtual model, which helped them simulated bacteria moving randomly over the virtual Petri dish, resulting in colonies that spread across the virtual dish in a pattern resembling that of the real dish.

Excerpts from the Episode

Student 1: Do we know if they move around randomly, though, or –
Student 2: How else would they move around?
Student 4: Maybe in specific ways that we could understand ...
Student 1: I guess, like, where they scooted, go toward the food, but it could just do that...
Student 2: What makes you think this is fine or not? How do you know?
Student 2: I think it doesn't really matter how they move.
Instructor: Doesn't really matter, for what?
Student 2: What do you mean?
Student 1: Like, how do you know it doesn't really matter, you know?
Student 2: Well, I mean, they'll eventually find food by moving randomly

Researcher's Observation

In this episode students face a specific conflict while comparing the virtual design results to the real colonies in the Petri dish; they discovered that bacteria (in the physical Petri dish) do not reproduce in the same location; rather, they migrate. This comparison and observation of the experimental results to the modeling suggested the idea that bacteria do not grow on top of each other. While observing the real Petri dish and confronting the mismatch with their virtual modeling results, students discussed possible mechanisms for bacterial motion in order to debug their model, which, in turn, made them seek explanations for the natural process. The discussion then progressed to physical mechanisms that might assist bacterial motion (e.g. the bacteria might be flagellates with whip like cilia at their anterior ends.)

Episode # 2: Do Bacteria Have an Infinite Life Cycle?

Context

While examining and running their virtual model, the students discovered that the bacteria would never die. Students then sought to understand how to make bacteria die by manipulating their food resources. After facing this discrepancy between what was observed in their virtual modeling as opposed to the physical experimentation, they began questioning the issues of bacterial death and life cycle. Following is students' design process which we broke down into the same three phases.

Design Process

- **Design:** students made their virtual bacteria move randomly and reproduce after encountering specific environmental conditions (food, water, etc.)

- *Observe:* It appears that the bacteria in the physical experiment are dying from different causes. Bacteria do not live forever. Students observed that in the Petri dish, the bacteria colony remained the same size for several days. Additionally, they noticed the bacterial growth curve, which includes a death phase.
- *Revise:* Students added the “food” variable. The corresponding rule is that when food is exhausted, and no new food resources become available, the bacteria die. The new design included bacteria that do not have unlimited lifespans.

Excerpts from the Episode:

Student 2: Look at the death.
 Student 1: Death?
 Student 4: What should happen is that they run out of food.
 Student 3: Okay. How should we – how should we do that? Can we make some – write some imaginary code for that?
 Student 2: Made some of the patches disappear [patches in the NetLogo code represent food]
 Instructor: So can you give me more? Imagine that I’m, like, really like a dumb computer. You need to tell me the steps I need to take. Is it, like, when all that is gone, then they all die?
 Student 1: The bacteria.
 Student 4: They slowly die. They still reproduce, but they slowly die.
 Student 1: Okay. And when is it, like, every tick or...?
 Student 2: Every ten.
 Student 3: If all patches – all 100 patches are gone, then bacteria die.
 Instructor: If all food is gone, then all bacteria die. Okay. Let’s run the model in our heads and think about how we’re going to do it. So all the food is gone...eventually when they run out of food, boom, they die. They all die. That’s the code we have right now.

Researcher’s Observation

In this episode the students explore the significant effect of food resources on the bacteria population. In this specific excerpt of the dialog, the students were asked to think of a way to “translate” the role of resources into code in Netlogo. They use “patches” (2) as “bacteria’s food”, and explained that when all food is gone, the bacteria die.

Episode # 3: Is There a Lag Phase in Bacteria Growth Pattern?

Context

At one point after “running” the virtual model, a student observed a mismatch: the growth curve was increasing exponentially from the start. She noted that this finding was incorrect because the real growth curve had an initial flat “lag phase” before beginning to grow. After a long discussion between group members, the students attempted to explain the lag phase of the bacterial growth, which commenced with the inoculation of the Petri dish.

Following is students’ design process:

Design Process

- *Design:* students made their virtual bacteria begin reproducing as soon as they are introduced into the Petri dish.
- *Observe:* In comparing with the physical experiment, students become aware of the lag time that occurs before the bacterial reproduction becomes apparent. The students discovered that it takes about five days before they are able to detect a colony on their Petri dish. The growth of the microscopic bacteria remains invisible until the population grows into the millions, at which point the colony has become sufficiently large to be visible. This discovery led the students to realize that specific conditions must be met for bacteria to reproduce.
- *Revise:* Students add variables and rank them so their modeled bacteria will reproduce only under favorable conditions. For example, if the model’s food value is greater than 10, the bacteria will reproduce. If this value is less than 10, the bacteria will first enter a lag phase.

Excerpts from the Episode

Instructor: What about the, I am asking again because I’m really trying to make a point here, remember they didn’t start like this in the graph? They didn’t just reproduce? ... and we did it like that and we had this phase which they don’t change, ... yeah. What happen there?

Student 4: The lag?
 Student 3: What is happening? Yeah, what is happening to them, the bacteria in real bacteria dish?
 Student 2: Because it takes a while for it to form and like reproduce, as soon as they get the hang of it, they're like, yeah, to make more.
 Student 3: So they get used to their, like they get used to their environment.
 Student 4: Their place.
 Instructor: So how can we do it in program? What do we need to add there?
 Student 3: Maybe like a spurt where they're having a bunch of babies and they kind of stop having babies, then they start having babies again. Have it slowly...
 Student 4: Slowly so they won't start at the beginning?
 Student 2: Yeah then they start and then they don't and then they start.
 Student 1: Are you trying to make it like this?
 Instructor: How can we turn this idea of the lag phase into a code?
 Student 2: I guess we can use a wait about like twenty ticks oh that's a lot, a lot of wait like ten ticks to get used to the environment so they can just say wait... ten days before starting to reproduce?.

Instructor: Good idea.

Researcher's Observation

It took time for students to realize that there is "lag phase" at the outset of the bacterial growth process. In the real Petri dish, it took five days before the students observed visible alterations and growth. However, in students' initial model bacteria grew and reproduced immediately. After the comparison between their computer model and the results of both the experiment and a bacterial growth curve, students realized that the initial stage of the physical experiment evidenced no apparent change in bacteria population. This conflict engaged the students in rethinking the phenomena they were attempting to model, and it also led them to revise their model according to their observations of the physical experiment. In order to succeed in this task, they had to find an explanation for the stable phase for inclusion in their virtual model in order to achieve a better correspondence with their observed results. In the process of generating a virtual model that better emulated the phenomena, the students added behavior parameters and behavior sequences in ways that related explicitly to real-world behaviors or included real-world constraints. In addition, they conducted similar processes to add the other relevant variables to their model.

Conclusion

This study demonstrates one of the main elements of the Bifocal Modeling framework: how discrepancies between a virtual model and physical experimentation can be generative. One of the main features of this framework is the explicit comparison of virtual models designed by students with physical experiment in real time. Our results suggest that the use of physical experimentation as a reference pattern in the creation and refinement of the virtual model is effective. In designing a virtual model that recreated the bacterial growth curve, students used their previously acquired knowledge about the curve and the physical cues of the bacteria colonies as an initial reference pattern, which indicated what their model should generate. When the model behaviors did not match the observed ones, students faced a discrepant event that required resolution (Piaget, 1985; Hewson, 1988). This mismatch led to debugging (Papert, 1980) and encouraged students to question their assumptions, to rethink the results, to consider alternative conceptions. In this process, students were actively engaged in hypothesis generation and testing. Throughout the entire activity, students acquired specific and detailed evidence regarding the behavior of bacteria, but the value of this evidence almost exclusively became apparent to them during their attempt to make their virtual model match their empirical observations. During their web research and physical experiment, the students took note of the fact that bacteria grow in specific patterns, do not remain in their original spot, and do not grow indefinitely. All these detailed bits of evidence would have remained inert knowledge had the students not engaged in "model matching"—and as a result, the students acknowledged this evidence opportunely when such knowledge was needed to design a more accurate virtual model.

In addition, without the real-time comparison with the physical model, their virtual model would have contained a number of misconceptions. Studies report that students are capable of designing correct models (Mulder, 2011), but that they often fail to relate their knowledge of natural events to their models (e.g., Sins, Savelsbergh, & van Joolingen, 2005). This incomplete understanding of the modeling process could present serious obstacles to learning through modeling. Traditionally, researchers have tried to argue that by making virtual models very similar to physical phenomena, and thus backgrounding their differences, learners could achieve equally in both virtual and physical experimentation. Conversely, we argue that these differences should be foregrounded and made apparent to students, and that combining the virtual and physical modalities and

encouraging students to look for mismatches offers a promising way to make learning with models more effective.

What is more, the process of model comparison encouraged the students to become engaged in the discovery of “discrepant events” in a manner that is congruent with scientific professional practice. Students’ desire to “fix” their models developed spontaneously throughout the activity. Even though we acknowledge that further research is needed to fully validate our framework, the data seems to suggest that the main feature of Bifocal Modeling—real-time model comparison—was an effective in the generation of model debugging moments that engaged students in rich, agentive, and generative intellectual work.

Endnotes

- (1) All phases can be slowed down by lowering the temperature – leaving the culture at an optimum temperature for growth over long periods will simply accelerate the death of the culture. Most dead bacteria cells closely resemble living cells, so a normal appearance of a colony on the Petri dish is no indication that its cells are actually alive.
- (2) The NetLogo world is a two dimensional world that is made up of “turtles” (moveable agents) and patches (stationary agents). The patches are the “ground” over which the turtles move.

References

- Blikstein, P. (2010). Connecting the science classroom and tangible interfaces: the Bifocal Modeling framework. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences (ICLS 2010)* (pp. 128-130). University of Illinois at Chicago: International Society of the Learning Sciences: Chicago, IL.
- Blikstein, P., Fuhrmann, T., Greene, D., & Salehi, S. (2012). Bifocal modeling: mixing real and virtual labs for advanced science learning. In *Proceedings of the 11th International Conference on Interaction Design and Children* (pp. 296-299). ACM.
- Blikstein, P. (2014). Bifocal Modeling: Promoting Authentic Scientific Inquiry Through Exploring and Comparing Real and Ideal Systems Linked in Real-Time. In A. Nijholt (Ed.), *Playful User Interfaces* (pp. 317-352): Springer Singapore.
- Gire, E., Carmichael, A., Chini, J. J., Rouinfar, A., Rebello, S., Smith, G., & Puntambekar, S. (2010). The Effects of Physical and Virtual Manipulatives on Students’ Conceptual Learning about Pulleys. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences (ICLS 2010)* (pp. 937-943). University of Illinois at Chicago: International Society of the Learning Sciences: Chicago, IL.
- Hewson, P., & Hewson, M. (1984). The role of conceptual conflict in conceptual change and the design of science instruction *Instructional Science* 13 (1), *American Journal of Physics*, Vol. 71, 618-29.
- Jaakkola, T., Nurmi, S., & Veermans, K. (2011). A comparison of students' conceptual understanding of electric circuits in simulation only and simulation-laboratory contexts. *Journal of research in science teaching*, 48(1), 71-93.
- Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching*, 44(1), 183-203.
- Liu, X. (2006). Effects of combined hands-on laboratory and computer modeling on student learning of gas laws: A quasi-experimental study. *Journal of Science Education and Technology*, vol. 15, 89-100.
- Levy, S. T., & Wilensky, U. (2008). Inventing a “mid-level” to make ends meet: Reasoning through the levels of complexity. *Cognition and Instruction*, 26(1), 1-47.
- Mulder, Y. G., Lazonder, A. W., & De Jong, T. (2011). Comparing two types of model progression in an inquiry learning environment with modelling facilities. *Learning and Instruction*, 21, 614-624.
- Olympiou, G., & Zacharia, Z. C. (2012). Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Science Education*, 96(1), 21-47.
- Piaget, J., (1985). *The equilibration of cognitive structures* (B. Terrance & K. J. Thampy, Trans.) Chicago, IL:University of Chicago Press.
- Sins, P., Savelsbergh, E., & Van Joolingen, W. (2005). The Difficult Process of Scientific Modelling: An analysis of novices' reasoning during computer based modeling. *International Journal of Science Education*, 27(14), 1695-1721.
- Wilensky, U. (1999). *NetLogo* [computer software]. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
- Zacharia, Z. C. & Constantinou, C. P. (2008). Comparing the influence of physical and virtual manipulatives in the context of the Physics by Inquiry curriculum: The case of undergraduate students’ conceptual understanding of heat and temperature. *American Journal of Physics*, 76(4&5), 425-430.

“Deep Hanging”: Mentors Learning and Teaching in Practice

Déana Scipio, University of Washington, Box 354941, WA 98195-3600, descipio@uw.edu

Abstract: This paper is a comparative case study focused on the experiences of mentors in a chemical oceanography afterschool program. The study explores “deep hanging”—the term mentors used to describe their experiences learning about the culture of science, scientific research processes, as well as learning they could be- and become- scientists. “Deep hanging” entails authentic tasks in rich contexts, providing access, capitalizing on opportunity, and building interpersonal relationships. Data include reconstructive history interviews with mentors and video of their interactions with youth in the afterschool program. The conceptual framework for this paper explores the ways that constellations of situated events lead to changes in sociomaterial practices overtime. Findings suggest that surfacing and exploring mentors’ self-position with respect to STEM is crucial to understanding how they will position youth. These findings have implications for the design of learning environments that seek to broaden participation for non-dominant groups in the sciences.

Major Issues Addressed

Bringing educators and experts together to create rich disciplinary focused learning environments for youth has the potential to broaden youths’ pictures of the types of people who can do science (Barab & Hay, 2001; Hsu et al., 2009; Polman & Miller, 2010). These programs operationalize learning in practice by making ways of knowing and doing in the sciences visible to youth and engaging them in contemporary scientific practices. Youth from non-dominant communities face barriers to Science, Technology, Engineering, and Mathematics (STEM) learning based upon the ways that they are typically positioned with respect to the domain of science. Positionality with respect to science domains and practices interacts with race, Socio-Economic Status (SES), gender, English language learning, epistemological commitments, family expectations, and cultural repertoires of practice (Bell et al., 2009; Calabrese Barton, Tan, & Rivet, 2008; Czujko, Ivie & Stith, 2008; Hanson, 2007; Lee, C., 2007; 2008; Lee, O. & Buxton, 2008; Nasir et al., 2006).

This paper is a comparative case study focused on the experiences of mentors in a chemical oceanography afterschool program. This study explores “deep hanging”—the term mentors used to describe their experiences learning about the culture of science, scientific research processes, as well as learning they could be-and become- scientists.

Potential Significance of the Work

One approach to building broadening participation programming for youth from non-dominant groups entails bringing scientists and youth together. Broadening STEM participation is more than teaching youth about STEM practices, it also entails making the mores, expectations, values, and ways of being (Herrenkohl & Mertl, 2011) visible to youth. Delpit (1995) would refer to these hidden ways of being as codes of power. One way that designers of learning environments seek to make these codes of power visible is by leveraging experts in STEM fields to serve as mentors for youth. Mentors serve as models to make the practices of sciences visible and accessible to youth from groups traditionally underrepresented in the sciences. Scientists from non-dominant communities who have successfully navigated the societal and disciplinary barriers to participation are in a unique position to help youth understand what else it takes to be and become a scientist. The experiences of mentors from non-dominant groups have the potential to serve as resources for youth from similar backgrounds entering into the sciences. The rationale is that scientists who have personal experiences with complex scientific practices are well positioned to make ways of knowing and doing STEM visible to youth.

Unfortunately, there is a metaphorical black box around the social and cultural contexts that influence the developmental trajectories of scientists from non-dominant groups. From the point of view of novices to STEM fields, this black box obscures disciplinary practices and cultural expectations about STEM domains. For youth from non-dominant groups it can make pursuing a career in science seem impossible. From the point of view of scientists from non-dominant groups it can obscure their connections to community. Our lack of knowledge about the processes that helped mentors come to participate in STEM means that designers of learning environments are ill equipped to design learning environments that leverage all aspects of mentors’ expertise to broaden participation for youth. Understanding what is happening for mentors in spaces where youth have access to identity building experiences has implications for the design of informal learning environments. As the learning sciences field strives to design “spaces” for youth to gain expertise in the sciences, a focus on knowing & doing is essential but fails to capture all aspects of learning. Studying the ways that learning environments foster the development of the kinds of people who can do science is integral to understanding the ways that they can help youth become the kinds of people who can *be* scientists.

This paper seeks to add to the growing literature on learning as identity development and deepening participation (Azevedo, 2013; Baron, 2006; Bricker & Bell, 2014; Herrenkohl & Mertl, 2011; Lee, 2007; Nasir, 2002; Nasir & Hand, 2006) to explore natural scientist mentors' experiences learning they could be and become scientists. Further, this paper seeks to establish a link between mentor's personal experiences and the ways they then want to introduce youth to the complex practices of STEM by focusing on identity development for youth in informal settings. Finally, there are implications for the ways designers, educators and researchers prepare mentors to work in informal environments designed to broaden STEM participation for youth.

Design Principles

The Chemical Oceanography Outside the Lab (COOL) Program is an ongoing collaboration between a Parks and Recreation afterschool space, an oceanography lab, and a Learning Sciences program. COOL is a design-based research initiative to introduce young women of color to practices of the geosciences through engagement with a six month long chemical oceanography afterschool program. The COOL program echoes design principles advanced by The Institute for Science and Math Education (<http://sciencemathpartnerships.org/>).

Designers positioned youth as developing experts and sought to build bridges from youth's everyday knowledge of science and technology to discipline specific modes of inquiry and participation. Designed learning spaces brought youth and experts into collaboration to accomplish projects that had personal, community and disciplinary relevance. We believed a learning ecology built around these commitments would place youth and adults into a robust and productive learning environment likely to enable shifts in youth identification with the domains of science. The theoretical framework used in this paper explores the types of places, actions, and positions that lead to long-term changes in youths' sociomaterial practices.

Sociomaterial practices are social arrangements and material resources that together support particular lines of practice. In the COOL program, we created constellations of social arrangements that included people, materials, and activities to make complex science content. We took a peer reviewed paper about fish feminization in the Pacific Northwest (Johnson et al., 2008) and created a mapping activity that would allow youth to collaboratively create a representation of the locations and numbers of feminized fish. Then we overlaid maps that showed population density and wastewater treatment overflows. Once students participated in building this representation they came to the same conclusions that the scientists did in the discussion section of their paper. This sociomaterial arrangement of materials, people and activities allowed youth to participate in data analysis and think in ways that scientists did about their data.

Theoretical and Methodological Framework

Theoretical Framework

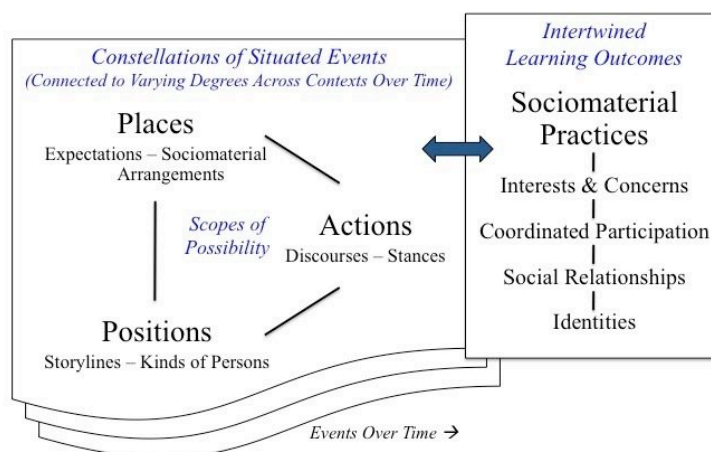


Figure 1. Cultural Learning Pathways Framework (Bell, Tzou, Bricker, & Baines, 2012).

Youth centered learning environments that leverage experts as mentors can offer youth opportunities to think like scientists, engage in authentic practices, negotiate identities, answer personally relevant questions, and learn about disciplinary specific cultural tools (Calabrese Barton, Tan, & Rivet, 2008; Cornelius & Herrenkohl, 2004; Tabak & Baumgartner, 2010; Cornelius & Herrenkohl, 2004; Polman & Miller, 2010). Harré et al. (2009)

defined positioning as a triangle of speech acts, storylines, and stances taken together these constructs shape the ways individuals develop within structures of social practice (Dreier, 2009). Positioning theory frames the ways that constellations of interpersonal interactions and social learning facilitate shifts in sociomaterial practices over time (Bell, et al., 2012). The conceptual framework for this paper focuses on the constellations of situated events on the right hand side of figure 1 below (Bell et al., 2012). As youth and mentors interact in a learning environment designed to allow for co-constructed moments of coordinated participation in chemical oceanography practices, we can observe the ways that mentors and youth access to places, positions, and discourse shape their scopes of possibility. Specifically we can observe ways that mentors' moves grant youth access or constrain opportunities for youth to participate in chemical oceanography practices.

Participants

Three professional scientists participated in the program as mentors to youth. These mentors were all women from groups traditionally underrepresented in the sciences (one German Latina, one Bolivian American, and one Diné). They were in the process of or had already attained graduate degrees in the sciences. There were fourteen female participants from sixth and eighth grades in the afterschool program (eight African-American, two White, one Filipina, one Asian-American, two Latina).

Table 1: Mentor Participants

Mentor	Ethnic Affiliations	Educational Background	Current Position
Eva	German & Latina	MS in Aquatic and Fisheries Sciences	Student and enrollment services director at local university, Director of a STEM enrichment program
Jessee	Diné, Hopi & Ute	MS in Molecular & Cellular Biology-Developmental Biology; MIT Secondary Education	Upward bound student services educator for local community college
Angelica	Bolivian American	MS in Botany; MA Museology	Research Analyst at Museum Consulting firm

The description of participants would not be complete without a description of my own positionality with respect to this project. As an Afro-Caribbean woman, I had many things in common with the largely African-American eighth grade COOL participants, as we were all members of the African Diaspora. Yet, as an Afro-Caribbean woman and an adult I was not a full community member able to leverage in-group language, references, or certain shared repertoires of practice with the youth participants. Within the COOL program, I was a full participant observer. I co-taught and co-designed the curriculum. During COOL sessions, I often took on the role of lead instructor. With the mentors in the program my role was that of pedagogical mentor and guide into the world of science education. Eva, Jessee, Angelica and I spent many hours debriefing the program sessions. We would talk about individual youth's developmental learning trajectories, flow of lessons, youth development, and pedagogical moves. Eva and Jessee have both gone onto teaching roles in the local community and I have continued my role as pedagogical mentor. This largely entails fielding questions about problems of teaching practice, approaches to teaching STEM content, and youth development work.

Outside of COOL sessions, I coordinated the program, liaised with mentors, school administrators, teachers, parents, and family members. I organized field trips, managed supplies, made calls to collaborators, and facilitated the successful completion of the COOL program from end to end. These positions gave me a unique perspective on the program activities and an intimate relationship with youth and mentors that had a positive impact on the interviews, my perspective on COOL sessions, and my analysis of program data.

Methods

The data used in this analysis were collected during the second year of the COOL program. Data sources for the analysis in this paper included field notes of mentor and student interactions, and 1.5 hours of interviews with youth and adult mentors. The entire data corpus includes over 33 hours of video, field notes for each program session, mentor and youth written reflections, and 6 hours of interviews. The interviews chosen for this analysis were the reconstructive history interviews with mentors conducted at the beginning of the COOL program in fall 2010. All interviews were transcribed and analyzed using a hybrid of emergent coding and codes built using the cultural learning pathway framework (Bell et al., 2012). This analysis surfaced a mentor generated analytical construct that will be explained in greater detail in the next section.

“Deep Hanging” became an analytical construct for understanding the ways that professional scientist mentors positioned both themselves and the COOL program participants with respect to science: as an enterprise and as a set of practices. Once a Deep Hanging profile had been built for each mentor, I returned to the video of mentor and youth interactions from the afterschool program sites and coded for ways that mentors positioned youth through stances, sociomaterial arrangements, actions. I also leveraged possible future selves (Markus & Nurius, 1986) as a way of linking mentors’ past experiences with ways that they envisioned themselves working with youth and the ways they were able to see themselves as people who could be and become scientists through their Deep Hanging experiences. We surfaced places, actions, youth, and mentor positions related to STEM practices.

Major Findings, Conclusions and Implications

Deep Hanging

“Deep Hanging” (DH) was Eva’s term. She used it to describe her learning experiences and the ways that she was brought into science and recognizing that she could be- and -become a scientist. She referred to the research cycle—i.e. the things that scientists do to complete a research project—in her description of her learning process. Eva maintained that while your advisor should facilitate this process it is actually through the interaction with peers and “just being part of it” that people learn complex scientific practices. Eva described a process of learning in practice. The question of where you can learn to fit into the culture of science was a recurring theme in interviews with other mentors and related directly to the ways they positioned themselves with respect to science. The following quote is Eva’s response to a question about how people learn complex science practices.

I think part of it is the “deep hanging” that you do. Ideally, it should be through the mentorship of your advisor...Your peers are going through the process, you’re discussing things...So being able to understand where you fit within that cycle, I think comes from just being part of it. I think a lot of it is the socialization that happens (Eva, 2011).

Eva described a socialization process that does not happen through direct instruction from a more competent other- from your advisor but it actually comes from “just being a part of it.” Learning to be and become a scientist happens in practice. As a research scientist, Eva participated in the culture of science within her lab and used science to make sense of the world.

She acknowledged that the professional definition of science that was not made accessible to everyone. “When we say science it has the Western academy picture behind it but that is not the only way of knowing or doing that helps us make sense of the world” (Eva, 2011). Eva believed that being able to see yourself as the type of person who can do science was key to the actual practice of doing science, and she is not alone; Herrenkohl & Mertl (2011) also described learning as a process that entailed knowing, doing, and being. Eva’s words challenge us to recognize that knowing and doing are not the only important aspects of becoming a scientist. DH recombines all three components- knowing through doing facilitated by being in relationship with others.

Deep Hanging as Eva described it was a process that helped change her sense of who she could be and become. This construct has four interconnected characteristics as Eva described them: Deep Hanging entails 1) authentic tasks in rich contexts, 2) direct access and engagement with novel practices, 3) leveraging interpersonal relationships to facilitate participation, and 4) interpersonal relationships that encourage shifts in identity and deepening identification with the discipline. I will use DH going forward to explore the STEM induction experiences of mentors in the 2011 COOL cohort.

Lave & Wenger’s (1991) Legitimate peripheral participation (LPP) might seem like an explanatory framework that could encompass Deep Hanging. Similar to LPP, DH entails having a purposeful activity within an ongoing complex practice. Eva was a technician in a local research laboratory and this role meant that she was engaged in purposeful activity during her interactions in the lab. However, there are a few key differences between LPP and DH. LPP has a definite telos- novices join the community of practice with the intention of becoming central participants i.e. moving ever closer to the center of practice. The LPP model when applied to mentoring in STEM presupposes that the goal of STEM education is to produce new STEM practitioners- specifically bench scientists. The “Pipeline” metaphor is an often used to describe the desired outcomes for broadening participation programs. Blickenstaff’s 2005 literature review describes the reasons that the pipeline intended to carry people from an interest in STEM through high school, college, and into a STEM career disproportionately leaks women at multiple points. Given this framing, interventions would aim to plug the “Pipeline” in such a way as to allow women to stay in the pipeline all the way into careers in STEM disciplines. This framing suggests that people who leave bench science to pursue other STEM related career paths (education, communication, or policy) have fallen out of the “leaky” pipeline. By this metric, none of the

mentors who participated in the 2011 COOL Cohort would have been qualified to mentor youth because they did not reach the center of STEM practice. I reject this framing of the COOL mentors because they were each negotiating new spaces for continued and broadened participation in STEM fields.

The mentors in the COOL program were not passive drips falling out of the STEM pipeline. They were agentive decision makers who chose to step away from the bench to bring youth into the practices of STEM by sharing their professional and disciplinary expertise. Their goals were to make visible and accessible to youth from groups traditionally underrepresented in the sciences the complex practices of science. By sharing their own journeys into science careers along with lessons learned along the way the COOL mentors shared more than the master novice relationship with youth in the program. The COOL mentors were pushing for the youth to have had enough experience with STEM practices and identities to make an informed decision about what they wanted out of their own interactions with STEM. Their stance echoes a more expansive version of STEM literacy advanced by Noah Feinstein (2009, 2010) and Joseph Polman et al., (2012) who suggest that STEM literacy needs to be functional in the lives of students as community members and individuals making decisions that have impacts on their health, communities, and career trajectories.

Both Angelica and Eva had mentors who provided direct access to novel practices. They both choose to engage in those practices with their mentors. Angelica spent time reading articles with her mentor in a context where she felt comfortable and encouraged to ask questions and develop new skills. Eva was invited to participate in lab meetings and join the full practice that was taking place in her research lab above and beyond her role as a scientist. Deep hanging also hinges upon interpersonal relationships in a nuanced way. Relationships with experts who serve as mentors helped Eva, Angelica, and Jessee deepen their participation in STEM pursuits. Jessee found her mentor in middle school. Her mentor spoke with Jessee's parents about letting her join a summer science program for girls. While their mentors provided access, it was up to Jessee, Eva, and Angelica to decide take up these practices. The other aspect of relationships translates directly into the COOL space. For Eva, Jessee and Angelica, relationships with scientists helped to shape their ideas of who they could be- and become. Eva got to know the scientist in her lab while they were working together on purposeful project; "there was that interpersonal space being brought into your professional setting. I really got to know them and to understand why they were in graduate school" (Eva, 2011).

Within Project COOL, findings from analyses of the impact of mentor Deep Hanging with youth show youths' broadening understanding of what it meant to participate in STEM, growing sense of identity as participants in the work of the chemical oceanography laboratory, and more generally an identification with the discipline. Findings on mentor learning indicate that mentors made shifts in their pedagogical and sociomaterial practices throughout the course of their involvement. In alignment with the cultural learning pathways framework (Bell et al., 2012) we saw changes in the interests and concerns of mentors, as well as the ways in which they coordinated participation in activities and deepened social relationships. This in turn allowed for mentors to reflect on how their identities as scientists, mentors, and educators changed through their participation in the hybrid learning environment.

What Deep Hanging Means for Mentors Positioning Themselves

The mentors in our program had their own complicated relationships with science. Mentors described the effects of the black box on their self-positioning with respect to science and the sense of responsibility they felt to their communities. Eva described science as a selfish proposition and one that caused culture clash "often in academic science, it's a very selfish endeavor that really goes against the grain...that's not the values we were raised with so you know there's lots of culture clash when you to get to academia" (Eva, 2010). Here Eva was discussing the ways that she found it difficult to reconcile her identities as an agent of change for her community and as a bench scientist.

COOL connected with Jessee, Eva, and Angelica, as they were each choosing to veer off of the traditional STEM career pipeline. Jessee came to our program after deciding to change her career path from a PhD in bench science to leaving with a Masters in science to join the Masters in Teaching (MIT) program at the same university. She used her hours in COOL to count towards the mandatory student contact hours necessary to enter the MIT program. Eva has just finished a Masters degree in fisheries science and was transitioning into a position in a laboratory as outreach and education manager. Angelica was finishing a Masters degree in biology and transitioning into a Masters in museology. These transitions away from exclusively pursuing bench science were connected to changing positioning for all three women (see Table 1 above).

Eva was the most directly aware of the impact of privileging the Western definition of science when working with youth from communities traditionally underrepresented in STEM.

There are people within the western academy who want to hold on to the profession, to the definition of science. I've seen that play out in a couple of different ways. What I think is so important for the girls' identity building is to see that they *can* do science, that there is an option, and that the practices they engage in are like science- in this way. When we say

science it has the western academy picture behind it but that is not the only way of knowing or doing that helps us make sense of the world (Eva, 2011).

In the above quote Eva positioned herself as someone who pays attention to privileged ways of knowing within the Western academy's definition of science but goes beyond to incorporate other ways of STEM sense making. This was a storyline for Eva, where she saw herself as a type of person who challenged what it meant to engage in science and inherently as the kind of person who facilitates a broader sense of learning for the youth she works with. Part of this came from Eva's position as a woman of color whose role as community builder was a central part of her identity, part of this stance came from her journey of learning a disciplinary specific set of scientific practices, and all of these things came into play as she took up her role in the COOL program.

Deep Hanging played a role in the lives of other mentors as well. Angelica came to the COOL program seeking to learn more about teaching science. She was finishing up her work in a biology lab with a Masters degree although she also came into the program seeking a PhD. Angelica was applying to a museology program at the same university. Angelica described a similar process of finding a mentor who helped her see herself as the kind of person who could be and become a scientist. However Angelica's learning in practice did not happen at the lab bench. Angelica foregrounded the sense of encouragement and comfort level she felt having conversations with her ecology professor about scientific papers. When she met with her mentor they would read and discuss articles together. Angelica described this time as pivotal to her development as a scientist because this helped her to feel confident as a consumer of scientific research. This comfort led her to envision a role for herself in creating scientific research.

These Deep Hanging experiences played out in the ways that mentors worked with youth in COOL. For Eva this meant that she valued the downtime, always came early to spend time with youth before the program started, and excelled at pulling in these of things into the COOL activities. Once the young women were talking about curling their hair and later in the day Eva used curlers as an analogy to describe the internal make-up of the Gas Chromatography machine. Angelica's DH experience meant that she was attuned to seeking and creating a sense of comfort paid most attention to the quietest participants and made sure that all students felt comfortable participating in games and scientific discussions. Angelica leveraged her language expertise and spoke Spanish with youth to make it more comfortable for them to participate in activities and encouraged them to share their ideas with the group.

What Deep Hanging Means for the Ways that Mentors Position Youth

The mentors all came to the COOL program at a time when they were questioning their own position with respect to STEM careers. Angelica and Jessee were particularly disillusioned as a result of negative interactions with their academic advisors- who were also running the labs they were participating in. They were in essence flirting with stepping outside of the black box. This tension made these mentors particularly aware of the ways that COOL as a broadening participation effort positioned youth. Broadening participation efforts can be built upon assimilationist metaphors, and focus on making sure that youth from non-dominant groups gain access to dominant ways of learning in practice. The COOL mentors viewed broadening participation in more expansive ways. Their vision included broadening the concept of what it meant to participate in STEM, suggesting that the ways that mentors position themselves with respect to science is crucial to understanding how they will go on to position youth.

All of the mentors in the program talked about the ways that they wanted the youth to see themselves as capable of becoming scientists. Youth from non-dominant groups may look at the black box of STEM from afar and make a decision without any knowledge of what kinds of possibilities it might hold. In essence mentors wanted to make sure that youth could make choices about participating in STEM careers with knowledge of what was inside the black box. In response to a question about the role she wanted to play Angelica said she wanted to:

Just to be the kind of person that would excite students about a subject and then make them feel that you know they could do it to and feel that they could really go on and be successful or even just think that they could be a masters student or a doctoral student. Kind of like a guide, that's how I saw myself (Angelica, 2010).

Another mentor, Jessee put it this way:

"Or even for them to realize, hey I don't really like science, I can do science. But it's not something I want to ...to be able to actually comprehend the long term interest in science would be amazing, for me to see that, and to actually see that that's actually possible" (Jessee, 2010).

Neither mentor discussed the pipeline as a goal nor STEM careers as the endpoint for youth STEM

participation. They wanted youth to leave the program with a new storyline about themselves as science participants who could go on to learn and do more science. Mentors' picture of broadening STEM participation for the youth in our program connected more deeply to the concept of "possible future selves" (Markus & Nurius, 1986).

Positioning youth as people capable of doing science and becoming scientists as opposed to future scientists within existing STEM paradigms balances the need to honor student agency while creating authentic access points to discipline-linked STEM participation. The mentors in our program were uniquely positioned to manage this tension. Here mentors' own experiences with learning scientific practices and managing tensions between assimilationist models and making their own meaning comes out. Jesse's comments highlight this tension, "hey, I don't really like science, I can do science" demonstrates this nuanced picture of what she wants for the students she mentors. Mentors in COOL pushed youth to find science personally relevant without letting go of the scientific rigor and this led to youth participants making comments like "scientists do what we did," when we asked them to describe a scientist in follow up interviews.

Implications

The above findings- specifically the role of Deep Hanging in shaping mentors' motivations can help us better understand what mentors are trying to do when they work with youth. By engaging mentors in reconstructive history interviews about their STEM induction experiences we can gain insight into the ways that they will work with young people. As a design strategy we have turned mentors into reflective practitioners with the ability to think about their own experiences of learning in practice. With programs like COOL, our design goal is to get youth into the black box, or at a vantage point to the black box for long enough to make informed decisions about their continued STEM participation. To achieve these ends, we want scientists with expansive attitudes to broadening participation and what it means to participate in STEM. Within the world of teacher education the apprenticeship of observation (Grossman, 1991) is the term used to describe the link between an individual's histories of learning which in turn prepare them to work with learners in particular ways. This study intimates that this holds true for disciplinary mentors. They learned they could be and become scientists through Deep Hanging and used this construct to position and work with youth in the afterschool program.

Warren, Ogonowski, and Pothier (2003) explained that researchers and teachers in their work with young people from non-dominant communities in the sciences, "had to work at learning to see and hear the intellectual substance" (p. 143) in their students' contributions. This reframing from "student as deficit" to placing the onus on teachers and researchers to remediate their visions of science learning and participation is a valuable change. It requires adults working with youth to make changes rather than expecting youth to do this alone. Thus creating another hybrid space where youth and disciplinary expert perceptions and participation are honored. The findings of this study go a step further to encourage the design of environments where mentors can position themselves as bridge builders not gatekeepers. Broadening participation in COOL means more than just teaching youth to do the things mentors or scientists do. It also means broadening the picture of what it means to participate in STEM.

Asking questions about disciplinary expert meaning making in programs intended to broaden participation is one way to begin this work. Future work could chart the ways that disciplinary experts are prepared to leverage youths' cultural and out of school identities into STEM learning contexts by exploring their past STEM learning experiences. Additionally, we can work to understand how experts and youth position one another with respect to STEM disciplines. Finally, we can look into the ways experts view the STEM practices of youth in relation to their own disciplinary practices.

References

- Azevedo, F. S. (2011). Lines of Practice: A Practice-Centered Theory of Interest Relationships. *Cognition and Instruction*, 29(2), 147–184.
- Barab, S. A. & Hay, K. E. (2001). Doing science at the elbows of experts: Issues related to the science apprenticeship camp. *Journal of Research in Science Teaching*, 38(1), 70–102.
- Barron, B. (2006). Interest and Self-Sustained Learning as Catalysts of Development: A Learning Ecology Perspective. *Human Development*, 49(4), 193–224.
- Bell, P., Lewenstein, B., Shouse, A., & Feder, M. (eds.) (2009). Diversity and Equity. In *Learning Science in Informal Environments: People, Places and Pursuits* (pp. 209-247). Washington, DC: National Academy Press.
- Bell, P., Tzou, C., Bricker, L., & Baines, A. M. D. (2012). Learning in Diversities of Structures of Social Practice: Accounting for How, Why and Where People Learn Science. *Human Development*, 55, 269-284.
- Blickenstaff, J. C. (2005). Women and science careers: leaky pipeline or gender filter? *Gender and Education*, 17(4), 369–386.
- Bricker, L. A., & Bell, P. (2014). "What comes to mind when you think of science? The perfumery!":

- Documenting science-related cultural learning pathways across contexts and timescales. *Journal of Research in Science Teaching*, 51, 3, 260-285.
- Calabrese Barton, A., 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.
- 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.
- Czujko, R., Ivie, R., & Stith, J. H. (2008). *Untapped Talent: The African American Presence in Physics and the Geosciences* (p. 1-22). College Park, MD.
- Delpit, L. D. (1995). *Other people's children: Cultural conflict in the classroom*. New York: New Press.
- Dreier, O. (2009). Persons in Structures of Social Practice. *Theory & Psychology*, 19(2), 193-212.
- Feinstein, N. (2009). Prepared for What? Why Teaching "Everyday Science" Makes Sense. *Phi Delta Kappan*, 762- 766.
- Feinstein, N. (2010). Salvaging science literacy. *Science Education*, 95(1), 168-185.
- Grossman, P. L. (1991). Overcoming the Apprenticeship of Observation in Teacher Education Coursework. *Teaching and Teacher Education*, 7, 4, 345-57.
- Harré, R., Moghaddam, F. M., Cairnie, T. P., Rothbart, D., & Sabat, S. R. (2009). Recent Advances in Positioning Theory. *Theory & Psychology*, 19(1), 5-31.
- Hanson, S. L. (2007). Success in Science Among Young African American Women: The Role of Minority Families. *Journal of Family Issues*, 28(3), 3-33.
- Herrenkohl, L. R., & Mertl, V. (2010). *How students come to be, know, and do : A case for a broad view of learning*. New York: Cambridge University Press.
- Hsu, P., Roth, W., & Mazumder, A. (2009). Natural Pedagogical Conversations in High School Students ' Internship. *Journal of Research in Science Teaching*, 46(5), 481-505.
- Johnson, L. L., Lomax, D. P., Myers, M. S., Olson, O. P., Sol, S. Y., O'Neill, S. M., ... Collier, T. K. (2008). Xenooestrogen exposure and effects in English sole (*Parophrys vetulus*) from Puget Sound, WA. *Aquatic Toxicology (Amsterdam, Netherlands)*, 88(1), 29-38.
- Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of scientific facts*. Beverly Hills: Sage Publications.
- Lave, J., & Wenger, E. (1991). *Situated learning : legitimate peripheral participation*. Cambridge [England]; New York: Cambridge University Press.
- Lee, C. D. (2007). *Culture, literacy & learning: Taking bloom in the midst of the whirlwind*. New York, NY: Teachers College Press.
- Lee, O., & Buxton, C. (2008). Science Curriculum and Student Diversity : A Framework for Equitable Learning Opportunities. *The Elementary School Journal*, 109(2), 123-137.
- Markus, H., & Nurius, P. (1986). Possible Selves. *American Psychologist*, 41(9), 954-969.
- Nasir, N.S. (2002). Identity , Goals , and Learning : Mathematics in Cultural Practice. *Mathematical Thinking and Learning*, 4(2), 213- 247.
- Nasir, N. S., & Hand, V. M. (2006). Exploring Sociocultural Perspectives on Race , Culture , and Learning. *Review of Educational Research*, 76(4), 449-475. Retrieved from <http://www.jstor.org/stable/412441>
- Nasir, N.S., Rosebery, A. S., Warren, B., & Lee, C. D. (2006). Learning as a Cultural Process: Achieving Equity through Diversity. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (1ST ed., pp. 489-504). Cambridge, Massachusetts: Cambridge University Press.
- Polman, J., & 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.
- Polman, J., Newman, A., Farrar, C., & Saul, E. W. (2012). Science Journalism. *Science Teacher*, 79, 1, 44-47.
- Tabak, I., & Baumgartner, E. (2010). The Teacher as Partner : Exploring Participant Structures , Symmetry , and Identity Work in Scaffolding. *Cognition and Instruction*, 22(4), 393-429.
- Warren, B., Ogonowski, M., & Pothier, S. (2003). "Everyday" and "Scientific": Rethinking Dichotomies in modes of thinking in science learning. In A. Nemirovsky, A. S. Rosebery, J. Solomon, & B. Warren (Eds.), *Everyday matters in mathematics and science education* (pp. 119-152). Mahwah, N.J.: Erlbaum.

Acknowledgments

This work is supported by a grant from the National Science Foundation (NSF#9014483) and is partly supported by the Learning in Informal and Formal Environments (LIFE) Science of Learning Center (<http://life-slc.org/>) (NSF# 0354453). All opinions, findings, and conclusions expressed here are those of the author. I would also like to thank my advisor Dr. Philip Bell, my research team- ESTG, and Shelley Stromholt my research collaborator.

Not a Magic Bullet: The Effect of Scaffolding on Knowledge and Attitudes in Online Simulations

Ido Roll, Adriana Briseno, Nikki Yee, Ashley Welsh, University of British Columbia,
220-1961 East Mall, Vancouver, BC, V6T 1Z1, Canada
{ido.roll; adriana.briseno}@ubc.ca, nikki.yee@alumni.ubc.ca, welshash@interchange.ubc.ca

Abstract: Common wisdom and prior research suggest that students with low prior knowledge are in greater need for scaffolding. However, some forms of scaffolding may overload novice-students' cognitive capacity or short-circuit productive exploration of the problem space. Hence, we evaluate the effectiveness of scaffolding in a virtual simulation in physics, considering students' attitudes and prior knowledge. 100 undergraduate students completed either a scaffolded or a relatively unstructured activity, followed by another unstructured activity. While the given scaffolding was beneficial for students with high prior knowledge, it did not assist students with low prior knowledge. Furthermore, the scaffolded activity increased students' attitudes towards memory-and-recall in a way that transferred to the later unsupported activity, where these goals were no longer appropriate. Last, prior knowledge did not contribute to learning outcomes in the presence of intuitive grounded feedback. We explain these results in terms of productive failure and cognitive load.

Introduction

The desired level of assistance in online exploratory learning environments has been the focus of an intensive debate. In general terms, proponents of high level of guidance suggest that novice learners learn better in situations where direct and comprehensive instructional guidance is provided (e.g., Kirschner, Sweller, & Clark, 2006). However, others assert that students should be given more agency over their learning processes within the supported environment (Hmelo-Silver, Duncan, & Chinn, 2007). It is increasingly clear that rather than asking how much support, a more viable question is what timing and forms of support are most appropriate (Koedinger & Aleven, 2007; Wise & O'Neill, 2009).

Many studies of inquiry learning have evaluated the effect of support on learning (e.g., de Jong & van Joolingen, 1998; Manlove, Lazonder, & de Jong, 2007) and motivation (e.g., Butler, 1998; Horner & Gaither, 2006). However, the integration of student characteristics such as incoming attitudes, prior knowledge, and expectations, in the context of inquiry learning, has not yet received much attention. This study adds to the assistance debate by investigating the relationship between level of scaffolding and student prior knowledge and motivation, in the context of an online physics simulation.

The Interaction between Prior Knowledge and Support

Students' prior knowledge and experiences have been long recognized as a critical factors in their learning processes and outcomes. Tuovinen and Sweller (1999) suggest that novice learners can benefit from higher levels of guidance, compared with more experienced learners, because guidance acts as a substitute for missing schemas. Thus, the scaffolding helps novices by minimizing working memory load. Experienced learners, on the other hand, are less likely to need additional instructional guidance as they already possess schemas that support the construction of mental representations. Kalyuga (2007) refers to this phenomenon as the "expertise reversal effect", where guidance becomes redundant and thus, ineffective for experienced learners.

Somewhat contrary to these results, a second strand of studies indicates that novice learners may find difficulties in interpreting the provided assistance. Instead, novice learners may benefit more from engaging in solution attempts, albeit failed ones, before they can comprehend and make use of the provided guidance. For example, we have previously found that novices may learn from their own explorations more than from using hints, even though the hints led them to quicker solutions (Roll, Baker, Aleven, & Koedinger, 2014). In another example, Vollmeyer, Burns, and Holyoak (1996) found that learners benefit more from answering their own questions, compared to being given scaffolding in the form of specific sub-goals. These lines of work highlight the distinction between *successful engagement* (that is, completing the task without errors), and *productive engagement* (that is, performing better on delayed measures of learning; Kapur, 2008; Schwartz & Martin, 2004; Roll, Aleven, & Koedinger, 2009). Letting novices explore the problem space also helps them acquire more flexible knowledge (Rittle-Johnson, Star, & Durkin, 2012). It seems that scaffolding the learning process by providing students with seemingly useful sub goals may put students in an "answer hunting" mode which may be counterproductive (Miller, Lehman, & Koedinger, 1999; Sweller, Mawer, & Ward, 1988).

While the debate around scaffolding is framed mainly in terms of given and withheld information (Koedinger & Aleven, 2007), students may be given implicit scaffolding in the form of grounded feedback (Nathan, 1998). Environments that offer grounded feedback support learning by showing the outcomes of

students' actions in an alternative, familiar, representation. One family of simulations that supports learning using grounded feedback is PhET Simulations (Wieman, Adams, & Perkins, 2008). Overall, PhET simulations are used over 45,000,000 times a year. Figure 1 shows the D/C Circuit Construction Kit, one of the more popular simulations of the PhET family (<http://phet.colorado.edu/en/simulation/circuit-construction-kit-dc>). In this specific simulation, students explore basic properties of D/C circuits by connecting wires, light bulbs, resistors, switches, and measurement instruments, on a virtual test bed. Grounded feedback in the simulation is given, for example, by matching the light intensity of the light bulbs to their current and voltage. In other cases, when the power is too high, elements in the circuit may catch on fire. In addition, students can see the speed in which electrons "travel" through the wires. This feedback gives students intuitive understanding on the outcomes of their actions. Other features of the simulation require more domain-level expertise, such as measuring using the voltmeter, shown in the bottom-left corner of all screenshots in Figure 1.



Figure 1. The D/C Circuit Construction Kit simulation. Grounded feedback is given in the form of light intensity, electron speed, and fire.

The Interaction between Attitudes and Support

The role of scaffolding in learning may be explored in further detail through an examination of motivation as measured through success attributions and self-efficacy. Success attributions are factors that students recognize as being the determinants of success in a particular context (Graham & Williams, 2009). They are important in helping learners understand what is required from them in order to succeed. Thus, success attribution should match the learning situation. For example, while students in a highly scaffolded environment can learn by answering questions, this strategy may be less appropriate for a less scaffolded environment, where students should determine their own sub-goals. Self-efficacy describes an individual's perceived capabilities and competencies for learning or acting in context (Bandura, 1997; Schunk & Pajares, 2009). Research suggests that students' perception of their ability is often inaccurate. In particular, novices often have unrealistic expectations for their own successes when confronted with novel tasks, as they are unaware of their personal knowledge gaps (Kruger & Dunning, 1999; Roll, Aleven, & Koedinger, 2011). Students tend to be more productive if they feel that success can be achieved, they think that they know what they need to do to be successful, and they possess the skills to do it (Butler & Cartier, 2004). Success attributions and self-efficacy research indicate a connection between these motivational attitudes and increased achievement (Bandura, 1997; Wilson & Linville, 1982). Understanding students' waxing and waning success attributions and self-efficacy can give us key insights into the role of scaffolding in the learning process and help us structure scholastic activities more strategically.

Similar to prior knowledge, students' attitudes coming into the learning environment may show a differentiated effect for different forms and levels of scaffolding. For example, Belenky and Nokes-Malackh (2013) found that students who enter a learning situation with low levels of mastery goals are more likely to benefit from low levels of scaffolding (in the form of explicit instruction), compared with students with high levels of mastery goals.

Research Questions

The current study explores how learning gains and attitudes, and in particular, students' expectations, change over time and in relationship to learners' prior knowledge and level of assistance provided. Specifically, we are interested in investigating the following questions:

- (i) What is the effect of scaffolding on learning, contingent on prior knowledge?
- (ii) How do students' expectations change with time, based on given scaffolding, and contingent on prior knowledge?

Method

One hundred post-secondary students from first-year physics classes in a large Canadian university participated in the study on the topic of D/C circuits. Data collection occurred in two waves: once during the school year as an add-on to the physics class, and again in the summer as a more integrated part of the physics course content.

The study procedure had five steps (see Figure 2): (i) Students first completed a 5-minute pre-test (assessing content knowledge) and pre-survey (assessing attitudes). (ii) Students then completed a 25-minute activity on light bulbs using the online simulation. Students were randomly assigned to one of two conditions with high- or low-levels of scaffolding, as described below. (iii) Following the learning activity, students were given a short break, and then completed a mid-survey which was identical to the pre-survey. (iv) A transfer inquiry activity on the topic of resistors was then administered for 25 minutes. Students from both groups received only low level of scaffolding in this activity. The activity used the same simulation that was used in the light bulb activity. (v) Last, a post-test of learning outcomes from both activities and a third survey of attitudes were administered.

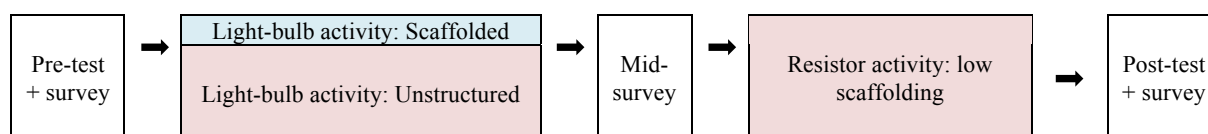
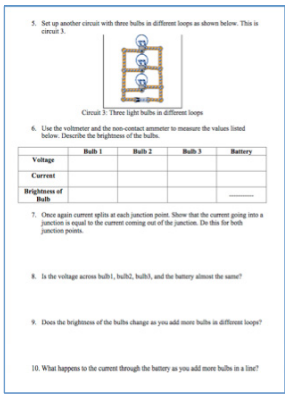


Figure 2. The study procedure.

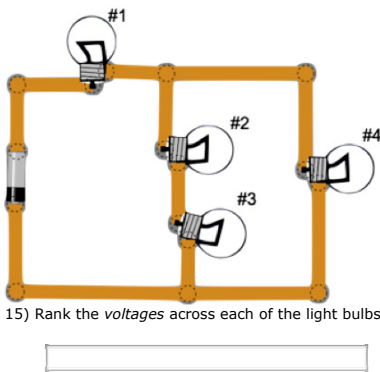
Materials

All students used the D/C Circuit Construction Kit (shown in Figure 1) throughout the study. For the first activity, students in both conditions received the same learning goal and some advice for their exploration: “Use the DC Circuit PhET simulation to explore how voltage, current, and the brightness of light bulbs depend on the number of light bulbs in a circuit and the arrangement of light bulbs in a circuit. For example, what happens when several light bulbs are connected in a line? What happens when light bulbs are sitting on different loops in the same circuit, and when electrons are moving through different loops? What happens when you use several batteries and switches?” Half of the students were assigned to the Scaffolded condition. In addition to the elaborated learning goal, students in this condition received a worksheet that guided their exploration of these topics. The activity was adapted from recommended activities by the designers of the simulation, using inputs from the instructor of the course in which data was collected. The activity included diagrams that showed students which circuits to build, tables for students to fill in their measurements, and prompts that asked students to compare and contrast their measurements for the different experimental set-ups (see Figure 3a). The rest of the students received only the elaborated learning goals, without the additional worksheet. Even though students in this condition were supported by the learning goal and its recommendations, for simplicity, we refer to this condition as Unstructured.

During the second activity, on the topic of resistors, all students received an (Unstructured) learning goal with no worksheet. Notably, this activity was much harder, as the topic of resistors is less intuitive than light bulbs. Also, the grounded feedback was less relevant in this activity: when working with light bulbs, light intensity offers useful information. However, when working with resistors, students should use the measurement instruments (voltmeter, ammeter) and interpret numeric values.



(a)



(b)

Figure 3. An example from the worksheet of the Scaffolded group (a) and an example of a post-test item (b).

The post-test covered both topics (light bulbs and resistors) and was administered online. All items in the test were conceptual and required no calculations (see Figure 3b). The post-test was a reliable measure of students' knowledge, with Cronbach $\alpha = 0.75$. The pre-test included a subset of the post-test items. Only items that had no diagrams in the post-test appeared also in the pre-test, in order to prevent a bias in students' inquiry towards regenerating pre-test questions. Three students had a perfect score on the pre-test and thus were removed from the analysis.

The attitudinal survey of success attributions and efficacy was adapted from Butler and Cartier (2004), and included ten Likert-scale items with two stems: 'I think that I can do a good job of...' and 'I think that I will succeed if I...' (see examples below). The same survey items were used before the first activity (pre-survey), between activities (mid-survey), and after the second activity (post-survey).

Analysis

Students were split to high- and low-prior knowledge groups, based on pre-test performance. The effects of condition and prior knowledge were evaluated using a MANOVA with light bulb and resistor scores from the post-test as the dependant variables and condition and prior knowledge (median split) as factors. Since there were two waves of data collection, and there were slight differences in the light bulb post-test items between the two waves, we use normalized z-scores throughout the analysis.

To analyze the attitudinal surveys, a factor analysis was run on each of the three time periods (pre-, mid-, and post-). The factors obtained in the post-survey were used to indicate which items can be averaged together to create summary scores for each student*time*factor. The factors from the final time period were used because they represent the final structure of the students' attitudes, and they are highly interpretable. This method allows us to see how students' answers to the items across the final factors changed over time, and to explore what variables might be influencing that change. Post-survey results indicated three, clear factors: The first factor, success expectations, includes items about students' expectations for success, such as '*Before I begin a PhET sim activity, I think that I will succeed if... I try hard*'. This factor included six items. The second factor, scientific reasoning, with two items, includes '*When I work on a PhET sim like the one in the example I think that I can do a good job of ...testing my ideas and theories*' and '*...exploring the topic*'. The last factor, memory and recall, also consists of two items: '*I think that I can do a good job of ...memorizing information about circuits*' and '*...answering given questions*'. Student scores on the items within each factor were averaged to give a summary score for each student on each of the three factors on the pre-, mid-, and post-survey. The summary scores for each of the time periods were then used as the repeated measures in a mixed design MANOVA, with prior knowledge and condition entered as independent variables, each with two levels.

An alpha level of 0.05 is used throughout the analysis.

Results

No significant differences were found between the Scaffolded and Unstructured conditions on the pre-test, $t_{(95)} = -.12$, $p = .91$. Identical items between pre- and post-test show significant learning. Mean (SD): Pre: 0.47 (0.17); post: 0.62 (0.23); $t_{(96)} = 6.1$, $p < 0.0005$. Analysis of learning by prior knowledge shows that students with low prior knowledge learned more than students with high prior knowledge, though learning for both groups was significant: Low prior knowledge: from 0.33 (0.78) to 0.57 (0.22), $t_{(44)} = 6.5$, $p < 0.0005$; High prior knowledge: from 0.59 (0.13) to 0.67 (0.24), $t_{(51)} = 0.015$.

Learning as a Function of Prior Knowledge and Condition

A MANOVA with normalized performance on the light bulb and resistor tests, as a function of condition and prior knowledge, found a marginally significant condition*prior knowledge interaction, motivating further analysis of the two separate ANOVAs: Wilks' $\lambda = .952$, $F_{(2, 92)} = 2.3$, $p = .10$. Descriptive statistics of students' z-scores on both post-test sections are listed in Table 1.

Table 1: Normalized mean (SD) of learning outcomes by condition

(a) Light bulb post-test

	Unstructured	Scaffolded
Low Prior Knowledge	-.09 (1.11)	-.53 (.76)
High Prior Knowledge ^a	.07 (1.01)	.41 (.87) ^b

(b) Resistor post-test (notably, support levels refer to the previous activity, as all students worked with Unstructured activity in this activity.)

	Unstructured	Scaffolded
Low Prior Knowledge	-.33 (1)	-.30 (.74)
High Prior Knowledge ^a	.20 (.92)	.33 (1.12)

^a Main effect for prior knowledge, $p < 0.01$; ^b effect for prior knowledge within Scaffolded condition, $p < 0.05$.

The ANOVA for the light bulb post-test showed a significant interaction effect for prior knowledge*condition, $F_{(1,93)} = 4.0$, $p < .05$, $\eta^2_p = .041$. Planned contrasts indicated that within the Scaffolded condition, students with high prior knowledge significantly outperformed students with low prior knowledge: $t_{(93)} = 3.42$, $p = .001$. The same difference was not found within the Unstructured condition: $t_{(93)} = -.59$, $p = .56$. The ANOVA for the resistor post-test found a significant main effect for prior knowledge, $F_{(1,93)} = 8.71$, $p < .05$, $\eta^2_p = .086$.

Attitudes as a Function of Prior Knowledge and Condition

A repeated measures MANOVA with the three factors as dependent measures found that time*condition*prior knowledge was significant, supporting a further exploration with the independent repeated measures ANOVAs. Wilks' $\lambda = .850$, $F_{(6, 88)} = 2.58$, $p = .024$. Table 2 shows the results for the three attitudinal factors.

Table 2: M (SD) for factor summary-scores by prior knowledge and condition for pre-, mid-, and post-survey

(a) Memory and Recall			(b) Scientific Reasoning			(c) Success Expectations		
	Unstructured	Scaffolded ^b		Unstructured	Scaffolded		Unstructured	Scaffolded
Low	2.63(.63)	2.67(.62)	Low	2.82(.55)	2.79(.64)	Low	3.02(.46)	3.02(.51)
Prior	2.73(.61)	2.98(.56)	Prior	3.06(.62)	3.11(.76)	Prior	3.03(.47)	3.10(.70)
	2.42(.46)	2.79(.51)		2.92(.76)	2.97(.76)		2.98(.50)	2.88(.70)
High	2.71(.62)	2.41(.49)	High	2.90(.55)	2.79(.49)	High	2.95(.55)	2.79(.55)
Prior ^a	2.79(.64)	2.77(.44)	Prior	3.25(.56)	2.95(.49)	Prior ^a	3.11(.62)	2.97(.49)
	2.85(.65)	2.79(.53)		3.15(.46)	3.14(.49)		3.08(.55)	2.98(.50)

^a A significant time * prior-knowledge interaction, $p < 0.05$; ^b a significant time * condition interaction, $p < 0.05$

With regard to *memory and recall*, a repeated-measures analysis revealed significant interaction for time*condition: $F_{(2,186)} = 3.20$, $p = .038$, $\eta^2_p = .033$, and for time*prior-knowledge: $F_{(2,186)} = 4.19$, $p = .032$, $\eta^2_p = .043$. As seen in Figure 4a, students in the Scaffolded condition associated success with memory and recall during the mid-survey, that is, right after completing the activity that required them to answer questions. Interestingly, their increased attitudes towards memory and recall remained high also after the second activity, which was Unstructured for all students.

With regard to *success expectations*, the repeated measures ANOVA found a significant time*prior knowledge interaction, $F_{(1.79,166)} = 4.94$, $p = .022$, $\eta^2_p = .05$. As seen in Figures 4b,c, students with low prior knowledge began the study much more certain of their ability. However, their confidence dropped after the resistor activity. Students with high prior knowledge showed the exact opposite trend, gaining confidence from pre- to post-survey.

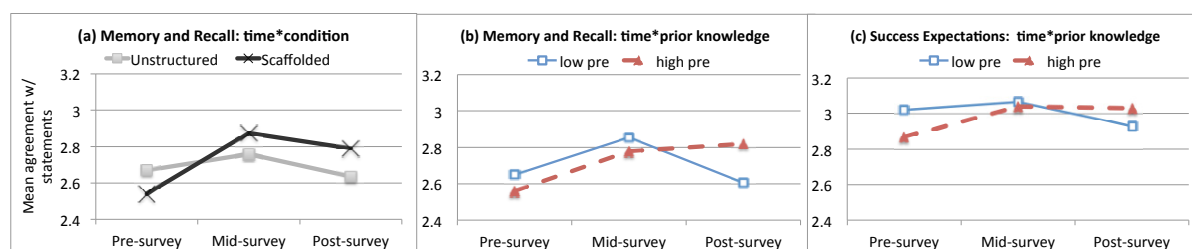


Figure 4. Significant interactions in the attitudinal data

No significant interactions were found with regard to the *scientific reasoning* factor. However, there was a significant main effect for time, $F_{(2,186)} = 13.64$, $p < .05$, $\eta^2_p = .13$.

Discussion

Our first research question focuses on the relationship between learning, prior knowledge, and scaffolding. In addition to finding a main effect for prior knowledge on learning outcomes, results show a somewhat surprising interaction of prior knowledge with condition. In the light-bulb section of the post-test, prior knowledge was very helpful for students who worked with the Scaffolded activity, but had no effect on performance among students working with the Unstructured activity. In fact, for students with low prior knowledge, performance on the Scaffolded condition was *lower* than the Unstructured condition. This result raises two questions: Why did the scaffolding help only students with high prior knowledge? And why did prior knowledge help students only in the Scaffolded condition?

With regard to the first question, only students with high prior knowledge benefited from the given scaffolding, suggesting that they were better able to infer relationships and generalize scientific rules based on the questions and diagrams they were given. Rather than limiting their inquiry, the scaffolding was used by high-prior students to engage in inquiry with better data. Novice learners, on the other hand, may have been cognitively overloaded by answering questions or filling-in tables so that they did not engage in the higher-order thinking that the worksheet had charted for them. In this case, it seems that the high scaffolding gave students with low prior knowledge a false sense of progress, as they could find values without understanding the patterns. This result highlights the tension between *successful behaviours* (i.e., completing the task), and *productive behaviours* (i.e., causing learning). While the high-level of scaffolding allowed students with low

prior knowledge to engage in more sophisticated experimentation, it did not support learning. We previously found similar results in learning in a problem solving environment, where hints helped students with low prior knowledge to complete the given problems, but did not improve their chances of succeeding on future problems (Roll et al., 2014). Another clue for the negative (yet insignificant) effect of scaffolding on novice students in the light-bulb activity may be found in their use of intuitive knowledge. As mentioned above, students could have made significant use of grounded feedback (light intensity, electron movement, etc). However, the scaffolding required students to use more formal physics and measurement instruments. Thus, pushing students towards formal notation and practices may pull them away from their intuitive ideas, and subsequently, hurts their learning. This is similar to the effect of equations vs. story problems among novice learners (Koedinger & Nathan, 2004).

Interestingly, these results are at odds with the expertise reversal effect (Kalyuga, 2007). While we found that students with high prior knowledge benefit more from high levels of scaffolding, the expertise reversal effect shows the opposite trend. One explanation for this difference is the given support. Our support gave students *more* responsibilities in terms of collecting data and running virtual experiments. Thus, the scaffolding that we used may have led to an increase in students' cognitive load. Kalyuga (2007), on the other hand, used worked examples. Thus, that scaffolding reduced students' cognitive load. Both forms of scaffolding can be useful – but apparently, for different populations.

The second finding shows that prior knowledge did not improve learning on the light bulb post-test for those in the Unstructured condition. It is possible that the prior knowledge of the high-prior students was too narrow and did not prepare them to conduct productive inquiry during this activity. While prior knowledge did not play a role in the light-bulb activity, it was a significant determinant of learning in the second activity. Given that both tasks were isomorphic for the Unstructured condition, why did prior knowledge improve learning in the resistor activity, but not in the light bulb activity? Notably, the resistor activity requires much better domain knowledge. Thus, the effect of prior knowledge on learning is not surprising. In the light bulb activity, grounded feedback may have helped low-prior students to overcome their lack of formal disciplinary knowledge.

Effect of Condition and Prior Knowledge on Attitudes

Our second research question focuses on the relationship between scaffolding, knowledge, and attitudes. The repeated-measures MANOVA found that all three attitudinal factors changed over time, supporting the notion that these motivational components are sensitive to contextual features. Interestingly, changes in attitudes generally tended to peak at the mid-survey, suggesting that students felt they knew what they needed to do to be successful based on their first activity. Perhaps this confidence was fostered by the grounded feedback from the PhET simulation on this activity, as the light bulb activity is much more intuitive than the measurement-intensive resistor activity.

Divergent patterns of success expectations based on prior knowledge can be noted from figures 4b-c. In the beginning, students with low prior knowledge tended to feel that they could navigate a path to success, whereas students with higher prior knowledge were not so confident. This difference in starting points may be due to a greater familiarity with the content or task demands by more expert students, while students with low prior knowledge may have suffered the dual burden of being unskilled and unaware (Kruger & Dunning, 1999; Roll et al., 2011). Students in both groups tended to peak at the mid-survey, and then switch trends: students with high prior knowledge maintained confidence, whereas more novice students saw a sharp decrease. This interaction may be interpreted as a realignment of success expectations where novice students have now gained the experience to adjust success expectations based on a more realistic understanding of the task and their abilities within the task. Although initial success expectations were not realistic, judging by performance on pre-test, students' expectations at the end of the study seem to reflect their actual knowledge.

Interestingly, students with high prior knowledge within the Unstructured condition did not perform as well on the light bulbs post-test, but did not seem to suffer much in terms of success expectations. Although this pattern may be explained by the pervasive influence of prior knowledge, it also seems to support the notion that students with high prior knowledge in the Unstructured condition had their content knowledge affirmed through the online simulation. Rather than pushing themselves to ask more advanced questions, these students appeared to be satisfied with their success and continued prospects for success, without actually making advancements in their knowledge about light bulbs.

Interactions were also noted in the memory and recall factor, where students predicted that they could be successful at memorizing information and answering questions. Figures 4a-b indicate significant interactions of time with condition and with prior knowledge. It is interesting to note that low ability students reported that they could be successful at memorizing information and answering questions, even though they did not seem to acquire much content knowledge by answering questions in the Scaffolded condition. Without basic conceptual understanding, merely filling in questions or memorizing facts is unlikely to support successful outcomes. This finding seems to support earlier suggestions that novice students, especially within the Scaffolded condition,

may be more pre-occupied with answering questions than focusing on deeper relationships. In addition, an inquiry activity requires a different kind of thinking than that encouraged by focusing on memory and recall.

Last, the interaction between time and condition is especially interesting, as it transfers to the post-survey, when all students worked with the same activity. Simply put, the level of support on the first activity determined students' attitudes following both the first and the second activity. We can split graph 3a to two: First, in the mid-survey, students who worked with the Scaffolded activity were confident in their ability to remember facts and answer questions. This makes sense, as the activity asked students to do exactly that. The second component of the graph is more surprising. Scaffolded students trusted their ability to remember and answer also after the resistor activity, and even though that activity was Unstructured for all students. This result demonstrates the contextual nature of students' attitudes, and how attitudes are shaped by experiences and expectations. At the same time, it also shows some maladaptive behaviour, as Scaffolded students did not readjust their expectations to the unstructured resistor activity.

The study presented above has several limitations. First, the sample size is fairly small. Second, analysis of the attitudinal surveys showed different factors for the pre-, mid-, and post-survey. While we chose only to use the post-survey distribution to factors, understanding the shift in the factor analysis is of interest. Last, the study relies on self reports and test scores, and lacks process measures. Examining students' online behaviours and paper materials, and inferring their strategies from these data, would improve our understanding of the effect of scaffolding, prior knowledge, and attitudes on learning (Kardan, Roll, & Conati, 2014).

Summary and Implications

We describe a study with 100 post-secondary students using an online simulation for D/C circuits. During the first activity, half of the students received a high level of scaffolding (using an elaborate worksheet), while the other half received only learning goals. During the second activity, which was much harder, all students received only learning goals. Results show that the level of support in the first activity had an effect on students' attitudes in the second activity. Thus, it is important for instructors to realize the contextual nature of success attributions and efficacy, which can work in a cyclical manner to influence learning outcomes (Butler & Cartier, 2004).

Secondly, results show that while students who received a high level of support benefited from their prior knowledge, prior knowledge was not helpful in the Unstructured condition. Yet, this was the case only for the first activity. Learning in the second, and more challenging, activity, benefited significantly from prior knowledge. We hypothesize that students' intuition and grounded feedback play a role. In the light-bulb activity, students could engage in productive inquiry using situational cues (such as light intensity and electron speed). In the resistor activity, learning required the use of virtual measurement instruments, which was much harder for students with low prior knowledge.

Results also show that while the given scaffolding helped students with high prior knowledge, it was not effective for students with low prior knowledge. It seems that the scaffolding that we used, and which focused on giving students sub-goals and reflection questions, required more technical proficiencies from students, thus, increasing their cognitive load. Also, it may have focused students on values and formalities rather than intuitive ideas. As a reminder, this support was modeled after common classroom activities. Thus, intuitions about the utility of different support mechanisms may not be confirmed by data (Koedinger & Nathan, 2004). Rather than debating high versus low levels of support, it may be beneficial to identify which support, for what tasks, is useful for which learners.

References

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Belenky, D. M., & Nokes-Malach, T. J. (2013). Mastery-approach goals and knowledge transfer: An investigation into the effects of task structure and framing instructions. *Learning and Individual Differences*, 25, 21-34.
- Butler, D. L. (1998). The Strategic Content Learning approach to promoting self-regulated learning: A report of three studies. *Journal of Educational Psychology*, 90(4), 682-97.
- Butler, D. L., & Cartier, S. C. (2004). Promoting effective task interpretation as an important work habit: A key to successful teaching and learning. *Teachers College Record*, 106, 1729-1758.
- de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68, 179-201.
- Graham, S. & Williams, C. (2009). An attributional approach to motivation in school. In K. Wentzel and A. Wigfield (Eds.), *Handbook of Motivation at School* (pp. 11-33) New York: Routledge.
- Hmelo-Silver, C. E., Golan Duncan, R., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99-107.

- Horner, S., & Gaither, S. (2006). Attribution retraining with a second grade class. *Early Childhood Education Journal*, 31, 165-170.
- Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19(4), 509-539.
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379-424.
- Kardan, Roll, & Conati (2014). The usefulness of log based clustering in a complex simulation environment. In *Proceedings of the International Conference on Intelligent Tutoring Systems (ITS)*.
- 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.
- Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, 19(3), 239-264.
- Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *The Journal of the Learning Sciences*, 13(2), 129-164.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134.
- Manlove, S., Lazonder, A. W., & de Jong, T. (2007). Software scaffolds to promote regulation during scientific inquiry learning. *Metacognition and Learning*, 2, 141-155.
- Miller, C. S., Lehman, J. F., & Koedinger, K. R. (1999). Goals and learning in microworlds. *Cognitive Science: A Multidisciplinary Journal*, 23(3), 305-336.
- Nathan, M. J. (1998). Knowledge and situational feedback in a learning environment for algebra story problem solving. *Interactive Learning Environments*, 5(1), 135-159.
- Rittle-Johnson, B., Star, J. R., & Durkin, K. (2012). Developing procedural flexibility: Are novices prepared to learn from comparing procedures? *British Journal of Educational Psychology*, 82(3), 436-55.
- Roll, I., Aleven, V., & Koedinger, K. R. (2009). Helping students know 'further' - increasing the flexibility of students' knowledge using symbolic invention tasks. In N. A. Taatgen, & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. (pp. 1169-74). Austin, TX: Cognitive Science Society.
- Roll, I., Aleven, V., & Koedinger, K. R. (2011). Metacognitive practice makes perfect: Improving students' self-assessment skills with an intelligent tutoring system. In G. Biswas (Ed.), *Proceedings of the International conference on artificial intelligence in education* (pp. 288-295). Berlin: Springer Verlag.
- Roll, I., Baker, R.S.J.d., Aleven, V., & Koedinger, K. R. (2014). On the benefits of seeking (and avoiding) help in online problem solving environment. *Journal of the Learning Sciences*, DOI: 10.1080/10508406.2014.883977
- Schunk, D. & Pajares, F. (2009). Self-efficacy theory. In K. Wentzel and A. Wigfield (Eds.), *Handbook of Motivation at School*. (pp. 35-53) New York: Routledge.
- Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, 22(2), 129-184.
- Sweller, J., Mawer, R. F., & Ward, M. R. (1988). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General*, 112(4), 639-661.
- Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, 91(2), 334-341. doi:10.1037/0022-0663.91.2.334
- Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science: A Multidisciplinary Journal*, 20(1), 75-100. doi:10.1207/s15516709cog2001_3
- Wieman, C. E., Adams, W. K., & Perkins, K. K. (2008). PhET: Simulations that enhance learning. *Science*, 322(5902), 682-683.
- Wilson, T., & Linville, P. (1982). Improving academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology*, 42, 367-376.
- Wise, A. F. & O'Neill, D. K. (2009). Beyond "more" versus "less": A reframing of the debate on instructional guidance. In S. Tobias, & T. M. Duffy (Eds.) *Constructivist instruction: Success or failure?* (pp. 82-105) New York: Routledge.

Acknowledgements

The work described in this paper is supported by the Social Sciences and Humanities Research Council of Canada (SSHRC) grant #430-2012-0521 and by the Betty and Gordon Moore Foundation. We would like to thank Kathy Perkins, John Blanco, and the PhET project team for their assistance

Explanations that Make Sense: Accounting for Students' Internal Evaluations of Explanations

Shulamit Kapon, Faculty of Education, University of Haifa, 199 Aba Khoushy Ave., Mount Carmel, Haifa 3498838, Israel, shulamit.kapon@edu.haifa.ac.il

Orit Parnafes, Education Department, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 8410501, Israel, Oritpa@bgu.ac.il

Abstract: This paper discusses students' internal evaluation of explanations, as opposed to external evaluations of explanations that are based on compatibility with formal scientific explanations. We propose a theoretical construct, metaphorically termed *the internal metric for certainty in explanation*, to account for the process of internal evaluation. We suggest that the internal metric is multidimensional, and discuss three dimensions: (1) intuition; (2) local coherence; (3) mechanistic reasoning. We argue that these three dimensions are highly active in self-evaluating explanations about phenomena in the physical world. We operationalize these dimensions into empirically traceable terms. We illustrate our framework through an analysis of an episode of a student's reasoning about what causes a plastic bottle to shrink when air is pumped out of it. The analysis demonstrates that the framework can explain conviction in an explanation, as well the preference for one explanation over another.

Introduction

One of the most common words in young children's vocabulary is "why." This word, indicating curiosity, points to children's essential need to understand and explain the world around them, a need that does not disappear as they grow older. From a very young age, humans try to explain the phenomena they encounter (Gopnik, 1998) for sense making purposes (Berk, 1985; Chi, Bassok, Lewis, Reimann, & Glaser, 1989) and for social purposes (Dagher & Cossman, 1992). Moreover, humans frequently evaluate various explanations of the same phenomenon, choosing the one that seems better to them and rejecting the one that seems worse (Harman, 1965; Kapon & diSessa, 2012; Keil, 2006). The assumption that guides this work is that the ability to self-evaluate explanations according to internal and not only external measures, is important for sense making and in the progression from one explanation to another. Although students' explanations are commonly evaluated by means of their compatibility to formal socially-constructed scientific explanations, students' internal evaluations of explanations may have different dimensions, and may rely on different resources. This study aims to explore the dimensions by which students evaluate their own explanations and those of others.

We suggest a theoretical construct that is aimed to account for the process of internal evaluation, which we metaphorically term *the internal metric for certainty in explanation*. Our goals in this paper are two-fold. First, we argue based on the literature that humans indeed possess such internal metric, and that this metric is highly contextually dependent. Second, we aim to operationalize this elusive metaphorical term into empirically traceable terms for the special case of explanations of phenomena in the physical world. We illustrate our framework through an analysis of an episode of a student's reasoning about what causes a plastic bottle to shrink when air is pumped out of it.

What Counts as a Good Explanation: Extrinsic and Intrinsic Considerations

The word metric implies standard measurement. Philosophers of science have long been debating the nature, characteristic features and developmental patterns of formal scientific explanations (e.g., Bechtel & Abrahamsen, 2005; Friedman, 1974; Machamer, Darden, & Craver, 2000; Salmon, 1989; Toulmin, 1972). Indeed, the most commonly used metric in educational practice to evaluate an explanation of a phenomenon in the physical world is the degree of its adherence to the formal scientific explanation, which comes to bear for instance as regards the nature of the argument students generate (e.g., Driver, Newton, & Osborne, 2000), and the way scholars conceptualize students' learning progressions (e.g., Mohan, Chen, & Anderson, 2009). The degree of adherence to the formal scientific narrative is an external metric that is culturally shaped by the nature of the discipline. In contrast, the metric that concerns us is an internal one that functions at the individual sense making level.

The basic idea that humans have an intrinsic sense of what counts as a good explanation is not new. Harman (1965), for instance, discussed people's inference of the best explanation, namely, the ability to draw inferences that do not follow logically from premises, but are rather based on temporal assumptions about what would be the best explanation. Keil, Rozenblit, and Mills (2004) noted the incomplete nature of understanding, terming it the "illusion of understanding", and suggested that asking people to explain makes people aware of gaps in their understanding. Gentner and colleagues pointed out that people evaluate analogical explanations based on their structural similarities (Gentner & Gentner, 1983), and Holyoak and Thagard (1995) emphasized

the role of pragmatic constraints (e.g., the goal of the explanation) in this evaluation. diSessa (1993) suggested that as humans interact with the physical world they develop simple unitary and non verbal schemas (phenomenological primitives, or p-prims) that when activated account for people's comfort with certain situations or surprise in others. Hence, explanations that resonate with an activated p-prim will be evaluated as more plausible (diSessa, 1993, forthcoming; Kapon & diSessa, 2012). Moreover, diSessa (1993) suggested that each p-prim has a *reliability priority* which results from processing feedback that can reinforce or undo the initial activation. Reliability priority expresses the degree of conviction a person attaches to a particular p-prim in a particular context of thought. High reliability priority means that it is unlikely that the p-prim will be deactivated by subsequent processing. Kapon and diSessa (2012) noted that often people activate conflicting explanatory primitives (not necessarily just p-prims) and they internally negotiate their relative priority. Parnafes (2012) conceptualized explanations as comprised of carefully assembled knowledge elements of different types. She argued that people are sensitive to the local coherence of these elements; namely, knowledge elements that comprise an explanation that makes sense cannot suggest conflicting inferences.

The above review suggests that people have an intrinsic sense of what counts as a good explanation, and thus the assumption that they have developed an internal metric for this purpose is reasonable. The question that concerns us is whether we can operationalize this internal metric into external traceable terms in sufficient detail that will allow us to explain the relative evaluations of competing explanations that individuals constantly make (i.e., why does a person prefer one explanation over another) by different values along this metric.

The Internal Metric for Certainty in Explanations: Three Dimensions

We believe that the internal metric for certainty in explanations are multidimensional, and that different dimensions are active in different contexts of thought. This is not a radical assumption. In a review entitled *Explanation and Understanding* Keil (2006) argued that "explanations can be contrasted in terms of domains roughly corresponding to academic disciplines" (p. 232). In fact, it has long been argued that even young children are sensitive to different domains of thought such as naive psychology (human behavior), naive biology (living things), and naive physics (the physical world around us) (Carey, 1985; diSessa, 1993; Inagaki & Hatano, 2002; Keil, 1994). Hence, it is likely that examining the internal metric for evaluating explanations in different domains of thought will help reveal different dimensions of this metric. Here we limit our discussion to a particular domain of thought, explanations about phenomena in the physical world; namely, phenomena that could, at least in principle, also be explained by formal scientific explanations. In this section we operationalize three dimensions of the internal metric that in our view are highly active in self-evaluating explanations about phenomena in the physical world: (1) Intuition; (2) Local coherence; (3) Mechanistic reasoning. We emphasize that these are *not* an exhaustive list of dimensions of the internal metric of certainty in explanations, and there are probably other dimensions. In what follows, we describe each of the three dimensions.

Intuition

Drawing and elaborating on diSessa's (1993) model of p-prims, Kapon and diSessa (2012) developed a model of explanation and change in explanation focusing on knowledge elements (KE) that provide a sense of satisfaction to those judging the explanation. They termed these KE explanatory primitives (e-prims). Like p-prims, e-prims are self-explanatory, unquestioned units of explanation, which students take as simply "the way things are," within the time span of evaluating an explanation. P-prims are a subcategory of e-prims, hence, not all e-prims are p-prims. For instance, a p-prim, by definition, is abstracted from experience in the physical world, whereas the source of an e-prim can be social interactions, language, or even explicit instruction. According to this model, explanations are accepted or rejected on the basis of the individual's convictions concerning particular e-prims and how these primitives fit the context of thought. The e-prims model suggests one dimension to our internal metric of certainty in explanations; namely, that at the base of any conviction lies an e-prim. Table 1 present a list of criteria for identifying e-prims.

Table 1: Criteria for identifying e-prims (Kapon & diSessa, 2012)

Criterion	Definition
Functionality	The knowledge element is explanatorily useful to the goal of reasoning and responsive to the context in which this reasoning takes place.
Obviousness	The knowledge element is regarded with explicit statements or unelaborated confidence and acceptance.
Developmental history	We are able to identify familiar experiences from which the knowledge element could have been abstracted.
Triangulation of expression	The knowledge element reappears frequently in a variety of manifestations during the reasoning process.

Criterion	Definition
Triangulation of form and content	The knowledge element matches documented p-prims or other documented intuitive notions. This criterion cannot always be satisfied, but when it does, we can consider our interpretation to be on a safer ground.

Local Coherence

The idea of consistency or coherence was used by Thagard (2007) for judging the quality of scientific explanations. In reference to Thagard's coherence, Hammer, Elby, Scherr, & Redish (2005) coined the term "local coherence", to highlight that the activation of knowledge elements and the formation of specific coherences are context sensitive. Sherin, Krakowsky, & Lee (2012) also supported this notion by suggesting that when a student explains a phenomenon, a temporary conceptual structure underlying the explanation exhibits consistency which may shift to a different explanation as the activity unfolds. Drawing on the notion of local coherence, Parnafes (2012) suggested that an explanation that makes sense to a student ties together a set of knowledge elements of different types that cohere, locally, in a specific context. A student's (temporary) explanation is viewed as a collection of activated knowledge elements that the student feels fit together at the particular time and for the particular purpose, hence the explanation make sense. Parnafes suggested that a student who is trying to understand a phenomenon goes through iterations of self-explanations, with temporarily stable stages of satisfaction with the explanation she generates in between. This comfort is interpreted as a temporary plateau of local coherence.

Mechanistic Reasoning

Causality plays a privileged role in explanations (Keil, 2003). Moreover, the ability to construct a mechanistic explanation for a physical phenomenon seems to be a developmental achievement (Metz, 1991). We hypothesize that the internal metric of certainty in explanations has a scale of causal strength of an explanation, and that this dimension is active in assessing explanations and self explanations about phenomena in the physical world. Russ, Scherr, Hammer, and Mikeska (2008) adapted an account of mechanism from the philosophy of science (Machamer et al., 2000) and developed a coding scheme that attempts to identify mechanistic reasoning in students' discourse in inquiry-based science classrooms. An underlying assumption of their coding scheme is that evidence of mechanistic reasoning in students' talk does not guarantee that the explanations they generate are scientifically correct. This assumption makes Russ et al.'s framework a promising tool to empirically assess the level of mechanistic reasoning in students' explanations, and whether it correlates with students' relative convictions in these explanations. Table 2 presents the coding scheme in Russ et al.. Note that Russ et al. conceptualized the first seven codes in Table 1 as hierarchically ordered, namely that code #7, chaining, is the highest evidence for mechanistic reasoning. The last two codes according to the Russ et al. framework are not necessarily connected to a particular level of mechanistic reasoning and the numbering of A & B are arbitrary.

Table 2: Coding scheme for mechanistic reasoning in students' talk (Russ et al., 2008)

Code	Definition
(1) Describing the Target Phenomenon	A clear statement or demonstration of the particular phenomenon or result that the student is trying to explain.
(2) Identifying Setup Conditions	Descriptions of the spatial and temporal organization of entities that enable the mechanism to run and produce the phenomenon.
(3) Identifying Entities	Recognizing the objects that affect the outcome of the phenomenon.
(4) Identifying Activities	Identifying the various doings in which the entities engage, articulate the actions and interactions that occur among entities, describe the things that entities do that cause changes in the surrounding entities.
(5) Identifying Properties of Entities	Identifying and articulating general properties of entities that are necessary for the particular mechanism to run.
(6) Identifying Organization of Entities	Identifying how the entities are spatially organized, where they are located, and how they are structured.
(7) Chaining Backward and Forward	Using knowledge about the causal structure of the world to make claims about what must have happened previously to bring about the current state of things (backward) or what will happen next given that certain entities or activities are present now (forward).
(A) Analogies	Using analogies to similar mechanisms in other contexts or fields as a framework for understanding new situations, comparing the target phenomenon to another phenomenon.

Code	Definition
(B) Animations	Using external animated models (gestures, body movements, etc.) to illustrate how certain entities act in the mechanism.

We now turn to a case study in which we highlight two explanations put forward by one student and analyze the internal metric in reference to these two explanations, using our framework.

Illustrative Case Study

The episode that will be analyzed here is drawn from a large corpus of data that was collected as part of the second author's project on the role of self-generated representations for promoting conceptual understanding. As part of the project, the second author and her research group interviewed pairs of students who were asked to generate, discuss and come to an agreement on explanations for phenomena in the physical world while generating representations and elaborating them. Interventions were kept to a minimum during students' discussions, and if made, were done for the purpose of clarifying meanings, or cuing students to consider overlooked conceptual issues. The episode (~ 3.5 minutes) is drawn from the first part of a 90 minute session. The students, two 6th graders, Reut and Natalie (pseudonyms), were trying to explain why a plastic bottle shrinks when air is pumped out of it. The analysis will focus on Reut's reasoning.

We chose the episode and the focus on Reut's reasoning since the explanations that Reut generated at the beginning of the episode and at the end of the episode were very different, and although she seemed quite sure of her explanation at the beginning of the episode she ended with a different one, allowing us to use our model to compare and contrast the two self-generated explanations, and account for Reut's internal evaluation. In each episode, we code and analyze the three dimensions, and then integrate the findings into one integrated metric. The three dimensions were analyzed as follows: (1) intuition: we listed all the knowledge elements in the explanation and those that were likely to be an e-prim, were checked using the criteria listed in Table 1. (2) Local coherence: we reexamined the knowledge elements that were coded for each explanation and examined the degree of fit. (3) Mechanistic reasoning was examined using the codes in Table 2.

Due to space limitations we cannot reproduce the full transcripts. To illustrate the function of the internal metric of certainty in explanation, we focus only on Reut's explanations and often regard the interaction with Natalie and the interviewer as part of the context in which Reut's explanation acquire its meaning, and add them as comments to Reut's explanation. We start by presenting the explanations that Reut generated, and then move to their analysis using our framework to explain why Reut preferred the second over the first. We systematically number each student's utterances to allow us to refer to specific utterances later in our analysis. Utterances are numbered as follows: student name, #of explanation, and a letter that numbers the utterances of this student. Thus, "Reut1B" denotes Reut's 1st explanation, second utterance. Italicized square brackets "[/]" denote our interpretive and other comments.

The aim of the analysis is to illustrate the operationalization of the internal metric of certainty in explanations. Hence, we examine the two explanations that Reut generated using the metric, explain her initial conviction in the first explanation, and the reason why she later preferred the second explanation over the first via differences between the two explanations along the three dimensions specified above.

Explanation 1

The students watched an experiment in which a plastic bottle shrinks when air is pumped out of it, and they then generated drawings to explain what happened. Using their own drawings they explained the phenomenon to each other, and occasionally the interviewers joined in. Reut and Natalie generated different explanations at the beginning of the episode. Table 3 presents Reut's first explanation.

Table 3: Reut's first explanation (three turns)

Reut1a:	It <i>[the bottle]</i> shrank in order to fill the vacuum, to reduce it.
Reut1b:	<i>[Answering a question from the interviewer "Why should the bottle care about the vacuum?"]</i> What... <i>[points to her drawing]</i> I don't know why this happens, but I know it does... I even don't know why I know it, but... if there is a vessel that has vacuum inside it and there are particles outside <i>[Natalie introduced "particles" to the discussion]</i> , or something... it always enters the vacuum to fill it. I don't know why it happens... hmmm...
Reut1c:	<i>Points again on her drawing</i> this is what I know, that always, if there is a vacuum so something will fill it. This is why here, there is more than – there is vacuum <i>[points to Natalie's drawing]</i> and the bottle shrank to fill it out <i>[demonstrates the change in the shape of the bottle with her hands]</i> . It's like, shrink inside.

Reut's Initial Evaluation to Natalie's Explanation

Natalie constructed a very different explanation. We provide here only what she said after Reut's turn 1C, since Reut later refers to this explanation.

Natalie 1D: Because the outside, it's like the pressure on the outside is bigger than the inside. A lot of pressure is activated towards the inside because the...ah... the number of particles outside, say for this area, is bigger than the number of particles inside for this area. So the outside pushes the bottle, to... shrink.

At this point the interviewer asked Reut: "What do you think Reut about what Natalie just said (Natalie 1D)?" Reut replied: "I have no idea if its right or wrong". In other words, at this point she did not think this explanation was better than her own.

Analysis of Explanation 1

Table 3 presents the development of Reut's first explanation during the discussion. We start by listing the knowledge elements that Reut used in her explanations:

Reut1A: (1) A vacuum tends to be filled; (2) A vacuum needs to be reduced.

Reut2A: (3) Particles are things (4) Things in the vicinity of vacuum will eventually enter in to fill it.

Reut 3A: (1) A vacuum tends to be filled

Intuition

We argue that the core of this explanation is an *e-prim* - "a vacuum has to be filled" - and that the knowledge elements # 1, 2, 4 are different manifestations of this. "A vacuum has to be filled" meets the first four mandatory criteria in Table 1 for identifying an *e-prim*: *Functionality*: it explains to Reut why the bottle shrank. *Obviousness*: Reut says: "I don't know why this happens, but I know it does... I even don't know why I know it" (Reut1b), "this is what I know" (Reut3c). *Developmental history*: source in language - the phrase "a vacuum will always be filled" is a common idiom in Hebrew (a nearly equivalent idiom in English is "nature abhors a vacuum"), source in interactions in the physical world - e.g., drinking from a straw, playing with a syringe. *Triangulation of expression*: KE 1, 2, 4 are different manifestations of the same *e-prim*. *Triangulation of form and content* - the notion that a vacuum has to be filled is mentioned in other studies. For instance Wiser and Smith (2008) discuss "a deeply held metaphysical principle that vacuum does not exist in nature" (p. 220), and diSessa (1993) identifies "vacuum impels" as a *p-prim*, but admits that this *p-prim* might be cultural to some extent. Given the fact that all five criteria were met we assumed that "a vacuum has to be filled" is an *e-prim* in Reut's knowledge system.

Local Coherence

All the knowledge elements that Reut activated (1, 2, 3, 4) locally cohere with one another and suggest coherent inferences that can explain the observed phenomenon in a convincing way: The bottle shrank to fill the vacuum because the vacuum had to be filled (Reut1c).

Mechanistic Reasoning

To establish this measure we used the codes listed in Table 2 (Russ et al., 2008).

Reut1a: (1) Describing the target phenomena ("the bottle shrank"); (3) identifying entities: ("vacuum");

Reut1b: (3) identifying entities ("vacuum", "vessel", "particles or something"); (4) identifying activities ("if there is a vessel that has vacuum inside it and there are particles outside, or something... it always enters the vacuum to fill it"); (5) identifying properties (the property of vacuum) (6) Identifying the organization of entities (the vacuum is inside the vessel the particles are outside the vessel).

Reut1c: (1) Describing the target phenomenon ("the bottle shrank"); (3) identifying entities: ("vacuum"); (5) identifying properties (the property of vacuum - "if there is a vacuum so something will fill it"), (6) Identifying organization of entities ("here there is more than...").

We argue that there is no chaining backwards (7) in this utterance, even though the structure of Reut's1c utterance approximates chaining backwards. We use the definition of chaining backwards in the following manner: using knowledge about the causal structure of the world (the property of a vacuum that it has to be filled) to make the claim (the bottle shrank to fill the vacuum) about what brings about the current state (the bottle has shrunk). In our view, this would be valid only if the bottle's "walls" were explicitly identified as

an entity, but they were not, and hence Reut's argument is circular. Without this identification, chaining is not part of Reut's explanation.

Summary

All in all, we argue that it is not surprising that Reut did not immediately abandon the explanation she generated. It was based on a strong intuition, it presented local coherence that was developed during the three utterances, and its causal strength was not complete but was quite impressive, since apart from chaining, almost all the components of mechanistic reasoning seemed to be present (#1, 3, 4, 5, 6 in Table 2).

Explanation 2

The interviewer asked Reut: "Can you repeat what you understood *[from Natalie 1D]*?" Table 4 presents the conversation between Natalie, Reut and the interviewer in which Reut constructed her second explanation, which started as a mere repetition of Natalie's idea and gradually developed and became "her own".

Table 4: Reut's second explanation

Reut2a:	What I understood... like, what I understood - there is the bottle... now... hm... one second <i>[looks to the side]</i> ... first the particles move, right? So what happens is that they bounce into the bottle and they push it inside <i>[gestures pushing with her hand]</i> . Now, before, there were more, there were many particles...
Int.:	Can you explain it using one of the drawings? Try doing it here <i>[gives them a blank page]</i>
Reut2b:	<i>[Reut and Natalie take the paper]</i> Can I try on this? Say this is a bottle <i>[draws a rectangle]</i> lovely <i>[laughs]</i> now what is happening, we have plenty of particles here that push in this direction, and try.... to push it out <i>[gesturing "out" with her hand]</i> and particles from this direction that try to push it inside <i>[draw arrows on the page and gesture with her hand "inside"]</i> . This is what I understood, this is what I know.
Natalie:	And now with the pump, we took out, like,
Reut2c:	Yeah
Natalie:	Lots of particle from here
Reut2d:	So now we have here, like just two particles...
Natalie:	More... more force pushing from the outside
Reut2e:	And less force to resist, so it shrinks. It has nothing to do with a vacuum.

Reut Prefers the Second Explanation Over the First

Note that in the second part of utterance Reut2e, Reut states "It has nothing to do with a vacuum." Thus she rejected her previous explanation and accepted the current one.

Analysis of Explanation 2

Table 4 presents the development of Reut's first explanation during the discussion. Before we start, we would like to note the interviewer response when Reut said that she had no idea whether Natalie's explanation is correct or not. The interviewer asked her to restate it. During her restatement and elaboration of this explanation she changed her mind and became convinced of the explanation that *she* constructed during this process. Again, we start by listing the knowledge elements that Reut used in her explanations:

- Reut2a: (1) the particles move; (2) the particles bounce on the walls of the bottle and push it; (3) particles outside of the bottle push the bottle's walls inside (4) pumping takes the particles out of the bottle.
- Reut2b: (2) the particles bounce on the walls of the bottle and push it; (3) particles outside of the bottle push the bottle's walls inside; (5) particles in the bottle push the bottle's walls outside.
- Reut2c: (4) pumping takes the particles out of the bottle.
- Reut2d: (4) pumping takes the particles out of the bottle.
- Reut2e: (6) fewer particles means a smaller pushing force; (7) the stronger force wins (overcoming p-prim; diSessa, 1993).

Intuition

At the core of Reut conviction lies the p-prim - "overcoming" - the intuitive schematization of one force or influence that overpowers another (diSessa, 1993). Note that Reut did not start the explanation with conviction. At first she was merely reconstructing Natalie's explanation, and her use of the word "push" (Reut2A) was to reference Natalie's explanation (Natalie1d). But in the next utterance Reut discussed conflicting forces: "we have plenty of particles here that push in this direction, and try.... to push it out *[gesturing "out" with her hand]*

and particles from this direction that try to push it inside [*draw arrows on the page and gesture with her hand "inside"*] (Reut2b). Note the triangulation of expressions for conflicting forces. Not only did Reut verbalize the conflicting forces, she also simulated them with her hands. P-prims are non verbally encoded, and it may well be that the gestures imitating the conflicting forces activated the notion of overcoming. After Reut explicitly expressed how pumping will result in fewer particles in (Reut2c & Reut2d, KE #4) and that fewer bouncing particles means less force (Reut2E, KE#6), the overcoming p-prim was activated, and was used with strong conviction.

Note that the reliability priority of overcoming is much higher than that of the e-prim "a vacuum has to be filled", and after it was activated Reut concluded that the shrinking of the bottle "has nothing to do with a vacuum" (Reut2e). The differences in the reliability priority could be attributed to the fact that e-prims that are also p-prims are abstracted from the very basic and frequent interactions with the physical world, whereas the roots of an e-prim such as "a vacuum has to be filled" are not as deep.

Coherence

The knowledge elements that Reut activated (#1-7) locally cohere and support one another, suggesting coherent inferences that can explain the observed phenomenon in a convincing way: The bottle shrunk since the particles out of the bottle pushed on the bottle harder than the opposing push of the particles in the bottle (Reut2e).

Mechanistic Reasoning

- Reut2a*: (3) Identifying entities (bottle, particles); (4) Identifying activities (the particles bounce, push); (5) Identifying properties of entities (particles move); (B) Animated model (Reut gestures with her hand to mimic the pushing of the particles)
- Reut2b*: (3) Identifying entities (bottle, particles); (4) Identifying activities (the particles push); (6) Identifying organization of entities (particles outside and inside of the bottle); (B) Animated model (Reut gestures with her hand to animate the conflicting directions of the particles being pushed in and out of the bottle)
- Reut2c*: (2) Identifying setup conditions (when we start pumping we take particles out of the bottle); (6) Identifying organization of entities (particles outside and inside)
- Reut2d*: (2) Identifying setup conditions (due to the pumping there are very few particles left in the bottle); (6) Identifying organization of entities ("we have *here* like two particles")
- Reut2e*: (7) Chaining backwards - using knowledge about the causal structure of the world (pumping the bottle took out particles from the bottle, thus fewer particles hit the inside of the bottle relative to those that hit the outer walls) to make claims (the force applied by more particles outside the bottle, overcame force applied by fewer particles inside the bottle) about what brought about the current state (thus, the bottle has shrunk).

Summary

Reut changed her mind and preferred the second explanation to the first one. This was accounted for in terms of the three dimensions of the internal metric for certainty in explanations. *Intuition*: The second explanation activated a stronger intuition than the first one (a p-prim vs. "just" an e-prim). *Local coherence*: The degree of coherence of the second explanation was higher than the first, since more knowledge elements cohered and supported one another. *Mechanistic reasoning*: The degree of mechanistic reasoning was higher since *all* the components of mechanistic reasoning were present.

Conclusion

In this paper we attempted to highlight students' internal evaluation of explanations, as opposed to external evaluations of explanations that are based on compatibility to scientific formulations. We explored the sources of students' sense of conviction of their own explanations and what guides their preferences for one explanation over another. We described the process of internal evaluation with a metaphorical theoretical construct *the internal metric for certainty in explanation*. The main contribution of this paper is the operationalization of this internal metric into empirical traceable terms. We envision the internal metric as composed of several dimensions. In the case of the internal evaluation of explanations of phenomena in the physical world, we suggested three dominant dimensions that form the internal metric: intuition, local coherence, and mechanistic reasoning. We operationalized each dimension and specified criteria for its evaluation based on students' self generated explanations. The illustrative analysis shows that the framework can explain conviction in an explanation, as well as relative conviction and the preference of one explanation over another.

It is important to note that paths of developing explanations guided by external evaluation and internal evaluation may not necessarily converge. This raises the the issue of the educational merits of the ability to

externally measure students' internal convictions. We argue that students' agency in the developmental path from intuitive to scientific explanation is crucial for maintaining a sense of understanding, and hence the ability to assess students' convictions in explanations is a key component in instructional design.

References

- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421-441.
- Berk, L. E. (1985). Why children talk to themselves. *Young Children*, 40(5), 46-52.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- Chi, M. T. H., 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.
- Dagher, Z. R., & Cossman, G. (1992). Verbal explanations given by science teachers: Their nature and implications. *Journal of Research in Science Teaching*, 29(4), 361-374.
- diSessa, A. A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, 10(2&3), 105-225.
- diSessa, A. A. (forthcoming). The Construction of Causal Schemes: Learning Mechanisms at the Knowledge Level. *Cognitive science*.
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science education*, 84(3), 287-312.
- Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy*, 71(1), 5-19.
- Gentner, D., & Gentner, D. R. (1983). Flowing waters or moving crowd: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental Models* (pp. 99-130). Hillsdale, NJ: Lawrence Erlbaum.
- Gopnik, A. (1998). Explanation as Orgasm*. *Minds and machines*, 8(1), 101-118.
- Harman, G. H. (1965). The inference to the best explanation. *The Philosophical Review*, 74(1), 88-95.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps : analogy in creative thought*. Cambridge, MA: MIT Press.
- Inagaki, K., & Hatano, G. (2002). *Young children's thinking about biological world*. New York, NY: Psychology Press.
- Kapon, S., & diSessa, A. A. (2012). Reasoning Through Instructional Analogies. *Cognition and Instruction*, 30(3), 261-310.
- Keil, F. C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. Hirschfield & S. Gelman (Eds.), *apping the mind: Domain specificity in cognition and culture* (pp. 234-254). Cambridge, UK: Cambridge University Press.
- Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in cognitive sciences*, 7(8), 368-373.
- Keil, F. C. (2006). Explanation and understanding. *Annual review of psychology*, 57, 227-254.
- Keil, F. C., Rozenblit, L., & Mills, C. (2004). What lies beneath? Understanding the limits of understanding. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 227-249). Cambridge, Mass: MIT Press.
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of science*, 1-25.
- Metz, K. E. (1991). Development of explanation: Incremental and fundamental change in children's physics knowledge. *Journal of Research in Science Teaching*, 28(9), 785-797.
- 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.
- Parnafes, O. (2012). Developing Explanations and Developing Understanding: Students Explain the Phases of the Moon Using Visual Representations. *Cognition and Instruction*, 30(4), 359-403.
- 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.
- Salmon, W. C. (1989). Four decade of scientific explanation. In P. Kitcher & W. C. Salmon (Eds.), *Scientific explanation, Minnesota studies in the philosophy of science* (Vol. 13, pp. 3-219). Minneapolis: University of Minnesota Press.
- Toulmin, S. E. (1972). *Human understanding*. Princeton, NJ: Princeton University Press.
- Wiser, M., & Smith, C. (2008). Learning and teaching about matter in grades K-8: When should the atomic-molecular theory be introduced? In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 205 – 237). NY: Routledge.

Acknowledgments

We would like to thank Rosemary Russ for a useful discussion about the application of the framework of mechanistic reasoning.

Exploring A Digital Tool for Exchanging Ideas During Science Inquiry

Camillia Matuk and Marcia C. Linn, University of California, Berkeley
Email: cmatuk@berkeley.edu, mclinn@berkeley.edu

Abstract: Practicing science increasingly involves knowing how to participate in a networked knowledge community. This includes expressing scientifically informed ideas, sharing ideas with peers, and evaluating multiple sources of information. Effective instruction builds on students' prior ideas, enables them to benefit from exchanging ideas with others, and supports them learning from one another. How might technology support these exchanges? And how might documenting these exchanges inform teachers' and researchers' improvements to their instruction and design? We describe the Public Idea Manager, a new curriculum-integrated tool that supports students exchanging ideas during web-based science inquiry. Our exploratory analyses show relationships between the diversity and sources of students' ideas and the quality of their explanations. We discuss implications for formative assessment, and for the role of technology in supporting students to engage more meaningfully with information and with each other.

Supporting the Exchange of Ideas During Science Inquiry

One of the most effective ways to support the understanding of science is to build upon learners' prior knowledge. But amid the multiple constraints of the classroom, it can be difficult for teachers to address students' diverse ideas, which are often conflicting and incomplete (diSessa, 2000). Tools that promote collaboration offer numerous advantages for students promoting one another's understanding. They also emphasize participation in authentic scientific communities. The potential for such tools to capture fine-grained information on students' exchanges can moreover help researchers and teachers better understand and support students' inquiry.

In this paper, we describe a curriculum-integrated tool that supports students exchanging ideas throughout the process of constructing scientific explanations. We present classroom findings on how middle school students used the tool during a life sciences unit, and how the sources and diversity of their ideas related to the quality of their explanations. We discuss implications of our findings for inquiry instruction, and next steps in research and design.

Theoretical Background

Explanation is a hallmark of science inquiry, but challenging for students. They have difficulty using evidence to support their claims (Sandoval, 2005), coordinating evidence from multiple sources and with multiple alternative hypotheses (Kuhn et al., 1995; Schauble, 1996), refining arguments in light of new evidence (Chinn & Brewer, 1998), and articulating ideas in writing (McNeill & Krajcik, 2008; Sandoval & Millwood, 2005). Planning activities, such as generating, organizing, and linearizing ideas can help, but students rarely do this (Andriessen et al., 1996). It is therefore not surprising that continuous guidance for developing explanations is useful (Quintana et al., 2004).

Our research and curriculum design is guided by the Knowledge Integration (KI) pattern (Linn & Eylon, 2011). This perspective assumes that learners hold multiple, often contradictory views of any one science topic and that learners deliberately distinguish and make connections between new and existing ideas (diSessa, 2000; Eylon & Linn, 1988; Linn & Hsi, 2000; Slotta et al., 1995). Instruction guided by KI thus emphasizes eliciting students' existing ideas; helping them explore additional; more normative ideas; and guiding them to reflect on distinguishing and sorting out alternatives as they build an integrated understanding.

In other words, the KI pattern would suggest learning benefits in initially diversifying and then converging on ideas as learners reflect upon and refine their understanding. Research on brainstorming across domains finds differential benefits of convergent and divergent thinking during different phases of problem solving (Crompton, 2006). Other research suggests that the kind of information encountered in collaborative knowledge building environments can affect how one revises their ideas. Exposure to incongruous information in Wikipedia, for instance, was more likely to prompt editors to revise their ideas than exposure to congruous information (Moskaliuk, Kimmerle, & Cress, 2012).

Existing collaborative environments offer tools with which learners can build upon one another's ideas (e.g., Scardamalia & Bereiter, 2006). Often, however, middle schools students require more explicit scaffolding to guide their construction of scientific explanations than are available in such environments. Questions also remain over how to make sense of what occurs during the free exchange of ideas. For instance, one might anticipate the exposure to peers' ideas to result in either the diversification or the reinforcement of students'

own ideas. But how do students make decisions over which ideas from their peers to incorporate? And how can technology capture these decisions to inform more effective instruction?

Goals and Research Questions

We designed a tool called the *Public Idea Manager (IM)*, which would (1) allow students to access each others' ideas as learning resources when refining their own understanding; and (2) capture students' exchanges of ideas in order to inform teachers and researchers of better ways to support students collaborative learning.

In this initial exploratory study of a larger design-based research program, we wished to know how students would take advantage of the tool. Our specific research questions were (1) how would students use a tool that supports the exchange of ideas among peers? That is, would they freely share ideas, or would they be reluctant to do so? Would they tend to copy ideas from their peers rather than generate their own? When students do copy peers' ideas, would they recognize good quality ideas? And (2) how would the diversity of students' individual repertoires of ideas change as a result of their access to a public idea repository? Moreover, how might the relative diversity of students' ideas, and the source of those ideas (generated by oneself or copied from one's peers), relate to the quality of students' scientific explanations?

Through the design of a tool for exchanging ideas, this research provides opportunities for students to participate in more authentic collaborative science inquiry. It moreover provides researchers and teachers ways to better understand the role of technology in mediating students engaging with each other's ideas, and drawing on their peers to support their explanations.

Methods

Participants and Setting

Participants were 297 students in a middle school in the western United States. Their two teachers had 5 class periods each, and collaborated closely to coordinate lesson plans and grading. One teacher had taught the WISE unit (described below) for more than 8 years, while the other teacher had taught it for 3 years. Students mostly worked as partners on the unit during one 50-minute class period on each of 7 consecutive school days. Two more days before and after this time were spent completing a pre and posttest (described below). Meanwhile, the teachers circulated the classroom to offer guidance on the material or assistance with the technology as needed. A researcher was also present in the room on most days, but mainly sat apart from the class to record observations.

WISE and the Public Idea Manager

The platform for our curriculum and technology development is the Web-based Inquiry Science Environment (WISE, wise.berkeley.edu). WISE is a free, open-source, and customizable online learning environment. With WISE, students engage in self-paced, collaborative investigations, in which they collect and interpret evidence from dynamic visualizations, and use various tools to express their understanding. WISE meanwhile logs students' interactions, which enables teachers to grade and give feedback, and designers and researchers to better study the impacts of technology on student learning.

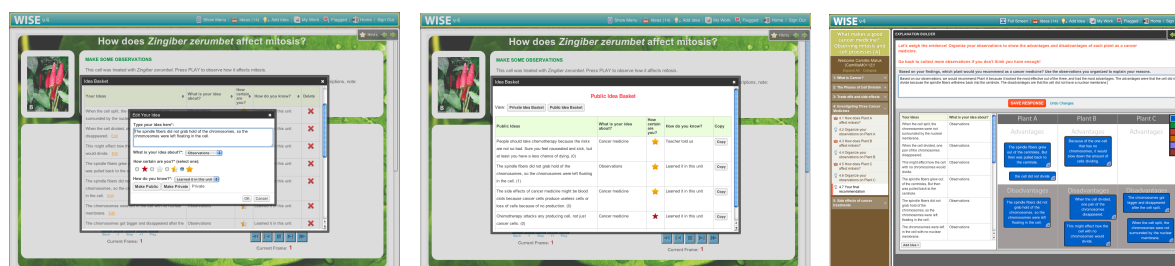


Figure 1. Over the course of a unit, students enter short text entries into their *Idea Baskets* (left), and specify attributes (e.g., source, tags, rating). Students can choose to share any private idea to the *Public Basket*, from which they may also copy any idea into their *Private Baskets*. Entries accumulate in a sortable list to which students can return and revise (middle). In the *Explanation Builder* (right), students drag ideas into author-defined categories, and refer to these as they write an explanation in response to a prompt.

The *Public Idea Manager* (Figure 1) is a feature of the *Idea Manager (IM)*, a tool integrated into WISE. The IM scaffolds explanation according to the KI pattern by making explicit the acts of gathering, distinguishing, and sorting ideas (Matuk, et al., 2012). As WISE logs these actions, the IM provides teachers and researchers a record of students' changing ideas that can inform instruction and design. For example, prior research used the IM in a chemistry unit to identify when in the process of explaining students found specific

ideas more or less difficult to apply (McElhaney, et al., 2011). Such analyses indicate the kind of assistance needed, as well as which students need it. As such, the IM adds value over typical pre and posttests of students' understanding. The Public IM extends the functionality of the IM from a private repository to a common space through which students can exchange ideas. First piloted in classrooms as a hybrid prototype (Matuk, et al., 2013), this feature makes it possible to track how ideas emerge and spread within a classroom, and how exchanging ideas affects students' developing understanding.

The design of the Public IM was inspired by previous and existing argumentation and explanation scaffolds (e.g., Bell, 2004; Zhang & Quintana, 2012). By supporting freeform documentation, the tool allows for the various kinds of ideas students may have, which do not always conform to teachers' and designers' expectations. The interface of the Public IM is moreover customizable, including the selection of idea attributes, the categories for sorting, and the names *Idea Basket* and *Explanation Builder*. As such, teachers and researchers can adapt the tool to align with their own research and instructional goals.

Integrating the Public Idea Manager into *Mitosis*

What Makes a Good Cancer Medicine? (aka, *Mitosis*) is a middle school unit publicly available at wise.berkeley.edu (Preview the unit at <http://wise.berkeley.edu/webapp/previewproject.html?projectId=6498>). In it, students assist a fictional scientist in investigating the potential of three different plant-derived chemicals in treating cancer. This sequence of activities was designed to follow the Knowledge Integration pattern by structuring activities to support eliciting, adding, distinguishing, organizing, and reflecting on ideas as students prepare to write recommendations for cancer medicines based on their observations (Matuk & Linn, 2013).

As students examine animations to compare the effects of each chemical on cell division, they learn the phases of mitosis and cell division. At several points throughout the unit, students are prompted to gather ideas in the IM about specific topics (e.g., cell division, cancer medicine, observations made of the animations). At four different points in the unit, students sort their ideas in preparation to write explanations (*What happens when cells divide? How might a medicine stop cancer cells from dividing? Which treatment would you recommend? and Why does hair fall out during cancer treatment?*). At these same points, students are encouraged to share at least one of their ideas to the Public Basket, and to copy at least one of the Public ideas into their Private Baskets. After each exchange, students are asked to write justifications for their choices of ideas.

Data and Analysis

Because this study is part of a larger research program on technology-enhanced curricula, we focus our analysis on just the relevant portion of all data collected. Specifically, our analysis aimed to explore how students exchanged ideas through the Public Idea Manager, and how the diversity of their ideas related to the quality of their explanations. Thus, we sought correlations in student log files on private ideas generated, private ideas shared to the Public Basket, and public ideas copied into students' Private Baskets. We analyzed students' written explanations to culminating questions in the final Idea Manager sequence of the unit, and analyzed their written justifications for exchanging ideas. Details of our analysis are described below. We triangulated this data with classroom field note observations, video of student pairs working together on the unit, and retrospective interviews with teachers, from which we selected cases to illustrate and explain our quantitative findings. Among other questions in the interviews, we asked teachers to reflect on students' uses of the tool, and to share their observations of its impact on students' learning and collaboration.

A pre and posttest included a paper-and-pencil test that asked students to draw and describe the phases of cell division, and a computer-based test with open-ended and multiple-choice questions that asked students to use their understanding of cell division to reason about the mechanism of an effective cancer treatment. These assessments were designed to measure the overall effectiveness of the unit on improving students' conceptual understanding of key ideas introduced in the unit. Each item was scored on a 5-point rubric that measured the degree to which students integrated ideas into a coherent explanation (Linn & Eylon, 2011).

Results

Pre and Post-Test Gains

Students showed significant average gains from the pre to the posttest ($M=1.38$, $SD=1.09$, $t(219)=-12.88$, $p<.0001$). This indicates that the unit had a positive impact on students' learning. In what follows, we discuss findings that illuminate students' uses of the Idea Manager.

How Did Students Exchange Ideas?

By the end of the unit, students' Private Baskets contained an average of 18.7 ideas ($SD = 6.8$). A significantly greater number of these ideas were self-generated ($M=14.53$, $SD=5.52$) than were chosen from the Public Basket ($M=4.17$, $SD=3.42$). This finding is encouraging because it suggests that in spite of their access to the

Public Basket, students were not reliant on their peers for ideas. Instead, they tended to start with ideas of their own.

Students appeared more inclined to contribute ideas to the Public Basket ($M=5.11$) than they were to copy ideas ($M=4.26$). They moreover appeared highly selective in the ideas they decided to exchange, as only 22.8 - 27.3% of the mean 18.7 ideas in their Private Baskets were ever in circulation. To determine whether students were making good decisions in their choices of Public Ideas, we scored each Public idea based on how well it integrated two or more relevant concepts into a supported statement (Table 1.)

Table 1. Scoring rubric for the quality of ideas in the Public Basket

Score	Description	Example
0	No response	(None found)
1	Off task	(None found)
2	Uninterpretable; nonnormative ideas or observations;	They divide during mitosis because cancer cell make them spread the sickness.
3	Declarative or factual statements or definitions; unconnected ideas; or a mixture of normative and nonnormative ideas.	its bad, you can die, theres many types of cancer.
4	A normative observation or claim, but lacks interpretation, explanation, or sufficient supporting evidence.	The medicine stops the centrioles and spindle fibers from pulling the choromosomes apart.
5	Integrates more than one concept into a well-supported, normative claim; offers a causal explanation for an observation.	The cell was able to undergo mitosis, but one of the cell's chromosomes was obliterated by the Zingiber zerumbet, possibly making that cell useless.

Public ideas had a mean quality score of 3.77, and were copied a mean number of 0.83 times. Results show that higher quality public ideas also tended to have been copied more frequently ($F(2, 721) = 10.76, p<.0001$) (Figure 2). This finding suggests that students appear to successfully recognize good quality ideas from among those available in the Public Basket.

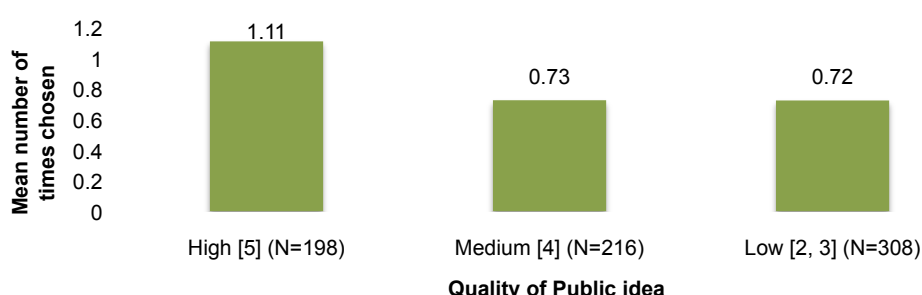


Figure 2. Students more frequently chose higher quality ideas than they chose medium or low quality ideas from the Public Basket. Quality scores are shown in square brackets.

Students' Reasons for Exchanging Public Ideas

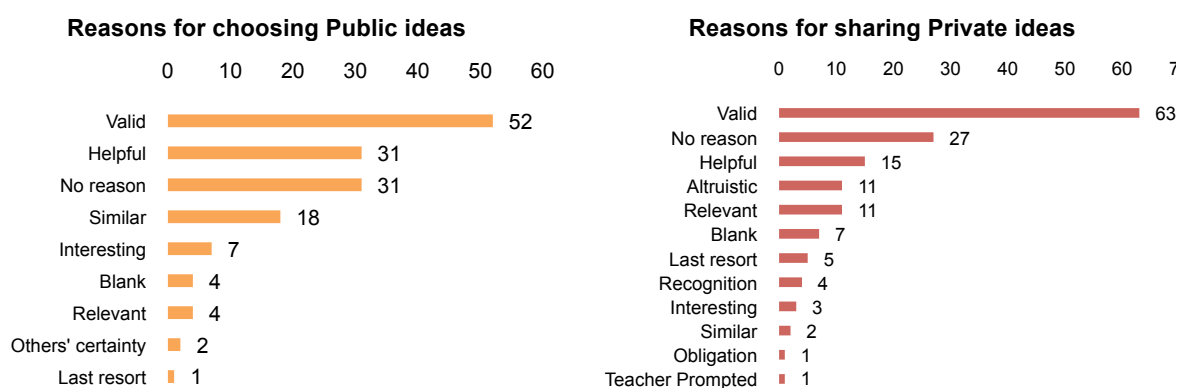


Figure 3. Categories of students' justifications for choosing and sharing ideas midway through the unit.

To explore whether students were consciously choosing ideas based on their quality, we examined their written responses to embedded prompts to justify their decisions. In these, students expressed a variety of reasons for

exchanging ideas in the Public Basket (Figures 3 and 4). Most commonly, students based their decisions to share and to copy ideas on the idea's perceived validity ("... because we agree with it the most." "...it was a very smart answer and it seemed like something that would be true.").

Reasons for sharing private ideas additionally included a desire for peer recognition ("...we wanted to... see what [our classmates] thought (if they [chose] our idea or not)."), and a desire to improve the Public Basket ("[our idea] seemed accurate [sic] compared to the other (students') ideas." "... we thought it was a good idea and nobody has it yet." "... it was a general statement that other classmates would understand."). Altogether, these findings suggest that students engaged in sophisticated decision-making when evaluating their ideas against those of their peers. In sharing ideas, some students even appeared to consciously consider the value of their contributions to this collaborative space.

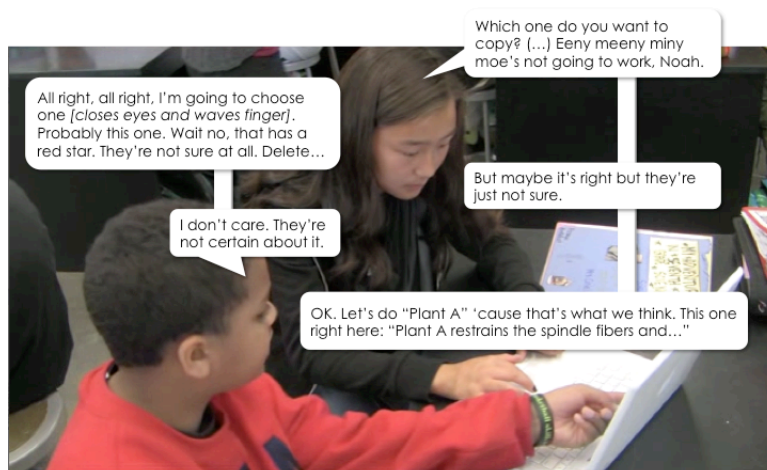


Figure 4. In choosing which Public idea to copy, these students consider various strategies, including random selection, the certainty of the idea's contributor, and the similarity of the idea to existing ideas of their own.

How do Idea Sources and Diversity Relate to Explanation Quality?

Notably, approximately 20.67% of students chose Public ideas because they were helpful; that is, because they added new information to their thinking ("... because it was well written and explained part of the lesson." "... because we thought that it was similar to our idea but different enough to provide food for thought."). In contrast, 12% of students chose ideas because of their similarity to existing private ideas ("...because we wrote a private idea similar to that...").

If students were indeed choosing Public ideas based on difference or similarity with respect to their own ideas, we might expect this to reflect in the actual diversity of students' Private ideas. The relative benefits of diversifying vs. converging on ideas might moreover be evident in the quality of students' explanations. Specifically, how well do students explain who tend to collect ideas that simply agree with their own? How well do students explain who tend to collect ideas that conflict with, or that otherwise diversify, their own ideas? To explore these questions, we coded each of students' Private Basket ideas according to whether it was unique or redundant relative to the rest of the ideas in the Basket. *Unique* ideas added new information not already present in the Basket, whereas *redundant* ideas restated already existing ideas. Results show that students' Private Idea Baskets contained a greater proportion of unique as opposed to redundant ideas, regardless of the source (i.e., self-generated or copied from the Public Basket) (Figure 5).

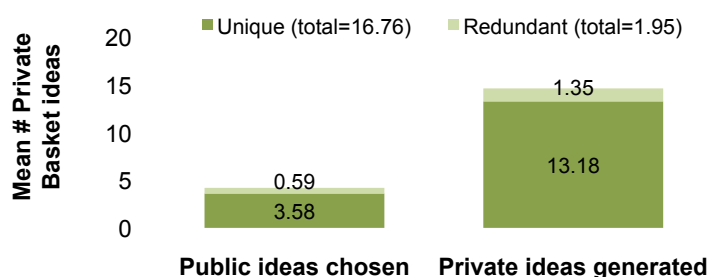


Figure 5. Proportion of unique and redundant ideas by source

To explore the relationship between idea diversity and students' abilities to explain, we scored students' explanations to the final prompt in the unit: "*Maya heard that her mother's hair might fall out during her cancer treatment. Why would this happen?*" Explanations were scored based on the number of links students made between key ideas presented in the unit (Table 2):

- Cancer is when cells divide rapidly/out of control.
- Cancer treatment stops cell division to treat cancer.
- Chemotherapy targets rapidly dividing cells.
- Chemotherapy also stops normal cells from dividing.
- Skin/hair cells are rapidly dividing cells.
- When hair cells aren't replaced with new ones, hair falls out.
- When skin/hair follicle cells die, they can no longer hold hair in the scalp and the hair falls out.

Table 2: Scoring rubric for explanations

Score	Description
0	Blank
1	Off-task, not possible to interpret
2	Non-normative, lacks explanation, doesn't address the question
3	1 normative idea
4	2 linked normative ideas
5	Elaborated response with 3+ linked normative ideas

Results show the relative diversity of ideas in students' Private Baskets to be correlated with the quality of their explanations at the unit's end. Specifically, students with the poorest explanations also tended to have a greater number of unique ideas in their Private Baskets, whereas students with the best explanations tended to have a greater number of redundant ideas ($F(2, 142) = 6.04, p < .005$) (Figure 6).

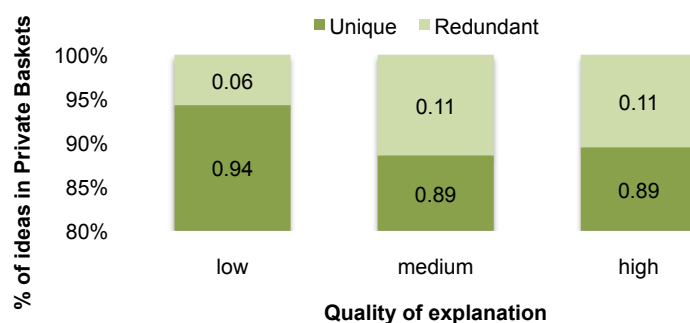


Figure 6. Proportion of unique and redundant ideas in students' Private Baskets, by explanation quality

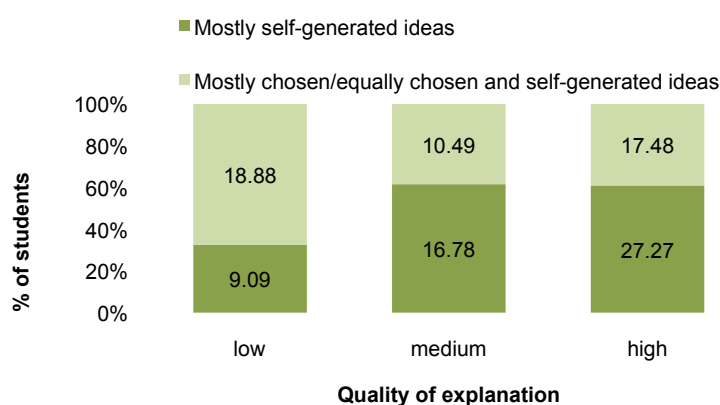


Figure 7. Sources of redundant ideas by quality of explanation

Further analysis shows a relationship between the quality of students' explanations and the sources of these redundant ideas: students themselves or their peers. That is, students who wrote better explanations also tended to have self-generated more of the redundant ideas in their Private Baskets. In contrast, students who wrote poorer explanations tended to have mostly chosen, or to have equally chosen and self-generated, their redundant ideas ($\chi^2 = 9.511, df = 2, p < .01$) (Figure 7).

How the Public IM Enhanced Teachers' Roles

In her interview, one teacher noted how prior to this tool, she struggled to keep her students from simply relying on the smart kids to give the answers. Because of its anonymity, the Public Basket placed all students' ideas on equal footing. It was such that the teacher noticed that her students' discussions are more interesting" as they made serious attempts to consider and understand their peers' ideas.

The tool also helped teachers provide formative feedback. Instead of waiting until students had gone too far down a misguided path, teachers would scan through WISE's grading interface, notice how students were organizing their ideas, and identify students in need of individual assistance. Teachers would also monitor students' ideas to identify which topics appeared most challenging overall. Based on this information, teachers would tailor whole-class opener activities to reinforce the unit's instruction. Future work will explore visualization tools to help teachers more easily monitor patterns in students' ideas; as well as a feature that allows teachers and researchers to moderate students' exchanges, such as by seeding the Public Basket with ideas, or promoting promising ideas for other students to notice.

Discussion and Implications

Using a new tool for students to exchange ideas during web-based science inquiry, we found relationships between the sources and diversity of the students' ideas, and the quality of their explanations. Interestingly, students who had generated more redundant ideas also tended to have constructed more coherent explanations. One reason for this observation is that students who were already likely to produce high quality explanations managed to identify the key ideas early in the unit. These students may thus have tended to rephrase these same key ideas whenever prompted to add new entries. Meanwhile, students who were already likely to produce poor quality explanations may have been less able to recognize relevant, normative ideas. It follows that a highly diverse set of ideas may indicate students who need more support distinguishing among their many ideas.

Another explanation for this finding is that generating redundant ideas has cognitive benefits. It may help students refine their understanding, as does self-explanation (Chi, DeLeeuw, Chiu, & LaVancher, 1994; Siegler, 2002). Elaborating by rearticulating ideas may be a metacognitive learning strategy that involves actively creating links between new and prior knowledge (Mayer, 2002; Weinstein & Mayer, 1986). As with rewriting and revision, students who generate redundancy may be re-evaluating and clarifying their thinking (Ladd, 2003; Fitzgerald, 1992; Scardamalia & Bereiter, 1994).

An advantage of this tool is that it allows careful engineering of students' interactions with information, and thus, ways to explore variations in scaffolding the exchange of ideas. As analysis of these data continues, we are investigating the effects of prompting students to use the Public IM either to diversify or to reinforce their ideas, and how exposure to either diverse or redundant ideas might influence how students revise their explanations. Among other questions, we will explore when in the process of explanation (e.g., gathering, distinguishing, or sorting ideas), and for which students (e.g., low vs. high prior knowledge) each strategy might be most effective. We will also trace specific self-generated and copied ideas over time to see how these become integrated into students' explanations.

Revisions to the technology may include features that support effective discourse between students, and more deliberative choices around the exchange of ideas. Future research might explore how lessons learned about the role of technology in the exchange of ideas might be applied to different contexts, such as engineering design, medical decision-making, and other problem-based learning scenarios.

Learning and Becoming in Practice

This research relates to the theme of *Learning and Becoming in Practice* in multiple ways. First, it attends to the notion that science inquiry entails participation in a global knowledge community. This involves developing and practicing various collaborative skills, including expressing scientifically informed ideas, sharing ideas with peers, and evaluating multiple sources of information. Second, our study explored the role of technology in helping students thus engage meaningfully with information, and with each other. We observed how our tool supported teachers attending to their students' ideas throughout their inquiry, thus giving them the means to focus assessment on learning processes and on the development of scientific practices, rather than just on outcomes. Finally, our approach to design is one that emphasizes sustainability. By being customizable, for example, the Public IM allows both researchers and teachers to explore different questions about how students learn, and how to support it. By also involving various stakeholders in the design of the tool, our design process invites others to contribute to the tool's improvement. These features help maintain the tool's relevance, and ensure its usability and usefulness over time.

References

- Andriessen, J. E. B., Erkens, G., Overeem, E., & Jaspers, J. (1996, September). Using complex information in argumentation for collaborative text production. Paper presented on the First Conference on Using Complex Information Systems (UCIS'96), Poitiers, France, 1996.

- Bell, P., & Linn, M. (2000). Scientific arguments as learning artifacts: designing for learning from the web with KIE. *International Journal of Science Education*, 22(8), 797-817.
- Bell, P. (2004). Promoting students' argument construction and collaborative debate in the science classroom. In M. C. Linn, E. A. Davis & P. Bell (Eds.), *Internet environments for science education* (pp. 115-143). Mahwah, NJ: Erlbaum.
- Chi, M. T., De Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, 18(3), 439-477.
- Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching*, 35(6), 623-654.
- Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal*, 18(3), 391-404.
- diSessa, A. A. (2000). *Changing minds: Computers, learning, and literacy*. Cambridge, MA: MIT Press.
- Fitzgerald, J. (1992). *Towards knowledge in writing*. New York, NY: Springer-Verlag.
- Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C. (1995). Strategies of knowledge acquisition. With Commentary by Sheldon H. White and by David Klahr and Sharon M. Carver; and a Reply by Deanna Kuhn. *Monographs of the Society for Research in Child Development*, 60(4, Serial No. 245).
- Ladd, B. C. (2003). It's all writing: Experience using rewriting to learn in introductory computer science. *Journal of Computing Sciences in Colleges*, 18(5), 57-64.
- Linn, M.C. & Eylon, B-S. (2011). *Science Learning and Instruction: Taking Advantage of Technology to Promote Knowledge Integration*. New York, NY: Routledge.
- Linn, M. C., Hsi, S. H., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Matuk, C. F. & Linn, M. C. (2013, April 27 - May 1). *Technology Integration to Scaffold and Assess Students Use of Visual Evidence In Science Inquiry*. Paper presented at the American Educational Research Association Meeting (AERA2013): Education and Poverty: Theory, Research, Policy and Praxis, San Francisco, CA, USA.
- Matuk, C., McElhaney, K., Miller, D., King Chen, et al. (2013). Reflectively prototyping a tool for exchanging ideas. In *CSCL'13: Proceedings of the 10th International Conference on Computer Supported Collaborative Learning*, (Vol. 2, pp. 101-104). Madison, WI, 2013. International Society of the Learning Sciences.
- Matuk, C. F., McElhaney, K., King Chen, J., Miller, D., Lim-Breitbart, J., & Linn, M. C. (2012). The Idea Manager: A tool to scaffold students in documenting, sorting, and distinguishing ideas during science inquiry. In *ICLS'12: Proceedings of the 10th international conference for the learning sciences*, Sydney, Australia, 2012. International Society of the Learning Sciences.
- Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85-139.
- McElhaney, K., Miller, D., Matuk, C., & Linn, M. (2012). Using the Idea Manager to promote coherent understanding of inquiry investigations. In *Proceedings of the 10th International Conference for the Learning Sciences*, Sydney, Australia, 2012. International Society of the Learning Sciences.
- McNeill, K. L., & Krajcik, J. (2008). Inquiry and scientific explanations: Helping students use evidence and reasoning. *Science as inquiry in the secondary setting*, 121-134.
- Moskaliuk, J. Kimmerle, J. & Cress, U. (2012) Collaborative knowledge building with wikis: The impact of redundancy and polarity, *Computers & Education*, 58(4), 1049-1057.
- Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., et al. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, 13(3), 337-386.
- Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction*, 23(1), 23-55.
- Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences*, 3(3), 265-283.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York, NY: Cambridge University Press.
- Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, 32(1), 102.
- Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31-58). New York, NY: Cambridge University Press.
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock & The American Educational Research Association (Eds.), *Handbook of research on teaching: A project of the American Educational Research Association* (pp. 315-327). New York, NY: MacMillan.
- Zhang, M., & Quintana, C. (2012). Scaffolding strategies for supporting middle school students' online inquiry processes. *Computers & Education*, 58(1), 181-196.

Framing Sociocultural Interactions to Design Equitable Learning Environments

Bryant Jensen, Ph.D., Brigham Young University
206-T MCKB Provo, UT 84602, bryant_jensen@byu.edu

Abstract: Many racial and ethnic minority (REM) students (i.e., African Americans, Hispanics, Native Americans, Pacific Islanders, and some Asian Americans) in the U.S. continue to underperform on academic tasks. Partially but not entirely explained by socioeconomic factors, REM underperformance is associated with different and inequitable sociocultural demands between informal and formal settings where REM children learn and develop. “Culturally responsive” classroom teaching is posited to help remedy this situation, but conceptual clarity and empirical support are limited. To address these limitations I recommend the learning sciences closely study cultural dimensions of social interactions across diverse classroom settings. Derived from a synthesis of research and theoretical traditions, I present an emergent framework of sociocultural interactions in classroom settings—10 dimensions organized into three domains: Life Applications, Self in Group, and Agency. I describe the constructs and address implications for design research: the need for reliable observation, and stronger theories of classroom settings/contexts.

21st Century Learning

Two important facts are important to bear in mind as researchers and practitioners continue to grapple with the purposes and goals of American Pk-12 schooling in the 21st century (e.g., Darling-Hammond, 2010; Rueda, 2013; Wagner, 2008). First, student demography is shifting toward a non-white majority. Already the case in California, New Mexico, and Texas, demographers project a non-white student majority nationwide within a couple decades (Hernandez, Denton, & Macartney, 2007). Second, the competencies demanded by employers and civic participation alike in the 21st century are expanding to include capabilities that are *intra-* and *inter-personal* in nature (Jensen, Under Review; NRC, 2012). Whereas *cognitive* skills like reading, writing, computation, and information processing are required at more complex levels than in the past (e.g., Carnegie, 2010); social (e.g., collaboration, adaptability, oral communication) and conative (e.g., self-regulation, self-efficacy) competencies are increasingly in demand (Wagner, 2008). This is because the sought-after innovations in government, business, and technology tend to be best nurtured in environments ripe in imagination, creativity, horizontal leadership, and purposeful group work (Murnane & Levy, 1996; Surowiecki, 2005).

REM Underperformance

A major challenge (and opportunity) for researchers and educators, therefore, is to find ways to meet 21st-century demands for a large and rapidly growing population of racial and ethnic minority (REM) students with a teacher workforce that is currently more than four-fifths white, non-Hispanic (U.S. Dept. of Education, 2009). Meeting the challenge means unpacking the varied explanatory factors for the pervasive REM underperformance (Lee, 2002) and designing curricula and instruction to address these factors.

Currently, the academic underperformance of racial and ethnic minority (REM) students in the US—including Blacks, Hispanics, Native Americans, Pacific Islander, and some Asian groups (Lee, 2002; Pang, Han & Pang, 2011)—is better documented than it is understood. Many REM students continue to perform lower in reading, writing, math, and other areas than their white, non-Hispanic peers. And racial/ethnic differences persist beyond the effects of socioeconomic status (Galindo, 2013; Miller, 1997, pp. 143-159). Whereas parent education, family income, occupation status, and associated factors account for a substantive portion of REM underperformance, studies demonstrate performance gaps above and beyond socioeconomic effects (Reardon & Galindo, 2009). For example, analyzing data from the Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K), Galindo (2013) found persistent math performance gaps (.36 to .65 SD) between middle-class Mexican-American and White, non-Hispanic students from kindergarten through fifth grade.

REM Developmental Assets

At the same time, studies in educational anthropology and developmental psychology identify a series of social and emotional assets nurtured in the out-of-school environments of many REM students. Research demonstrates childrearing practices in Mexican-origin families, for example, to be rooted in cultural values of respect for authority, hard work, and group solidarity (Bridges et al., 2012; Livas Stein, Garcia Coll, & Huq, 2013; Reese, 2013; Reese, Balzano, Gallimore & Goldenberg, 1995). These values tend to translate into strong development

for children, especially interpersonal competencies like cooperation and communication, as well as intrapersonal abilities like self-beliefs and a robust work ethic (Fuller & Garcia Coll, 2010; Knight & Carlo, 2012).

Researchers refer to the poor academic performance and relatively strong socioemotional competence of REM students as a “developmental paradox” (Fuller & Garcia Coll, 2010) because in mainstream society children demonstrating stronger cognitive skills also tend to demonstrate stronger socioemotional competencies. Some argue for building on REM children’s intrapersonal and interpersonal strengths as a way of enhancing their cognitive and academic competence (Galindo & Fuller, 2010; Jensen, 2013; Jensen et al., Under Review; Reese, Jensen & Ramirez, 2014). This recommendation underscores the idea that complex academic tasks like problem solving and critical thinking demand a combination of social, emotional, and cognitive skills.

Classroom Culture

Many have argued that classrooms should build on REM students’ socioemotional competencies to improve their academic performance (Crosnoe, 2006; DiPerna, Volpe & Elliott, 2005; Galindo & Fuller, 2010; Livas Stein, Garcia Coll & Huq, 2013), but little evidence demonstrates how. “Culturally responsive” classrooms for REM students propose useful approaches and offer claims, yet have not demonstrated evidence of narrowing performance gaps.

Sociocultural Learning Theory

Whereas there is no unitary theory of sociocultural learning, we situate our central arguments for greater specificity in the design of classroom improvement for REM students within the theoretical tradition of Vygotsky, Luria, and Leont’ev. Their work and corollary extensions accentuate inseparable relationships between human thought, action, and identity (John-Steiner & Mahn, 1996; Palincsar, 1998). Indeed, socioculturalists define human learning in terms of change in what the person knows and does, as well as changes in the *knower* herself (i.e., identity—“who I am”). Packer and Goicoechea (2000) make the case that sociocultural learning is marked not only by epistemological assumptions (*what* is known) but also ontological assumptions (*who* is the knower) of the world. Perhaps Dewey (1916) said it best when he described “the mind *with* the self” (p. 293, italics added). Learning, from the sociocultural perspective, comprises community practices carried out in various settings (Lave & Wenger, 1998). It builds on prior cultural knowledge and experience (Cobb & Yackel, 1996).

Importantly, “culture” from this perspective should not be thought of as a series of “individual traits”, but rather as “community practices” associated with children’s daily routines (Gutierrez & Rogoff, 2003, pp. 20-21). Culture does not consist of a fixed and static set of rituals, beliefs, and traits; to conceptualize students’ culture this way runs the risk of reinforcing simplistic group stereotypes (Irizarry, 2007; Reese, 2013). This way culture is not synonymous with race, ethnicity, social class or any other group label. We view culture as both produced and reproduced in daily interactions within diverse settings, rather than necessarily a function of group difference. This conception allows for changes in culture over time and for within-group variation, in response to changing conditions (Nasir & Hand, 2006).

Social interactions—i.e., the behaviors or actions that occur between two or more persons in response to one another—are a central analytic unit in sociocultural theory (Vygotsky, 1978, pp. 24-30). Culturally bound rules, norms, values, and expectations shape how we relate, communicate, question, assist, and generally engage with one another. Sociocultural interactions underlie our learning and development. Indeed, according to Vygotsky (1978), differences between a child’s “actual developmental level as determined by independent problem solving” and her potential level are “determined through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky, 1978, p. 86). We learn through prompts, clues, modeling, conversation, observation, imitation, joint participation, encouragement, etc.

Culturally Responsive Teaching

Literature on “culturally responsive pedagogy”—connecting student knowledge and experience from non-school to classroom settings (Bowers & Flinders, 1990; Gay, 2000; Irizarry, 2007; Stuart & Volk, 2002; Wortham & Contreras, 2002)—builds loosely on sociocultural learning theory. It also draws on a social justice (Brown-Jessy & Cooper, 2011) and socioaffective (Ladson-Billings, 1995) frameworks. Culturally responsive teaching includes explicit (e.g., content personalization) and implicit (e.g., social motivations) efforts to connect students’ out-of-school lives with classroom experiences.

Culturally responsive pedagogy is more than generic attributes of “good teaching” (Ladson-Billings, 1995). It is important, thus, to distinguish universal (or global) from cultural (or local) dimensions of classroom interactions (Tharp, 1989). Global dimensions demonstrate positive effects on student development and learning in general, whereas local dimensions tend to favor some students’ over others within the same classroom. Typically the curricular content and forms of social interaction in school settings favor White and middle to

upper-middle class children over underperforming REM students. Thus the culture of classroom learning is hypothesized to be more “responsive” for the already privileged than for REM students.

Two issues undermine the uptake and influence of culturally responsive practices in REM classrooms. First, little inferential evidence exists to demonstrate how—i.e., casual pathways—they influence student academic learning (Goldenberg, Rueda & August, 2006). Most research on culturally responsive teaching—and sociocultural studies in classrooms more generally—are interpretive and non-interventionist. Without associating sociocultural practices with student learning, however operationalized, it is difficult to ascertain instructional principles that could be replicated and tested in across classroom settings. Second, there is a lot of jargon and term confusion in the literature on culturally responsive teaching and related sociocultural studies. Semantic deviations can undermine a knowledge base for subsequent study and improvement.

Global and Local Dimensions of Classroom Interactions

A significant hurdle to expand the knowledge base on culturally responsive teaching is to disentangle the cultural (or local) from universal (or global) nature of classroom interactions (i.e., what Tharp [1989] termed “psychocultural variables” and “constants”, respectively). These distinctions are necessary to understand how to fit academic content with REM students’ out-of-school knowledge and experiences. One of the problems with sociocultural research in education generally is the ambiguity between quality (i.e., universal action for all students) and cultural practices that inadvertently privilege certain students over others (Goldenberg & Gallimore, 1989). Literature on culturally responsive pedagogy (e.g., Bowers & Flinders, 1990; Gay, 2000; Ladson-Billings, 1995) often presents local dimensions of classroom quality as different from or somehow indifferent to global dimensions.

Colleagues and I have argued that what is needed is a clearer articulation of how local dimensions are *instantiated* by global dimensions of classroom interactions (Reese, Jensen & Ramirez, 2014). In general, global dimensions address *whether* and to *what extent* quality interactions occur to spur student learning. Global dimensions include affective quality of teacher-student interactions, classroom management, and generic features (e.g., feedback, scaffolding, analytic discussion) of cognitive support during classroom instruction (Hamre & Pianta, 2001, 2005). Local dimensions address *how* that same interaction is socialized to draw on students’ previous knowledge, experience, routines, values, and traditions.

We make three points to elucidate differences between global and local dimensions of classroom interactions (Jensen et al., Under Review). Each highlights the centrality of classroom context to make appropriate distinctions. First, the global vs. local nature of classroom interactions depends on the complexity (e.g., memory vs. inference) of learning objectives. Cooperation (a local dimension below), for example, can be considered a feature of global quality when teaching complex scientific concepts (Sinatra & Chinn, 2011).

Second, dimensions of classroom interactions can be considered local in terms of the differential degree of their influence (to student learning and development). Some work has found nurturing student choice and autonomy in classroom discussion to be associated with stronger language and reading gains for native Hawaiian children than for white, middle-class students in the same classrooms (Au & Mason, 1981).

Third, dimensions of classroom interactions are considered “local” in terms of setting differences that are functionally qualitative. That is, meaning from social interactions is made not exclusively in terms of “the extent to which” classrooms help students make personal connections through culturally responsive practice, but also “how” practices are contextually enacted—i.e., qualitative differences in how teachers motivate and socialize children. These differences are analytically inductive rather than deductive and have significant implications for designing culturally responsive environments. Thus we assert that classroom dimensions considered “local” are contextual and should incorporate qualitative information for valid interpretation.

Our contention is that the amalgamation of local and global dimensions of classroom interactions matters to relevant learning for REM students. Classrooms deemed “high quality” are not necessarily responsive, and vice versa. Whereas REM students can certainly gain understanding of academic content through universal quality alone, understanding is enhanced and intrinsic value for academic learning are greatly increased, in theory, when classroom interactions are also responsive to students’ out-of-school lives (Brophy, 1999). On the other hand, classroom interactions that are highly responsive yet low in quality are characterized as maintaining some student interest but little academic learning.

Sociocultural Interactions in Classrooms

Shown in Figure 1, we identify ten “local” dimensions of sociocultural interactions in Pk-6 classroom settings, and organize them into three abstract domains: Life Applications, Self in Group, and Agency. We draw these constructs from a variety of fields, including cultural psychology, educational anthropology, developmental psychology, social psychology, sociolinguistics, communication, and multicultural education. The constructs themselves are not new, but the way we organize them into a single model sets the stage for a host of activities to design more relevant learning opportunities for underperforming REM students. Described at length below, our model provides a common schematic and nomenclature.

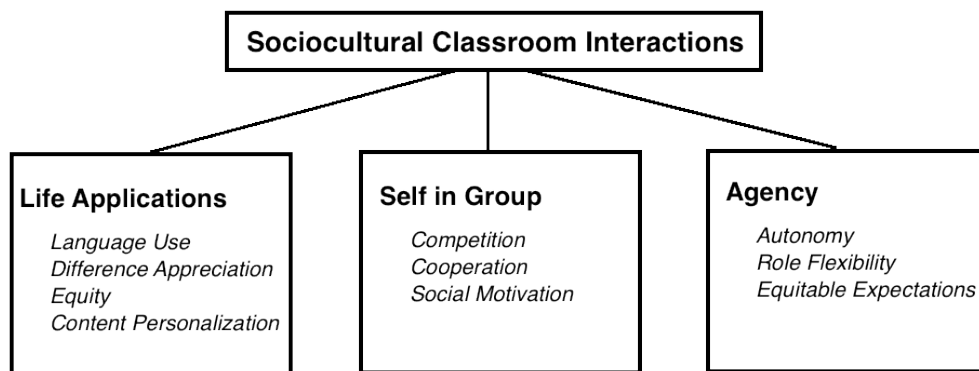


Figure 1. Model of sociocultural interactions in classroom settings

We established a priori criteria to construct inclusion for the model. Namely, we were interested in capturing constructs that a) have received attention in the research literature, b) are at least purported to be associated with children’s socioemotional and/or cognitive development, c) are detectable to the trained observer, d) are commonplace enough to be observed within most hour-long classroom observations, and e) demonstrate malleability for subsequent improvement.

Life Applications

The first domain addresses whether and how classroom interactions explore and value students’ interests, beliefs, knowledge, and experiences in order to make personal connections with classroom content (Aguirre, 1988; Cazden, 2001; Gee, 2001; Gonzalez, Moll & Amanti, 2005; Wertsch & Toma, 1995). Interactions are scaled on a continuum from “disconnected” to “well connected.” Teachers in “well connected” classrooms demonstrate commitment to learn about students’ lives, and they search for ways to incorporate what they learn. Dimensions include: Language Use, Difference Appreciation, Equity, and Content Personalization.

Language Use refers to the extent to which classroom interactions gauge and incorporate the natal, non-school languages (varieties or systems) of students to enhance content understanding and social relations (Heath, 1983). *Varieties* refer to the organization of language (e.g., style, structure, tone, vernacular) whereas *language system* refers to the actual language code (e.g., Portuguese, Korean, Spanish). Connecting classroom learning objectives with student language repertoires requires teachers to “code-switch” (Aguirre, 1988; Gumperz, 1982; Hakuta & Garcia, 1989; Valdes-Fallis, 1978) between the school and the non-school languages. Switching can be relational (Garcia, 2005, pp. 27-29) and instructional (Cazden, 2001).

Difference Appreciation refers to the extent to which teachers and peers value and address the diverse experiences, knowledge, beliefs, and interests of students in the classroom. “Well-connected” classrooms frequently discuss the out-of-school hobbies, activities, social roles, responsibilities, traditions, and peer and family relationships of children. This way, students learn more about one another, thereby developing their “cultural competence” (Ladson-Billings, 1995, pp. 160-161). They become more aware of their classmates’ differences, and appreciate them rather than overlook, dismiss, or look down on them.

Equity addresses how classroom interactions address societal injustices (past and present) associated with student differences. “Well-connected” classroom interactions in terms of Equity develop teacher and student consciousness of past and present prejudice and discrimination—real or perceived—in order to imagine social transformations (Freire, 1970). Equitable classroom interactions promote social justice and, according to Ogbu (1987, 1992), are especially important for “involuntary” minorities—i.e., descendants of groups of persons thrust into a new society “against their will” (Ogbu & Simons, 1998, p. 165)—who identify with historical oppression (e.g., African American slavery, Native American subjugation).

Content Personalization refers the connections made between students’ lives—routines, perspectives, social relationships, expertise, values, and traditions—and classroom learning objectives. Tharp and colleagues (2000) refer to Content Personalization as “making meaning” (p. 26-29). They argue that children “are willing to struggle with [...] abstract notions in science, math, and other content areas when they are motivated by [...] activities they and their families value” (p. 26). To do so, teachers orchestrate classroom interactions that “are situated in problems and issues of students’ everyday lives [and] provide vivid opportunities [for] students to stretch their informal understandings to more abstract levels” (p. 26).

Self in Group

Self in Group, or “self-construal” (Kim, 2002), refers to how classroom interactions socialize students to relate to and work with one another to motivate learning and establish identities (Au & Mason, 1981; Greenfield,

1994). Do classroom conversations, for example, nurture a communal climate where team success is more significant than individual performance (Kim et al., 1996), or a setting in which the learning and achievement of students are at odds with one another? We scale scores dimensions of this domain from “independent” to “interdependent.” Three dimensions of Self in Group include Competition, Cooperation, and Social Motivation.

Competition refers to the extent to which student success in the classroom depends on the failure of another (individuals or groups). In competitive classrooms, students strive to perform well simply to outperform their peers. Doing well in relation to one’s peers is more important than understanding content (McInerney, Roche, McInerney & Marsh, 1997). Competitive classroom interactions are scaled as “independent.”

Cooperation is defined as collaborative student effort to achieve a shared objective. Shared work, common goals, and social cohesion are indicators of Cooperation. Extant research on “cooperative learning” (Slavin, 2010) has found increased student learning outcomes, though some students benefit more than others (Stevens & Slavin, 1995). Webb & Farivar (1994), for example, found that Latino and African American middle school students demonstrated greater mathematics gains and classroom participation from academic skill training in cooperative groups than their white, non-Hispanic peers. Depending on the cultural norms and community practices, students may require close scaffolding to foster the pro-social skills (e.g., listening, turn-taking, assertion) required to cooperate.

Social Motivation refers to how classroom interactions incentivize student effort and participation. Are motivations oriented toward independent or interdependent effort and accomplishment? The idea is that social interactions and interpersonal relationships are inherently motivating for student engagement and learning (Martin & Dowson, 2009). But some rewards are more authentic than others, and cultural values across communities vary with regard to social cohesion.

Agency

The last domain refers to how student choice and freedom are managed (Holland et al., 1998). It addresses how active the classroom environment allows students to be—to exercise choice, undertake responsibility, take on different social roles, and internalize learning expectations. This domain is concerned with who makes the instructional, curricular, and organizational decisions in the classroom and how those decisions are enacted. Agency dimensions include Autonomy, Role Flexibility, and Equitable Expectations, scaled from “not distributed” to “well distributed.”

Autonomy addresses how much choice children are granted to select materials and tasks, monitor their learning progress, collaborate with peers, engage in relevant conversation, etc. It refers to an “internal perceived locus of causality” (deCharms, 1968). It does *not* refer to any ideal toward group activity (Chirkov, Ryan, Kim & Kaplan, 2003), though Western perspectives often associate autonomy with independence (Deci & Ryan, 2008). Autonomy, however, “refers not to being independent, detached, or selfish but rather to the feeling of volition than can accompany any act, whether [...] collectivist or individualist” (Ryan & Deci, 2000, p. 74).

Role Flexibility refers to how rigid or malleable roles are between experts (e.g., teachers) and novices (e.g., students) to complete a task. Cultural psychologists find role flexibility to vary across communities (Paradise & Rogoff, 2009; Rogoff et al., 2007). In some communities, roles are rigid—divided and assigned by the expert. In others, roles are more fluid: though with limited experience, the novice attempts to complete a task with guidance from the expert (Rogoff, 1990). “Distributed” agency is associated with flexible roles.

Equitable Expectations are defined as the enacted beliefs that teachers have about their students’ learning aptitude in the classroom. These beliefs—developed and communicated mostly through subtle, indirect, and ongoing interactions (Good & Brophy, 2008, pp. 54-57)—are associated with the distribution of student agency. That is, teachers with equitable expectations induce autonomy for *all* students through complex tasks, opportunities for self-evaluation, individualized feedback, and shared respect (Good & Weinstein, 1986).

A Stronger Sociocultural Design Science

We encourage further study to apply attributes of design research (van den Akker et al., 2006) to study (and improve) these sociocultural dimensions in classroom settings. Design research provides opportunities not only for establishing a broader base of empirical evidence regarding the influence of Life Applications, Self in Group, and Agency in classrooms, but also information on when these features matter, for whom, and for what purposes. It does so by focusing on contextualized problems of practice through iterative study, cultivating close collaborations between practitioners and researchers, blending multiple research methods as needed, and providing “fine-grained” (diSessa & Cobb, 2004, p. 89) theoretical advances.

Two major advances are needed in the field to establish a stronger design science on associations between REM learning and sociocultural interactions in classroom settings. First, we need a reliable observational system to capture “local” dimensions of classroom quality. This system should demonstrate internal consistency, face and construct validity, a strong, interpretable factor structure. This task is certainly daunting, as it requires capturing historically interpretivist constructs within a psychometric paradigm.

Thoughtful efforts should capture context-sensitive markers for the dimensions described, yet clear enough scoring rubrics for raters to reliably observe them across diverse classroom settings and cultural traditions.

Second, by extension, stronger conceptions of classroom contexts associated with “cultural responsive” practice are needed (i.e., ecological validity). Whereas research findings associating sociocultural interactions and REM student learning would generalize at the classroom or setting level, we need stronger schematics of context attributes (i.e., covariates) that matter to these relationships. Said schematics are likely to include attributes of the academic task (e.g., developmental demand, artifacts/tools), timing (e.g., duration, time of day, day of week), rules (e.g., participation structure, values) and, of course, participants—fixed qualities like race, age, and gender, as well as malleable qualities like identity, beliefs, and dispositions. To add further complexity, setting attributes of the classrooms are necessarily interwoven with characteristics at the intuitional (programs, policies, dominant beliefs) and community (socioeconomic status, racial/ethnic integration) levels that constrain and afford culturally responsive practice (Cobb, McClain, de Silvia Lamberg & Dean, 2003).

References

- Aguirre, A. (1988). Code switching, intuitive knowledge, and the bilingual classroom. In H. S. Garcia and R. Chavez (Eds.), *Ethnolinguistic issues in education*. Lubbock, TX: Texas Tech University.
- Au, K., & Mason, J. (1981). Social organization in learning to read: The balance of rights hypothesis. *Reading Research Quarterly*, 17, 115-152.
- Bowers, C. A., & Flinders, D. J. (1990). *Responsive teaching: An ecological approach to classroom patterns of language, culture, and thought*. New York: Teachers College Press.
- Bridges, M., Cohen, S. R., McGuire, L. W., Yamada, H., Fuller, B., Mireles, L., & Scott, L. (2012). Measuring the social behavior of Mexican American children. *Early Childhood Research Quarterly*, 27, 555-567.
- Brophy, J. (1999). Toward a model of the value aspects of motivation in education: Developing appreciation for particular learning domains and activities. *Educational Psychologist*, 34(2), 75-85.
- Brown-Jessy, S. & Cooper, J.E. (2011). Toward a conceptual framework of culturally relevant pedagogy: An overview of the conceptual and theoretical literature. *Teacher Education Quarterly*, 38(1), 65-84.
- Carnegie Council on Advancing Adolescent Literacy. (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success*. New York, NY: Carnegie Corporation of New York.
- Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinemann.
- Chirkov, V. Ryan, R. M., Kim, Y., & Kaplan, U. (2003). Differentiating autonomy from individualism and independence: A self-determination theory perspective on internalization of cultural orientations and wellbeing. *Journal of Personality and Social Psychology*, 84, 97-110.
- Cobb, P., McClain, K., de Silvia Lamberg, T., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district, *Educational Researcher*, 32(6), 13-24.
- Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context and developmental research. *Educational Psychologist*, 31(3/4), 175-190.
- Crosnoe, R. (2006). *Mexican roots, American schools: Helping Mexican immigrant children succeed*. Palo Alto, CA: Stanford University Press
- Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. New York: Teachers College Press.
- deCharms, R. (1968). *Personal causation*. New York: Academic Press.
- Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology*, 49(1), 14-23.
- Dewey, J. (1916). *Democracy and education*. The Macmillan Company.
- DiPerna, J. C., Volpe, R. J., & Elliott, S. N. (2005). A model of academic enablers and mathematics achievement in the elementary grades. *Journal of School Psychology*, 43, 379-392
- diSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *Journal of the Learning Sciences*, 13(1), 77-103.
- Freire, P. (1970). *Pedagogy of the oppressed*. New York: Continuum.
- Fuller, B., & Garcia-Coll, C. (2010). Learning from Latinos: Contexts, families, and child development in motion. *Developmental Psychology*, 46(3), 559-565.
- Galindo, C. (2013). Math performance of young Mexican-origin children in the United States: Socioeconomic status, immigrant generation, and English proficiency. In B. Jensen and A. Sawyer (Eds.), *Regarding Educación: Mexican-American Schooling, Immigration, and Bi-national Improvement*. New York: Teachers College Press.
- Galindo, C., & Fuller, B. (2010). The social competence of Latino kindergartners and growth in mathematical understanding. *Developmental Psychology*, 46(3), 579-592.
- Garcia, E. (2005). *Student cultural diversity: Understanding and meeting the challenge*, 3rd ed. Boston, MA: Houghton Mifflin.

- Gay, G. (2000). *Culturally responsive teaching: Theory, research, & practice*. New York: Teachers College.
- Gee, J. P. (1989). What is literacy? *Journal of Education*, 171, 18-25.
- Goldenberg, C. & Gallimore, R. (1989). Teaching California's diverse student population: The common ground between educational and cultural research. *California Public Schools Forum*, 3, 41-56.
- Goldenberg, C., Rueda, R., & August, D. (2006). Sociocultural influences on the literacy attainment of language-minority children and youth. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report to the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum.
- Gonzalez, N., Moll, L., & Amanti, C. (2005). *Funds of knowledge: Theorizing practices, households, communities, and classrooms*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Good, T. L., & Brophy, J. (2008). *Looking into classrooms (10th ed.)*. Boston, MA: Allyn & Bacon.
- Good, T.L., & Weinstein, R. (1986). Teacher expectations: A framework for exploring classrooms. In K. Zumwalt (Ed.), *Improving Teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Greenfield, P. M. (1994). Independence and interdependence as developmental scripts: Implications for theory, research, and practice. In P. M. Greenfield & R. R. Cocking (Eds.), *Cross-cultural roots of minority child development*. Hillsdale, NJ: Erlbaum.
- Gumperz, J. J., (1982). *Discourse strategies*. Cambridge, UK: Cambridge University Press.
- Gutierrez, K., & Rogoff, B. (2003). Cultural ways of learning: Individual traits and repertoires of practice. *Educational Researcher*, 32(5), 19-25.
- Hakuta, K., & Garcia, E. (1989). Bilingualism and education. *American Psychologist*, 44(2), 374-379.
- Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, 72, 625-638.
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first grade classroom make a difference for children at risk of school failure? *Child Development*, 76(5), 949-967.
- Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. New York: Cambridge University Press.
- Hernandez, D. J., Denton, N. A. & Macartney, S. E. (2007). Young Hispanic children in the 21st Century. *Journal of Latinos and Education*, 6(3), 209-28.
- Holland, D., Lachicotte, W., Skinner, D., & Cain, C. (1998). *Identity and agency in cultural worlds*. Cambridge, MA: Harvard University Press.
- Irizarry, J.G. (2007). Ethnic and urban intersections in the classroom: Latino students, hybrid identities, and culturally responsive pedagogy. *Multicultural Perspectives*, 9(3), 21-28.
- Jensen, B. (2013). Finding synergy to improve learning opportunities for Mexican-origin students. In B. Jensen & A. Sawyer (Eds.), *Regarding educación: Mexican American schooling, immigration, and binational improvement*. New York: Teachers College Press.
- Jensen, B. (Under Review). Developmental-capabilities pedagogy to ameliorate the Mexican paradox.
- Jensen, B., Rueda, R., Reese, L., & Garcia, E. (Under review). Designing sociocultural interactions to improve relevant learning opportunities for underperforming minority students.
- John-Steiner, V., & Mahn, H. (1996). Sociocultural approaches to learning and development: A Vygotskian approach. *Educational Psychologist*, 31(3/4), 191-206.
- Kim, M. S. (2002). *Non-western perspectives on human communication: Implications for theory and practice*. Thousand Oaks, CA: Sage.
- Knight, G. P., Kagan, S., & Buriel, R. (1982). Perceived parental practices and prosocial development. *Journal of Genetic Psychology*, 141, 57-65.
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Journal of Educational Research*, 32(3), 465-491.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge.
- Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher*, 31(1), 3-12.
- Livas Stein, G., García Coll, C., & Huq, N. (2013). Fostering resilience in Mexican American youth through cultural and family assets. In B. Jensen & A. Sawyer (Eds.), *Regarding Educación: Mexican-American Schooling, Immigration, and Bi-national Improvement*. New York: Teachers College Press.
- Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and educational practice. *Review of Educational Research*, 79(1), 327-365.
- McInerney, D. M., Roche, L. A., McInerney, V. & Marsh, H. W. (1997). Cultural perspectives on school motivation: The relevance and application of goal theory. *American Educational Research Journal*, 34(1), 207-236.
- Miller, L. S. (1997). *An American imperative: Accelerating minority educational advancement*. New Haven,

- CT: Yale University Press.
- Murnane, R., & Levy, F. (1996). *Teaching the new basic skills*. New York: The Free Press.
- Nasir, N. S., & Hand, V. M. (2006). Exploring sociocultural perspectives on race, culture, and learning. *Review of Educational Research*, 76(4), 449-475.
- National Research Council (2012). *Education for life and work: Developing transferrable knowledge and skills in the 21st century*. Washington, DC: National Academies Press.
- Ogbu, J. (1987). Variability in minority school performance: A problem in search of an explanation. *Anthropology & Education Quarterly*, 18(4), 312-334.
- Ogbu, J. (1992). Understanding cultural diversity and learning. *Educational Researcher*, 21(8), 5-24.
- Ogbu, J., & Simons, H. (1998). Voluntary and involuntary minorities: A cultural-ecological theory of school performance with some implications for education. *Anthropology & Education Quarterly*, 29, 155-188.
- Pang, V. O., Han, P. H., & Pang, J. M. (2011). Asian American and Pacific Islander students: Equity and the achievement gap. *Educational Researcher*, 40(8), 378-389.
- Paradise, R., & Rogoff, B. (2009). Side by side: Learning by observing and pitching it. *Ethos*, 37(1), 102-138.
- Reese, L. (2013). Cultural change and continuity in U.S. and Mexican settings. In B. Jensen & A. Sawyer (Eds.), *Regarding educación: Mexican American schooling, immigration, and binational improvement*. New York: Teachers College Press.
- Reese, L., Balzano, S., Gallimore, R. and Goldenberg, C. (1995) The concept of “educación”: Latino family values and American schooling. *International Journal of Education Research*, 23 (1), 57-81.
- Reese, L., Jensen, B., & Ramirez, D. (2014). Emotionally supportive classroom contexts for young Latino children in rural California. *Elementary School Journal*, 114(4)
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Rogoff, B. et al. (2007). Children’s development of cultural repertoires through participation in everyday routines and practices. In J. E. Grusec and P. D. Hastings (Eds.), *Handbook of Socialization: Theory and Research*. New York: Guilford.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78.
- Rueda, R. (2013). Twenty first century skills: Cultural, linguistic, and motivational perspectives. In D. Alverman & N. Unrau, (Eds.), *Theoretical Models & Processes of Reading*. Newark, DE: International Reading Association.
- Sinatra, G. M., & Chinn, C. (2011). Thinking and reasoning in science: Promoting epistemic conceptual change. In K. Harris, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *Educational psychology: Contributions to education*. Washington, DC: American Psychological Association.
- Slavin, R. E. (2010). Instruction based on cooperative learning. In R. E. Mayer and P. A. Alexander (Eds.), *Handbook of Research on Learning and Instruction*. New York: Routledge.
- Stevens, R. J., & Slavin, R. E. (1995). The cooperative elementary school: Effects on students’ achievement, attitudes, and social relations. *American Educational Research Journal*, 32(2), 321-351.
- Stuart, D. and Volk, D. (2002). Collaboration in a culturally responsive literacy pedagogy: Educating teachers and Latino children. *Reading Literacy and Language*, 36(3), 127-134.
- Surowiecki, J. (2005). *The wisdom of crowds*. New York: Anchor Books.
- Tharp, R. (1989). Psychocultural variables and constants: Effects on teaching and learning in schools. *American Psychologist*, 44(2), 349-359.
- Tharp, R., Estrada, P., Dalton, S., & Yamauchi, L. A. (2000). *Teaching transformed: Achieving excellence, fairness, inclusion, and harmony*. Boulder, CO: Westview Press.
- U.S. Department of Education (2009). *Our future, our teachers*. Washington, DC: U.S. DOE.
- Valdes-Fallis, G. (1978). Code-switching among bilingual Mexican-American women: Towards an understanding of sex-related language alternation. *International Journal of the Sociology of Language*, 17, 65-72.
- van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.) (2006). *Educational design research*. New York: Routledge.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Wagner, T. (2008). *The global achievement gap*. New York: Basic Books.
- 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.
- Wertsch, J. V., & Toma, C. (1995). Discourse and learning in the classroom: A sociocultural approach. In L. Steffe & J. Gale (Eds.), *Constructivism in education*. Hillsdale, NJ: Lawrence Erlbaum.
- Wortham, S., & Contreras, M. (2002). Struggling towards culturally relevant pedagogy in the Latino diaspora. *Journal of Latinos and Education*, 1(2), 133-144.

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

Raymond Kang, University of Illinois at Chicago, 1240 W Harrison St, Ste. 1570, rkang2@uic.edu

Abstract: Within schools, data-driven practices have served as the centerpiece for educational reform initiatives focused on improving instruction and student learning, but this process has largely remained insufficiently conceptualized. This case study follows an Instructional Leadership Team (ILT) in the Chicago Public Schools (CPS) during the 2012-2013 school year as they enacted the data-driven practice of Instructional Rounds (IR). Taking a Cultural-Historical Activity Theory with Institutional Theory framework (CHAT-IT), a discursive analysis of ILT meetings displays the seeming resolution of manifestations of cultural-historical contradictions within the ILT's enactment of the IR practice. Yet, further analysis of and interviews with individual members troubles these resolutions, particularly when seeing IR as a coercive, mimetic, and normative mechanism. As one of few studies adopting the CHAT-IT framework, this case study supports the fruitful, dialogical understanding of organizational learning as culturally and historically situated activity.

Introduction

As an increasingly popular approach to improving teaching and student learning, teachers' use of data has become central to public policy discourse for educational reform. In their review of the literature, Young and Kim (2010) assert that the "importance of using 'data' is now taken-for-granted as an essential strategy for educational improvement" (p. 3). Despite or, perhaps, due to the assumption of its importance, numerous writers contend that the connections between data and instructional practice remain insufficiently conceptualized (Spillane, 2012; Cosner, 2012). This assumption further stems from other problematic assumptions, particularly the notion that data, per se, offers clear guidance to practice as opposed to understanding how data use is situated within activity systems (Spillane, 2012). Essentially, data-driven practices cannot be viewed as an unequivocal good but must, instead, be viewed as an activity system comprising multiple, inherent contradictions that must be resolved if an organization is to learn (Engeström, 1999).

Ogawa and colleagues (2008) put forth a theoretical framework that draws upon conceptual tools from cultural-historical activity theory with institutional theory (CHAT-IT) for understanding organizational learning. Understanding the nature of change within any activity system begins with locating the manifestations of *contradictions* within its enactment, following their "dynamic movement and evolution" over time (Engeström & Sannino, 2011, p. 384). This study aims to adopt the CHAT-IT framework to guide the analysis of an Instructional Leadership Team (ILT) as they work to design and implement the data-based practice of Instructional Rounds (IR). IR involve an iterative cycle of four steps: (1) identifying a problem of instructional practice; (2) observing practice in small groups; (3) debriefing together after the observation; and (4) focusing on the next level of work needed to improve instruction relative to the problem of practice (City, Elmore, Fiarman & Teitel, 2011). Within this context, our specific research questions are:

1. How and what does the ILT learn from engaging in the practice of IR?
2. What contradictions manifest in the activity of IR, and how are they (if they are) resolved?

Theoretical Framework

The CHAT-IT framework guides the analysis reported here. In merging CHAT and IT, this framework takes on the assumptions of those theoretical frameworks. By adopting CHAT, we assume that the individual actions and social structures are mutually constitutive, resulting in an activity system that is both culturally and historically situated. Due to this situativity, activity systems also inherently contain *contradictions*, and activity must go through cycles of internal, critical self-reflection and external searches for solutions to become expansive, or, to learn (Engeström, 1999). Moreover, this assumes that the contradictions of activity systems must be "creatively and painfully resolved by working out a qualitatively new 'thirdness', something qualitatively different from a mere combination or compromise between two competing forces" (Engeström & Sannino, 2011, p. 371). Organizational learning thus becomes an activity in which contradictions manifest, and attempts to resolve these manifestations ideally lead to a qualitatively transformed organization.

Any activity system comprises multiple, mutually constitutive elements. Ogawa et al. (2008) adopt the model first put forward by Engeström (1999), which describes six elements: *subject*, *object*, *mediating artifacts*, *community*, *rules*, and *division of labor*. Within this model, the *object* of an activity system is the purpose that organizes individual actions and connects them to the collective activity, leading to sought after and unforeseen outcomes. The *subject* of the activity system can be either individuals or groups of individuals that are more or less organized. CHAT acknowledges that the subject of an activity system exists within dialectical relationships

with all other elements of the activity system, where the subject both influences and is influenced by the other elements. *Mediating artifacts* enable this dialectical relationship, through both physical tools and symbolic signs. These mediating artifacts afford and constrain activity, and the creation or appropriation of new mediating artifacts allow for subjects to shape their activity.

These three elements form the core of activity, but CHAT expands upon this model by incorporating elements of *community*, *rules* and *division of labor* in its model. The subject-artifact-object relationships occur within *communities*, which are defined by collective social and cultural practices that organize these dialectical relationships (Lave & Wenger, 1991). These practices become enacted and embodied through *rules* and *division of labor*. Ogawa et al. (2008) see *rules* as what “embody and reproduce ideologies in the broader societal context,” constraining and affording participation within the activity system (p. 87). Similarly, *division of labor* legitimates certain forms of participation based upon the subject’s culturally and historically constrained position within the activity system relative to the object. In other words, the central or peripheral nature of a subject’s participation as well as its legitimacy stem from the subject’s dialectical relationships with the other elements of the activity system. Critically, all of these elements dialectically relate to each other ultimately resulting in an *outcome* of activity

With this necessarily brief overview of activity systems, Ogawa et al. (2008) add an additional layer of mechanisms that carry *institutions*, “the systematic expression of dominant cultural values” into formal organizations such as schools (p. 89). By expanding this dialectical view from CHAT to include how institutions sustain and legitimate the practices of organizations, they argue that how organizations learn must include acknowledgement of the *coercive*, *mimetic*, and *normative* mechanisms that enact and carry institutions into activity. *Coercion* includes the formalized rules and regulations of the state. Within the context of schooling, this can be seen embodied within accountability and threats of school turnaround or closure (Smylie, 2010). *Mimicry* occurs when organizations attempt to adopt practices and structures that are perceived as key to success within their field. This results in “isomorphic behaviors” where organizations within a field begin to increasingly resemble one another (Smylie, 2010, p. 32). *Norms* include “professional codes” as well as organizational values that afford and constrain actions deemed as acceptable within an activity system (Ogawa et al., 2008, p. 89).

Ogawa et al. (2008) argue that “CHAT and IT identify conceptually similar elements of social systems, which, taken together, link the immediate settings in which learning occurs to the social contexts of organizations and to the social and cultural forces that shape organizations” (p. 84). A dialectical view of organizational change, or learning, “offers a way to parse, in depth, the multiple contributing pieces that give shape to any particular activity” (Ogawa et al., 2008, p. 87). Following the “dynamic movement” of an activity system within the formal organization of schooling must consider the potential affects of these institutional carriers. Critical to this conception is the dialectical nature of these relationships. Any account of organizational learning with CHAT-IT must account for how these carriers affect and, in turn, are affected by the activity system of a school.

Methodology

Methodological Framework

While organizational learning can conceptually be understood as the resolution of contradictions within an activity system, we have no direct access to contradictions and must instead attend to their manifestations, particularly within the discourse of practitioners. Engeström and Sannino (2011) presented the evolution of the discourse of individuals and groups as “windows into systemic contradictions in the activity system” (p. 385), suggesting a potentially fruitful methodology for exploring the dialectics inherent within the work of the ILT. In their preliminary framework, they discuss four types of discursive manifestations of contradictions: *conflicts*, *dilemmas*, *crucial conflicts*, and *double binds*.

Within their guiding framework, the authors expanded upon four discursive manifestations of contradictions: (1) *dilemmas*, or “an expression or exchange of incompatible evaluations, either between people or within the discourse of a single person,” (p. 373); (2) *conflicts*, which “take the form of resistance, disagreement, argument and criticism,” (p. 373); (3) *critical conflicts*, which are the “situations in which people face inner doubts that paralyze them in front of contradictory motives unsolvable by the subject alone,” (p. 374); and (4) *double binds*, or “processes in which actors repeatedly face pressing and equally unacceptable alternatives in their activity system, with seemingly no way out” (p. 374).

In their analysis of the management group for the city of Helsinki’s municipal home care for the elderly, Engeström and Sannino (2011) linked rudimentary linguistic cues from discourse to these manifestations of contradictions. For example, *dilemmas* were characterized by acknowledgements of competing statements (e.g., “yes, but”; “on the one hand ... on the other”). *Conflicts* were characterized by disagreements in the discourse, linked to cues such as “I disagree” or “this I can accept.” *Critical conflicts* often presented as personal accounts and/or the use of vivid metaphors (e.g., “I now realize that...”). Finally, *double*

binds manifested in rhetorical questions or expressions of helplessness (e.g., “we must”). Further guidance for locating contradictions within activity systems comes from an earlier work by Engeström (1999), where he located contradictions between the central components of the activity system (i.e., subject, object, mediating artifacts, rules, community, and division of labor). Adopting this guiding framework, this report attempts to locate manifestations of contradictions within the discourse of the ILT in an effort to further understand what contradictions it needed to resolve in order to transform its activity and, thus, learn as an organization.

Data Collection and Analysis

To examine the potential of the theoretical and methodological frameworks for gaining insight into data-use practices in schools, the author participated in a single-site case study during the 2012-2013 school year in the Chicago Public Schools (CPS), a large urban school district in the Midwest United States. The ILT of Douglas M. Wilson High School (Wilson; all names including the school’s are pseudonyms) was followed for the course of the 2012-2013 school year. During the course of observation, Wilson enrolled approximately 1500 students, 85% of whom were low-income and 17% were of limited English proficiency.

The specific organizational practice studied was that of the meetings of the ILT that took place once every two weeks, particularly as they attempted to design and implement the new routine of IR for the school. The ILT comprised the department chairs of the school (Math, Science, Social Studies, Fine Arts, English, Foreign Language, Special Education); heads of other organizations in the school (Service Learning, Junior ROTC); administrators (primarily the principal and one assistant principal); an Instructional Support Lead (ISL) from the Network, a partition of CPS that included Wilson; and, at times, would be joined by members from the counseling department. A typical meeting would be organized by an agenda established during the week that the ILT did not meet, and all members of the ILT were able to place items on the agenda for discussion. Topics of discussion primarily focused on instructional matters, including implementation of IR (which constitutes the bulk of this analysis), but also including discussion of school metrics (e.g., freshman on-track rate, attendance rates), reorganizing the utilization of the schools facilities, and student test-scores on EPAS (EXPLORE, PLAN, ACT system). Meetings lasted for one class period (50 minutes) and usually discussions took the entire time.

Analysis occurred through an examination of four data sources: (1) transcripts from audio-recordings of 16 ILT meetings that occurred throughout the 2012-2013 school year and three IR Subcommittee meetings held outside of regular ILT meeting time; (2) transcripts from audio-recordings of semi-structured interviews (one to four interviews per informant) with fifteen individuals comprising members of the ILT as well as other district personnel (i.e., the consultant hired to train this school and others in the practice of instructional rounds as well as a district official overseeing a subset of district schools that included Wilson) that occurred one to four times over the course of the 2012-2013 school year; (3) documents or instruments produced or used by the ILT in their practice, including but not limited to their implementation of IR (e.g., agendas for meetings, evaluation forms, test score reports); and (4) field notes and memos created by the author.

These data sources were triangulated in order to create an account of the activity of IR over time. Focusing on the activity of IR constrained the data. While many discussions about data occurred within the ILT (e.g., student test scores), this analysis focuses on the use of data collected through IR. Excerpts from transcripts of ILT and IR Subcommittee meetings were first identified using Engeström and Sannino’s (2011) framework, and then further coded using a grounded methodology (Charmaz, 2006). As the case study progressed, focus codes such as *evaluation*, *purpose of data use*, *purpose of IR*, *history of CPS*, and *history of Wilson* emerged. These codes informed the structure of subsequent interviews, which primarily took place at the beginning, middle, and end of the school year. Interviews were then used to saturate these focus codes (Charmaz, 2006). The transcripts were then arranged chronologically, in order to provide an account of how the ILT engaged in the activity of IR over the course of the entire school year.

An Enactment of Instructional Rounds

The cultural and historical context of the research context is critical for identifying manifestations of contradictions in the ILT’s work. Prior to this school year, the Chicago Teachers Union (CTU) voted to go on strike, which lasted for eight days and delayed the beginning of the school year by seven days. In the resolution of this “bitter dispute” (McCune, 2012, ¶1), the CTU agreed to a new teacher evaluation system called REACH Students (Recognizing Educators Advancing Chicago’s Students). Based on Danielson’s (2013) framework for teaching, this new evaluation system places significant emphasis on students’ growth on teacher- and district-developed performance assessments administered at the beginning and end of the school year (CPS, 2012).

This new accountability system is part of the dominant political discourse in the United States around using data to measure the quality of instruction teachers provide. Beginning with No Child Left Behind and continuing with the current administration’s initiative of Race to the Top, differentiating the most- and least-effective teachers continues to be a significant priority for national educational policy. The CTU strike embodies a growing movement against these accountability-focused policies. In the press release for the strike, CTU argued that using students’ test-score data “is no way to measure the effectiveness of an educator” (CTU,

2012, ¶5). While ultimately agreeing to the new evaluation system, the use of student test-score data for evaluating teachers remains a contentious issue, both socio-politically and in the research literature (Haertel, Rothstein, Amrein-Beardsley, & Darling-Hammond, 2011).

The preceding context serves two purposes. First, it foregrounds the *contradiction of data-use*. Data-use can be characterized as serving either formative, improvement-centered interpretations or summative, evaluation-centered interpretations. Within the context of a formal organization, both interpretations are always present, and a constant negotiation must occur to focus an iterative cycle of data-based practice. Second, it frames IR as a carrier for CPS to “touch” the activity of its constituent schools, including Wilson. As a carrier, IR may serve *coercive*, *mimetic*, and/or *normative* purposes for influencing the object of data-based activity. This manifests another dialectical contradiction: the *contradiction of organization*. Schools must always and already be seen as comprising agentic individuals while simultaneously constituting social, cultural, and historical institutions. Both contradictions significantly manifested in the work of the ILT, leading to a transformation of activity seen as ambiguously expansive.

Manifestations of the Contradiction of Data-Use

During the previous year, the ILT initiated a series of “learning walks,” where a small group (three to four) of teachers visited another teacher’s classroom focusing on a “problem of practice,” which the ILT decided was student engagement for that year. Although similar to IR, City et al. (2011) note that this learning walks “has been corrupted in many ways by confounding it with the supervision and evaluation of teachers” (p. 4). At Wilson, however, the ILT made significant effort to distinguish the practice of learning walks from administrators’ evaluation of teachers. An assistant principal, Ms. Day, reiterates their learning walk protocol:

So, if we want to have the same protocol, where the data team—without me being present at that time, at the data team at all—receives all the documents, and processes, and reports only data that cannot be traced to any particular teacher, any particular classroom. So, we want to keep this part, I think it's an easy to agree part. I believe that we've worked out the issues last year, and that might be easiest to agree on. (ILT Meeting, November 26, 2012)

Despite the decision to retain the previous year’s protocol, disambiguating evaluation from instructional improvement remained a significant conflict for the ILT into this school year. As the ILT discussed their implementation of IR, the introduction of REACH played a significant role in constraining the problem of practice that would become the focus for this year. Initially, the administrators suggested that the problem of practice be connected to REACH, specifically questioning techniques. Several teachers balked at this proposal. Ms. Hentges, the Fine Arts department chair, stated “this is how we’re also going to be evaluated—actually evaluated ... it might actually be disconcerting if this becomes the focus of the learning walk, where it’s their peers.” Ms. Walters, the Special Education department chair agreed, saying, “I think writing kind of objectifies it a little bit more, makes it just a little less personal, a little less threatening ... and we can say whatever we want, there's still a grain, a seed in the back of everyone's head saying, ‘This is evaluating me’” (ILT Meeting, November 26, 2012).

Here, we see that despite the work done in the previous year, the *critical conflict of evaluation* required a re-resolution in the new school year manifesting in this emotional, generalized account of teachers’ reactions to IR. Ms. Hentges echoed this sentiment in an interview, saying, “[the CPS Board] know they can’t come in here and just spy on us, so they’re getting us to spy on each other,” presenting an extreme metaphor for the introduction of IR (Interview, January 10, 2013). Borrowing Roth & Lee’s (2007) notation for dialectical relationships, this critical conflict touches upon both *subject|rules* and *subject|division of labor*. With the introduction of REACH, teachers and administrators needed to renegotiate how data-based practices, whether they were learning walks or IR, could threaten their positions as teachers and their relationships with administrators. Instead, several members of the ILT proposed focusing on the school-wide initiative of writing in the classroom, creating a new meaning for the data collected, and which became the focus for the IR.

Although resolving the critical conflict of evaluation, once decided, the ILT faced another manifestation of the *contradiction of data-use*: the *conflict of purpose*, or defining the purpose of the data collected and, thus, the form it takes. Within an activity system where the *subject* is a collective, negotiating the *subject|object* and *subject|mediating artifact* relationship became the focus of considerable discussion, because data as a *mediating artifact* affords and constrains *object* of the ILT. Initially, the ILT had decided to focus on the presence of informal and formal writing tasks across classrooms. Yet, this decision did not resolve the *conflict of purpose*, specifically of defining the distinction between informal and formal writing tasks. The following excerpt from a subsequent ILT meeting represents one manifestation of this conflict. In this exchange, Ms. Tolbert, the Social Studies department chair and Ms. Marston, the English department chair, argue for the need of a common definition of informal/formal writing with the principal, Mr. Osborn and Ms. Day:

Ms. Tolbert: We have a list of all the things, but we need that definition in order to move forward.

- Ms. Marston: I agree. I don't think we, for English, I don't think we can move forward until we have that. Because, even within English we have different definitions of what that is.
- Ms. Waters: So, that definition would be common to all the departments.
- Ms. Marston: Yeah.
- Ms. Tolbert: So this could be meaningful.
- Mr. Osborn: Yes. Right. That's the thing that links. (ILT Meeting, January 7, 2013)

Despite the trepidation that bringing the decision to the departments would lead to an untenable number of different definitions, the ILT decided to have individual departments assign writing tasks to informal or formal categories. Examples of these writing tasks that were ultimately returned from the departments include: (1) formal writing tasks such as written explanations to mathematical problems (Math), annotated bibliographies (English), REACH style written responses from a prompt (Social Studies), research paper (Science); and (2) informal writing tasks such as showing work and mathematical reasoning when solving problems (Math), reflection journals (English), source content acquisition, analysis tools, and worksheets (Social Studies), Frayer Model (Science) (IR Document, January 28, 2013).

At another ILT meeting, Ms. Hillman, the Instructional Support Leader (ISL) from the district challenged this purpose: "I'm not sure if knowing if it's happening or not matters, because you're going to see formal and informal writing in the building ... so I just really want to get back to what is the purpose of looking for it" (ILT Meeting, January 7, 2013). Other members of the ILT pushed back, supporting the idea of understanding what types of writing were occurring across the school:

- Ms. Walters: I mean, I think part of the-, [Ms. Marston] and I both agree that, at the beginning of all of this, when we started talking about writing, and [Ms. Marston] brought it up was, that this is our school-wide initiative and, is it really happening school-wide was the question where it started with [Ms. Marston], and I think just getting a baseline of how much writing is actually going on outside of an English or a Social Studies classroom, particularly, these two subjects lend themselves to more writing than others, might be helpful.
- Ms. Hillman: Okay, so then, for me, that clarifies it because then your purpose is to see if your school-wide initiative is implemented. So, that's actually a good purpose for the walk. "Did we implement the school-wide strategy of writing and how are we seeing it?" So, that, to me, becomes, now, I have a sense of your purpose for collecting this information. (ILT Meeting, January 7, 2013)

While this resolution of form for the IR allowed the ILT to move forward with their implementation, a further *conflict of authenticity* arose around the validity and reliability of the data collected. During one meeting, Ms. Tolbert worried that "the teachers can produce a writing assignment for that day and forget about it, and I'm afraid that's what's going to happen, I mean, I know that's what's going to happen" (IR Subcommittee Meeting, December 18, 2012). Ms. Day agreed, stating, "I definitely don't want it, again, to be a dog-and-pony show," suggesting that she believed this was the case during their learning walks last year (IR Subcommittee Meeting, December 18, 2012). From this, we see that the *conflict of authenticity* involves the *subject|community* relationship, as the ILT members presumed inauthentic engagement from the faculty with the IR leading to questionable, if not worthless, data. In a later IR Subcommittee meeting, Mr. Metts, the Math department chair, reiterated these concerns:

- Mr. Metts: This is not a good way of collecting data. I mean, we get information but, you know, the best is, as you said, over time, or at random, are a little bit more valuable. And you don't really look at a little better idea of how much is it happening. I don't know if that's what you want to do, as that will cause a lot of stress or whatever. You don't know when they're coming in, but, then again, you know, you're really influencing what you're trying to observe.
- Ms. Walters: You're definitely influencing the frequency of the writing by saying, "We're coming to look for writing," but you're not necessarily going to influence the types of writing that people are doing, so you still might gather data, I mean, you still will, may see some, you know, information about is it informal or is it formal. (IR Subcommittee Meeting, January 18, 2013)

After this point, the discourse within ILT meetings did not question the potential quality (or lack thereof) of the data collected through IR, seeming to provide a resolution to the *conflict of authenticity* as Ms. Walters argued that the data-form would allow for authentic data regardless of the authenticity of engagement from faculty members. At this time, though, a *conflict of outcome* arose around the *outcome* of IR, specifically

with regards to the actions that would arise within the school based on the data collected. Earlier in the year, the ILT spoke about the data from IR “[driving] future professional development” (ILT Meeting, November 26, 2012). Ms. Day reinforced this idea of the IR data providing a focus for professional development (PD) during an interview, where she stated the data from IR “might give us a clear indication as to how those PDs might be structured” (Interview, January 9, 2013).

In spite of this seeming consensus on IR informing future PD for the faculty, other members of the ILT had reservations about how the data they were to collect could serve that purpose. During an IR Subcommittee meeting, Ms. Tolbert raised this concern by saying, “I’m saying that if we’re hoping to have a clear direction for PD out of this learning walk, I don’t think it will be there. I think that it will do a wonderful job in stimulating conversation and reflection within the departments.” (IR Subcommittee Meeting, January 18, 2013). Mr. Osborn agreed with this statement, “And that’s instructional rounds. That’s reflection.” Consequently, the overarching purpose of IR, the *object* of the activity, for Wilson shifted from an initial focus on guiding PD toward one of promoting reflective dialogue within and across departments about the school-wide writing initiative.

By February, the ILT had resolved numerous conflicts around IR: within the *division of labor* and *rules* administration would not see any of the raw data or be involved in its analysis; the *mediating artifact* of data was to categorize tasks as informal or formal to assess the implementation of the school-wide writing initiative; and IR had the *object* of fostering reflective dialogue amongst the faculty. Through these resolutions, the dialectical relationships constituting the activity system of the ILT changed and, critically, an *object* of the activity of IR had been negotiated. After these conflicts and critical conflicts were resolved, discussions of IR were absent in further ILT meetings as the staff began implementing the process. ILT meetings became centered on topics such as reorganizing classroom utilization, upcoming performance management reviews with district representatives, and ACT preparation.

Manifestations of the Contradiction of Institutions

The next and last time IR were discussed in an ILT meeting was several months later when the aggregate results were presented. The entirety of the ILT’s discussion about the results lasted less than ten minutes, with three of those taken up by a discussion of observing IR at another school. Ms. Tolbert, who aggregated and analyzed the data from the IR, summarizes these results “so, I think we can say we writing across the board in all our disciplines here at [Wilson]” (ILT Meeting, April 22, 2013).

Curiously, the ILT never discusses IR during their meetings after this report of the data. Of further note here is the fact that department heads did not share their teachers’ reflections on the data nor did they even suggest that such discussions were had. These observations structured the interviews held at the end of the year, where teachers were asked to discuss what value they gained from participating in IR. Within these interviews, ILT members discussed their perspectives on the IR process, revealing a disconnect between the resolution of conflicts during the ILT and IR Subcommittee meetings and individual members’ ongoing struggle with these manifestations of contradictions. Throughout their accounts, an underlying *contradiction of organization* manifested in a *double bind of agent|worker*. Eliciting this double bind were the different perceptions of IR as a mechanism of institutions that threatened their identities as agents and reinforced their vulnerability as workers. ILT members believed IR to be either *coercive*, *mimetic*, *normative* or some combination of the three, leading to the perception of a lack of autonomy or helplessness.

When asked about what challenges they encountered when engaging in IR and addressing the conflicts manifested during that process, many ILT members pointed to the *coercion* from the Network, the partition of the CPS district that Wilson was a part of, for having to engage in IR in the first place. For example, Ms. Hentges expressed a perceived lack of agency for the ILT as a whole when she said, “we became just a function of the Network, which isn’t bad, of course, we want to please our bosses, but it didn’t really work on the instructional issues that are in this building” (Interview, June 12, 2013). Ms. Tolbert made a similar statement, “learning walks are driven by our Network, so it took up a lot of our time where I don’t think most of us would have voted to put our time there” (Interview, June 18, 2013). From these statements, participation in IR became merely an act of compliance and was not perceived as a means for exercising their agency as teachers and directing the improvement of their instruction.

Exacerbating the *double bind of agent|worker* were further issues pertaining to the *conflict of evaluation*. Although the previous year’s work established a protocol that separated administration from analysis of the data, the confluence of the new REACH evaluation system with the mandate for engaging in IR unsettled this resolution for several members of the ILT. Both administrators from the ILT mentioned this issue when asked about the challenges of implementing IR. Ms. Day summarized this concern and connected it to the institution of schooling, saying, “obviously, again with the strike and the union involvement, it is a part in teachers feeling uneasy about anything that looks like evaluation” (Interview, June 6, 2013).

Recall from the enactment of IR, the *conflict of evaluation* initially became manifest through administrators raising the idea of IR being focused on questioning techniques. However, this connection between REACH and IR did not begin with the administrators at Wilson. Rather, connecting the two was a

deliberate push from the Network and this was communicated within the PD sessions some ILT members attended during the school year. Ms. Smith, a high-ranking member of the Network, characterized this connection by saying, “one [series of professional development] is dealing with incorporating question and discussion techniques in the classroom, and we’ve kind of done that through the lens of instructional rounds” (Interview, May 24, 2013). Accordingly, tying IR to a domain of the REACH evaluation process was a conscious effort from the Network embodied within their messaging and training they conducted for administrators and teachers.

From the perspective of the Network, this connection was seen as a way of establishing IR as a means for improving teachers’ instruction. Mr. Hensen, the consultant who ran several of these PD sessions, also noted this connection between questioning and IR, framed this connection as a focal point for collaboration, saying from a teacher’s perspective, “let’s sit down and hammer those things out together, so that when an ILT comes through my classroom on instructional rounds, or if my principal or my assistant principal comes in for my REACH evaluation, it’s all kind of connected” (Interview, June 15, 2013). Furthermore, Ms. Smith underscored the efforts of the Network to establish this connection in all of its schools, “so we’ve been meeting with school teams one on one to discuss how that can be implemented ... and so, really we’ve been planning on how to incorporate it for the next school year in all the schools” (Interview, May 24, 2013).

In this sense, the implementation of IR was a form of leveraging the institutional mechanism of *mimicry*, attempting to push all Network schools towards focusing on a specific domain of REACH. At Wilson, ILT members immediately rejected this connection to their professional evaluations. In so doing, the ILT resolved the *conflict of evaluation*, yet were left with the challenge of resolving the conflicts of purpose and authenticity that such a connection may have facilitated. This led to multiple ILT members experiencing the *double bind of agent/worker* as they had to comply with a practice that they struggled to authentically connect with their instruction or acquiesce to a connection with their evaluations and a potential breach of the CTU contract. Mr. Robinson made this clear in an interview earlier in the year, stating, “there’s the whole differentiation, which has to be made, between evaluation in the legalistic sense of the term as the Board applies it, and observation ... because there’s only four people in the building [the principal and assistant principals] ... who can make evaluative judgments” (Interview, January 14, 2013). Ms. Marston stated her discomfort with this position, saying, “the Network really wanted, was really trying to push this idea that department chairs are pseudo-administrators, and we’re not, we’re teachers working under the same contract as every other teacher in the building” (Interview, June 10, 2013).

Despite the imperfect transformation of IR, many members of the ILT viewed their participation as valuable in establishing *norms* around classroom visitation and observation and all members of the ILT saw potential in this type of activity, though perhaps not in IR, per se. Ms. Tolbert described this normative process as “reducing, lowering the barrier ... concerns seem to relax over time as they become less threatening because the data isn’t used in a negative way,” suggesting that engagement in IR eased teachers’ apprehensions toward classroom observation and data collection (Interview, June 18, 2013). Ms. Walters saw this value in their collective work, saying, “there was true collaboration [in the IR] and I think [the ILT] has become more collaborative” (Interview, June 17, 2013). From these perspectives, engaging in IR became part of a normative process, lowering barriers to sharing classroom practices and finding value in that activity.

The ISL and administrators shared similar views. Ms. Hillman observed, “[the ILT] don’t want to say that a purpose for something [like IR] until they know that that purpose they set is going to fit in the school’s vision and fit with each teacher, and that each teacher will actually implement it” (Interview, May 28, 2013). Ms. Day expressed hope for future iterations of IR because it “opens the discussion about the teaching practices, ... so there’s a lot to be gained, a lot that we never had a chance to implement because of this tendency for teachers to stay in their classroom” (Interview, June 5, 2013). From these statements, IR provided the ILT another opportunity to increase comfort with having observers in classrooms, helping to establish similar situations as the norm for teachers at Wilson.

Indeed, all the interviewed ILT members made note of the power of classroom observations. While also noting the increased comfort of having people enter and exit the classroom, Mr. Metts suggested that, “if we actually had some serious program we were trying to implement, we would be really interested in how it was being used in other classes and what were the results from the students” (Interview, May 22, 2013). Ms. Marston made a similar remark, stating, “I think that instead of instr- maybe a version of instructional rounds ... if I can go into a teacher’s room who’s teaching the same content that I am and focus on something that our team specifically needs” (Interview, June 10, 2013). Ms. Hentges asserted that regardless of the form it takes, classroom visitation is valuable because, “teachers can’t help themselves when they go into a classroom, they just love learning and looking at other teachers and what they’re doing” (Interview, June 12, 2013). Collectively, it seems that while IR may not have achieved the *outcome* established by the ILT, engaging in the activity was valuable in that it supported establishing classroom observation as a normative practice at Wilson.

Conclusion

At the beginning of this paper, we asked two questions: How and what does the ILT learn from engaging in the practice of IR? What contradictions manifest in the activity of IR, and how are they (if they are) resolved? From our analysis, we saw that the ILT at Wilson engaged in a cycle of internalization, renegotiating and redefining the dialectical relationships comprising their activity system as they engaged in IR. We then saw an attempt at externalization, where data was gathered from classrooms, yet the *outcome* of reflective dialogue never materialized. This resulted in an incomplete transformation of the activity system and, thus, imperfect learning. However, through increasing the internalization of *norms of observation*, the ILT was able to find value and learn from engaging in IR.

Within the theoretical framework of CHAT-IT and the methodological framework targeting discursive manifestations of contradictions, organizational learning requires the resolution of inner contradictions. While the ILT was able to resolve manifestations of the *contradiction of data-use*, in so doing, its members also opened itself up to being placed in the *double bind of agent|worker*, a manifestation of the *contradiction of organization*. In many ways, this double bind undermined the potential learning that may have occurred through IR, but its tentative resolution through establishing new norms was critical to the learning of the ILT.

As one of few studies adopting Ogawa et al.'s (2008) CHAT-IT framework and applying the methodological framework proposed by Engeström and Sannino (2011), this case study supports the fruitful dialogue between these research communities in understanding organization learning as culturally and historically situated activity. By applying these frameworks in this case study we begin to see the power of conceptualizing organizational learning as an activity occurring within social, cultural, and historical contexts.

References

- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: SAGE.
- Chicago Public Schools. (2012, March 30). CPS unveils new comprehensive teacher evaluation system collaboratively designed to support and empower educators to boost student learning. Retrieved from http://www.cps.edu/News/Press_releases/Pages/03_30_2012_PR1.aspx
- Chicago Teachers Union. (2012, September 9). Press release: CPS fails to negotiate fair contract to prevent first strike in 25 years. Retrieved from <http://www.ctunet.com/blog/cps-fails-to-negotiate-fair-contract-to-prevent-first-labor-strike-in-25-years>
- City, E. A., Elmore, R. F., Fiarman, S. E., & Teitel, L. (2011). *Instructional rounds in education: A network approach to improving teaching and learning*. Cambridge, MA: Harvard University Press.
- Cosner, S. (2012). Leading the ongoing development of collaborative data practices: Advancing a schema for diagnosis and intervention. *Leadership and Policy in Schools*, 11(1): 26-65.
- Danielson, C. (2013). *The framework for teaching: Evaluation instrument*. Princeton, NJ: The Danielson Group.
- Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R.-L. Punamäki (Eds). *Perspectives on Activity Theory* (pp. 19-38). New York: Cambridge University Press.
- Engeström, Y., & Sannino, A. (2011). Discursive manifestations of contradictions in organizational change efforts: A methodological framework. *Journal of Organizational Change Management*, 24(3): 368-387.
- Haertel, E., Rothstein, J., Amrein-Beardsley, A., & Darling-Hammond, L. (2011, September 14). Getting teacher evaluation right: A challenge for policy makers. Retrieved from http://www.aera.net/uploadedFiles/Gov_Relations/Getting_Teacher_Evaluation_Right_summary_brief-FINAL.pdf
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.
- McCune, G. (2012, October 4). Chicago Teachers Union ratifies deal that ended strike. *Chicago Tribune*. Retrieved from http://articles.chicagotribune.com/2012-10-04/news/sns-rt-us-usa-chicago-schoolsbre89309m-20121003_1_chicago-teachers-union-karen-lewis-part-on-student-performance
- Ogawa, R. T., Crain, R., Loomis, M., & Ball, T. (2008). CHAT-IT: Toward conceptualizing learning in the context of formal organizations. *Educational Researcher*, 37(2): 83-95.
- Roth, W.-M., & Lee, Y.-J. (2007). "Vygotsky's neglected legacy": Cultural-historical activity theory. *Review of Educational Research*, 77(2): 186-232.
- Smylie, M. A. (2010). *Continuous school improvement*. Thousand Oaks, CA: Corwin.
- Spillane, J. P. (2012). Data in practice: Conceptualizing the data-based decision-making phenomena. *American Journal of Education*, 118(2): 113-141.
- Young, V. M., & Kim, D. H. (2010). Using assessments for instructional improvement: A literature review. *Educational Policy Analysis Archives*, 18(9). Retrieved July 22, 2013, from <http://epaa.asu.edu/ojs/article/view/809>

Tensions and Possibilities for Political Work in the Learning Sciences

Angela N. Booker, UC San Diego, 9500 Gilman Dr., La Jolla, CA 92093, angelabooker@ucsd.edu
Shirin Vossoughi, Stanford University, 485 Lasuen Mall, Stanford, CA, 94305, shirinvossoughi@gmail.com
Paula K. Hooper, Exploratorium, Pier 15, San Francisco, CA 94111, phooper@exploratorium.edu

Abstract: How can the learning sciences engage more directly with the political dimensions of defining and studying learning? What might this engagement offer for democratizing learning? This paper delineates a tension between deep studies of learning and explicit attention to issues of power, inequality and human dignity. We frame this as a productive tension that will generate new insights, as well as conceptual and methodological tools that contribute to the democratization of learning. We identify a history of ideas inside and outside the learning sciences that inform this objective, including the political dimensions of the field's founding theorists. We then offer examples of ways these tensions manifest in our own empirical work, and conclude by considering how explicit attention to political dimensions of learning can advance our theories about what learning is, about what it is for, and about the conditions that give rise to deep forms of learning for all.

Expanding Space for Politics in the Learning Sciences?

In this article we are grappling with the following questions: How can the learning sciences engage more directly with the political dimensions of defining and studying learning? What might that engagement offer for democratizing learning? Addressing these questions is crucial to educators and designers of learning environments who share a commitment to working with youth and communities contending with marginalization. The work is underway (Bang, et. al, 2012; Gutiérrez, 2008; Lee, 2001; Nasir, Roseberry, Warren & Lee, 2006). We believe the time is ripe for making this a more central preoccupation of the field. In our view the learning sciences has a political tension—a tension that has emerged as a shared thread across our work. Here, politics refers to explicit attention to issues of power, hierarchy and inequity, and to the roots of those issues. The field also has an edge that can ground and inform the work of colleagues in other fields who directly address political dimensions of education—we know how to investigate *learning* with methods that trust and are informed by locally situated social actors and their multiple forms of practice and knowledge. Here, we will describe the boundaries of this tension, ground the discussion in examples from our own work and propose a theoretical stance that privileges human dignity as a central concern.

A key strength of scholarship in the learning sciences is in the combined commitment to theories that explore learning as situated in the lives and practices of people (Dewey, 1942; Vygotsky, 1978; Lave, 1987; Gutiérrez & Rogoff, 2003) while drawing on grassroots methods in research design (Hawkins & Pea, 1987; Barab, et al., 2004; Barab & Squire, 2004;). Theory offers a way of explaining phenomena of concern, but we sense an apparent contradiction within our field about what we want to explain. At times, learning sciences can orient toward a dominant frame regarding the purposes of education and learning: educational achievement and competition in a global marketplace. This frame was not created by the learning sciences, but the field is responsive to it. This frame is expressed in a variety of ways, yet competition remains the organizing feature. In this view individuals prepare to compete within an economic system and, through that system, collectively contribute to a country's economic standing—and in turn, degree of control (Eckert, 1989; Varenne & McDermott, 2008). In the learning sciences, this looks like a commitment to developing expert knowledge and deep conceptual understanding without a broader attention to the political and economic factors that shape and constrain trajectories of learning. We have operated as a future-oriented field, researching and designing toward learning environments that can consistently yield deep conceptual understanding, reflection, and expert knowledge that can serve our practice in the world (Sawyer, 2006). The challenge arises when we fail to or choose not to articulate to what end. In those cases, the dominant narrative is ready and waiting to absorb that future as its own. Put another way, it is one thing to marshal support for strong systems of education through calls to prepare people to be effectively competitive—people tend to envision being on the succeeding end of imagined competitions. It is the *same* thing, although not often articulated, to establish losers in the competition. This is where the political tension becomes taut.

An alternative frame places its emphasis on human dignity as a mode of inclusion. Again, this alternative frame did not originate within the learning sciences, but the field clearly desires to be responsive to it (Esmonde & Caswell, 2010; Nasir & Hand, 2006; Nasir, et al., 2006). In this view, varied, localized, cultural ways of knowing can yield respect, reflection, and cooperation—if not riches. Notably, voices in the field organized in this way tend to express the politics at play in the work of learning and knowing. In this frame, deep conceptual understanding and the practices of novices and experts are still key features of the work. The

distinction is in making human dignity and social equity the primary commitments. With this commitment comes a more explicit attention to the kind of future that is embodied and potentially engendered by alternative educational designs and practices. Where new social and political visions are made explicit, research on deepening conceptual understanding and expertise also takes shape differently: intellectual activity is understood as embedded in social relations; those social relations can either reproduce or reimagine and transform the hierarchies (raced, classed, gendered, aged, nationalized, etc.) and forms of competition that uphold the status quo. Thus, in addition to treating learning as a cognitive, affective and social process, understanding human learning as a fundamentally political process can lead to distinct empirical insights, designs and methods. Where we do not actively attend to political dimensions in learning, we can reproduce depoliticizing currents that weaken our analyses, making it more difficult to scale our work across settings.

Building on Two Active Modes of Theorizing

One way to illuminate this distinction is to consider two active modes of theorizing in the learning sciences. One approach is to theorize in ways that yield scalable designs to support learning in various disciplinary contexts:

Unlike these previous generations of educational research, learning scientists spend a lot of time in schools—many of us were full time teachers before we became researchers. And learning scientists are committed to improving classroom teaching and learning—many are in schools every week, working directly with teachers and districts. Some even take time off from university duties and return to the classroom, teaching alongside teachers and learning how to make theories work in the real world. This is a new kind of science, with a goal of providing a sound scientific foundation for education (Sawyer, 2006, p. 15).

When we theorize from this place, we emphasize science that is grounded in real practices and lived experiences of students, teachers, novices, experts, and professionals. We also tend to prioritize schooling, even as we draw from understandings of learning in everyday life. In and of themselves, these are not overtly political moves. Rather they allow us to reveal practices—local or disciplinary—around which knowledge can be jointly produced. Learning, then, is understood in terms of its depth and effectiveness for yielding flexible and adaptive expertise. This is the bread and butter of the field and we are making meaningful progress. Still, systems that privilege some and marginalize others persist. This is a concern the learning sciences also takes seriously.

A second mode of theorizing holds tightly to local practices that are culturally mediated. With this lens, practices that support learning are somewhat freed from the disciplines and can be rooted in a wide variety of activities and cultural spaces (e.g. sports and games, shopping, organizing and activism, etc.). When our lens is focused on cultural ways of knowing and doing, politics have more freedom to emerge through processes of negotiating meaning. We are interested in extending these situated, sociocultural modes of theorizing by investigating what it would mean for the *method* to scale while the design itself may not. If the learning sciences is to emphasize human dignity, we need to theorize in ways that may not be scalable in a strictly scientific sense. That is, “a sound scientific foundation for education” might be upgraded to sound human-centered methods for learning where knowledge is not only co-constructed but also politically active.

A Good Problem

This is a problem—a good one—because the tension here is productive. The field is primed to debate the merits of how learning is characterized—as learning for its own sake, as a necessity for democratic life, as a necessity for economic participation, as a necessity for humanization. Ours is not the first field to take up these overlapping roles, but ours is positioned to provide leadership on what to privilege, why to privilege it, and how what we design leads us to these ends. This is so because our methods merge cognitive and computational ways of investigating understanding, ethnographic attention to the meaning people make together, and design practices that are iterative and rooted in a sociocultural theoretical stance. For these reasons, we are arguing for the learning sciences to become more attuned to the field’s own dominant and marginal voices. Our interdisciplinary work is still situated in a stratified world, and we must take seriously the reality that we are also susceptible to being organized by dominant ideologies. This practice will organize us to think about how to include work that engages directly with political aspects of learning in ways that push and grow the field in directions that urgently reimagine social relations and the systems that support them. This is a critical intention, so we will dwell here for a moment.

The tension we are highlighting has a longer history in arguments that pit science against politics and emotion and ensnare scientific discourse and rationality. In search of universal principles, the humans can get lost. Yet too much emphasis on the local particularities of lived experiences might impede the effective application of what is learned more broadly. Harris and Shultz (1993) took this up in the field of law from the standpoint of reason and emotion—in the law, it seems, emotion has had an unscientific reputation. In the article, the authors called for reunification of reason and emotion:

Emotion ruled off the official educational agenda remains unchallenged, unexamined, and undisciplined by reflection and analysis. Emotion made impermissible becomes emotion disowned, emotion for which no one is responsible....If emotions are successfully repressed, rationality suffers in any case. Draining intellectual arguments of emotion also drains them of meaning....Rationality unchallenged by emotion makes legal analysis an abstract exercise. The result is not only boredom but legal thought this both technically precious and practically irrelevant (Harris & Shultz, 1993, p. 1780).

Similarly, this phenomenon is reflected in research on learning and education – where politics may have an unscientific reputation. This is a tension that persists between researchers and educators. Indeed, it can interrupt practical and intellectual momentum. Happily, the learning sciences are organized to confront this reality.

The wedding of interdisciplinary theories (e.g. sociocultural and cognitive), with interdisciplinary methods (e.g. ethnographic, design-based, and computational) allows for ecologically grounded empirical research that can effectively take up politics that permeate learning both as practice and becoming. This is why we are concerned with what can at times feel like a depoliticizing context of research and theorizing. Studying learning in a highly stratified and competitive world is not a neutral act. All research has a perspective and is therefore political. But some research gets seen from a positivist perspective as “political” and thus subjective and advocacy driven—qualities which are in turn framed as less legitimate. We are suggesting this is one of the mechanisms through which more political voices get marginalized. In this way, we are viewing ourselves—scholars engaged with the learning sciences—as learners and our sites where intellectual work is shared and debated, as a learning environment. If we are interested in what the learning sciences can offer for how to democratize learning, we need our theoretical and methodological tools to help us wrestle with inequity and dehumanization. We need them to organize our investigations to ensure that teaching and learning are humanizing experiences and to understand the myriad ways this can look. When political approaches are embraced as a valid and necessary thread of work, the learning sciences themselves will be strengthened.

Concepts Related to Political Dimensions of Learning: A History of Ideas

In arguing for a more direct engagement with the political dimensions of learning, we situate ourselves within a history of ideas and research in the learning sciences that has sought to develop adequately complex and human-centered conceptual and methodological tools. This includes, for example, work that treats learning as a fundamentally cultural process and draws explicit connections between this approach and the wider social, political and economic constraints on learning for youth and communities (Gonzalez, 2005; Nasir, et. al., 2006). The recent special issue of *Human Development* organized by Nasir and Bang (Eds., 2012) is an example of this approach. As Lee (2012) writes: “The problem space the contributors to this volume have taken on is inherently complex because it entails multiple levels of context: the immediate setting in which individuals participate; relationships across the multiple settings that people navigate; the broader cultural, political, economic, and, indeed, ideological belief systems and institutional configurations in which these micro-level settings exist.”

In our own work, we are drawn to micro-level analysis for the ways it helps ground research in moment-to-moment pedagogical practice and the perspectives of children, youth and adults in educational settings—both inside and outside of schools. This close attention to micro-level practice is one of the hallmarks of research in the learning sciences. At the same time, if our analysis only focuses on the learning of disciplinary content without attending to the ways developing new understandings intersects with questions of epistemology and identity that in turn reflect broader political and economic inequalities, then the theories that emerge from our analysis may be inadequate as tools for equity oriented design and pedagogy. This is where embracing political dimensions of learning can help imbue the learning sciences with analytic tools that allow us to study educational relationships, interactions and experiences as deeply connected to broader social problems and visions. The theories and methods we develop are our technologies for reconnecting with the more political side of the history of ideas around learning. These tools are essential to carrying forward the field’s stated goals.

What are some of these tools? Within the field, we are drawn to socio-cultural or cultural-historical approaches for their attention to the centrality of cultural mediation (Wertsch, 1998), to human activity or cultural practices (rather than individual behavior) as units of analysis (Cole, 1996), to the deeply social dimensions of learning and their relationship to human potential and possibility (Vygotsky, 1978; 1986), to tensions and contradictions as progenitors of change (Engeström, 1991; Gutiérrez, Baquedano-Lopez, & Tejada, 1999) and to ecologically valid forms of research (Cole, Hood, & McDermott, 1997). Many of these concepts have their roots in political problems and concerns. For example, ecological validity emerged in part from a critique of research that measured human intelligence based on normative cultural tasks, leading to deeply problematic conclusions about the intellectual capacities of non-Western communities, both inside and outside the United States. This political and humanistic critique was attuned to relationships between psychological/educational research and colonial ideologies and practices and to the need for researchers to

reflect on their own assumptions. Such critical and reflective modes of research engendered new methods that understood cognition as embedded in cultural practices (Scribner & Cole, 1988) and studied learning as a situated activity, an approach that revealed intelligence and ingenuity where deficit had been assumed.

Thus, while some of these concepts and approaches sit comfortably within the learning sciences, we are interested in what can be learned from making explicit the political critiques that brought them into being—and what openings might emerge in future strands of work. Similarly, we seek to explore both the political commitments of founding figures within the field, such as Dewey, Piaget, Vygotsky, Papert and Bruner, and how these concerns gave shape to their ideas about learning. Highlighting these concerns can help connect their theoretical contributions to the economic and political tensions outlined above—namely the purpose of education in a democracy, and the role alternative assumptions and approaches to learning can play in helping bring about an egalitarian future. This theoretical work can help counter currents of de-politicization—where theories of learning are explicitly or implicitly stripped of their critical bent. Revisiting the political context within which such foundational contributions emerged can also help us understand their/our blind spots (around race, or gender for example) and what these mean for our thinking today.

At the same time, some of the ideas and scholars we are drawn to may hold a more marginal place within the field. For example, the work of Paulo Freire (1985) and the traditions that grew from his contributions directly attend to problems of dehumanization and humanization, oppression and liberation as tied to educational theory and practice. Similarly, critical race and decolonizing approaches to education make central the role of race and racism in the educational experiences of students of color, and offer alternative methodologies for community-based research (Paris & Winn, 2014; Smith, 1999; Tejeda, Espinosa, & Gutiérrez, 2003). Do these ideas have a place within the learning sciences? Are they seen as overtly political and therefore less “scientific”? What new tools might emerge from greater dialogue across these traditions? In his commentary on the aforementioned special issue of *Human Development* Cole (2012) writes,

What is much less certain is whether it is possible, on a mass scale, to so reform schooling that the kinds of exceptional interactions that appear to unleash student creativity become the norm, not the exception. However persuasive the arguments for reframing and desettling educational practices may be, they run directly against the powerful forces that seek to amplify the effectiveness of a scientific world view in which triumph over nature, and over other humans considered less-than-human, is considered an economic and political imperative (p. 346).

Though many within the field of learning sciences would agree with the goal of making ‘exceptional interactions that appear to unleash student creativity the norm rather than the exception’—and are actively involved in research that works towards these ends—we are interested in the kinds of shared thinking, dialogue and research that would become possible through equal and explicit engagement with the ‘powerful forces’ that blunt the fulfillment of this possibility for all students.

Political Possibilities: Some Examples from Our Empirical Work

Our proposed framework therefore problematizes the overt or subtle de-politicization of learning—including dominant assumptions about what constitutes legitimate empirical research—and seeks to illuminate when and how moving towards rather than away from politics may yield practice-based alignment with the theoretical goals and premises of the learning sciences. One narrative suggests foregrounding the researcher’s values makes both a weaker researcher and weaker research. But if, in turn, we ignore the presence of values in our data and analysis, that also weakens the research. This tension opens a kind of critique that expands where the field can go and what it can do. When and how are political perspectives necessary for thorough research on learning, and what kind of future are we implicitly working to bring forward? In our work, these questions emerge in a wide range of contexts and disciplinary orientations. In this section we consider what we gain by taking a political view of learning in each of our settings and how that is connected to democratizing learning in those settings and more broadly.

Civic and Political Practice

Booker conducts community-based research with youth and attends to how they learn to engage as civic and political participants. Her studies are framed by a combination of sociocultural theory, cultural studies, and political sociology. Sites for research have included community, family, and school contexts, particularly at their points of intersection where learners can privilege their own expertise while generating meaningful learning opportunities (Goldman, Booker, & McDermott, 2007; Booker, et al. 2011). The focus on civic participation and political development derives from a concern with democratized learning and barriers that limit access to political analyses and critiques.

In a study of a student advisory board of high school students attempting to influence school district

policy making, three aspects of civic and political development became critical: technical, pragmatic, and political framing of their work (Booker, 2010). Technical frames were based in learning and applying civic knowledge (e.g. how policy is made and implemented). Pragmatic and political frames defined actions (e.g. organizing supporters, engaging the media, explicitly engaging in power analysis, etc.). Two analytical understandings of their work emerged. When examining the data for learning and outcomes, a tidy linear progression emerged that suggested students first had to learn the technical requirements for doing policy work. This was followed by a period of pragmatism where the group acted in ways most likely to receive support from adult decision makers while addressing important student concerns. Finally, they used political frames to make their views known in a contentious public debate, criticizing the Superintendent's signature school reform program and calling on her to renegotiate her contract in response to severe budget cuts. A second analytical strategy foregrounded politics to examine where youth, community groups, district leaders, adult allies, and members of the press gave explicit attention to issues of power, hierarchy and inequity. This analysis revealed the students' nearly constant use of technical, pragmatic, and political frames to negotiate meaning and action.

The linear picture, by itself, left substantial room for interpretation. The journey from technical to pragmatic and then political frame could be a straightforward learning trajectory. Or, it could suggest that political frames came into play as a form of hubris at best or co-opting by opportunistic adults at worst; the actions of oppositional adults implied this interpretation. A third read was that youth were selecting political frames to advance their agenda, and it was a sign of sophistication. This was an interpretation shared by youth and allied adults. Analyzing data with specific attention to politics as a fundamental aspect of the learning environment as well as policy-making practice revealed all three frames were evident in practically every event in which the student advisory board engaged. They constantly negotiated the meaning of their collective work by filtering possible actions through these frames and addressing the vulnerability of their positions in hierarchies of power—as youth from a variety of backgrounds, students who were subject to the authority of adults, elected representatives of their peers, advisors with no real vote, etc. They made political choices, even when the outcomes of those choices appeared technical or pragmatic.

What was gained by engaging political dimensions of learning? The yield distinguished practices employed by youth to democratize their own learning and participation experiences and the consequences they suffered for it. Possibilities for youth and communities alike were constrained simply by defining the student advisory board as a place for learning—as opposed to participation. Politics were implied in the actions, negotiations of meaning, and contradictions that emerged as the students discovered their influence, worked to wield it, and encountered resistance and support. It was necessary but not sufficient to examine what happened—it was also imperative to examine competing historical interpretations in relation to contested visions for possible futures. Dreaming possible futures and negotiating for their existence are fundamentally democratic practices. Nurturing possible futures requires the development of political practices because those futures do not develop in a vacuum. They develop largely to the degree that they displace existing practices.

Programmable Media

Hooper has studied children's learning with programmable media and construction of computational ideas from a sociocultural perspective. Her work has focused on classroom settings where teachers are explicitly engaged in supporting the learning of African American and Latino elementary school children. She has examined how the theoretical and design tools of constructionism (Papert, 1993) explain from a cognitive perspective that creating projects with classic constructionist tools such as Logo and Lego Mindstorms supports children's learning of math and science ideas. Her work has illuminated the role of a political stance through studies of programmable media in classrooms where teachers seek to include cultural experience in their pedagogical choices.

One study of second-grade students' explorations of geometry with programming in Logo revealed that students' use of turtle graphics supported their construction of geometric designs. It also revealed that the students' engagement with problem-solving and extending each others' ideas expanded when the pedagogical context included the use of discourse practices that were familiar to them. The teacher had a socio-political value orientation to using these practices as well as a constructivist one. From this stance, she instituted a sharing time for students to present and discuss successes and challenges in their projects with the whole class. This was similar to a math talk or poetry sharing time that had immersed the class in practices of sharing approaches to solving problems, offering writing for communal critique and review, and affirming each others' progress. She encouraged a discourse formed by children's everyday ways of expressing their ideas and developing academic discourse that was mutually constituted by her modeling and interactions between students that were comfortable and respectful. Her stated goal was for each student to "have a voice" in sharing their ideas, no matter what the subject-matter or level of academic achievement or English-language proficiency the student brought to the experience. In this class, all the students learned to believe they could share their ideas and their ideas would be heard and respected.

For example, during one of the sessions where this sharing practice was used, a student who was low in her academic achievement initiated a lively discussion of the trial-and-error process that she used to figure out

how to draw a hexagon using turtle graphics commands. This discussion placed her mathematical ideas in a valued position in the co-construction of computational ideas within class. This example illustrates the political nature of the teacher's choice to apply the sharing practice that had been appropriated by the class to work with programmable media. This pedagogical move supported students' negotiation of everyday voices with their learning in the new academic realm of computational thinking. Over time and close examination of these sharing sessions, it became clear that this practice was a scaffold for all students sharing ideas that took their learning with programmable media to levels of complexity they could not easily reach by using the tools on their own. Analysis of data from this study suggests the need for explicit attention to: 1) elements of learning environments such as socio-politically valued classroom practices, and 2) the designs of programmable media because these aspects are operating in concert to support children's learning of and with computational ideas.

Hooper has extended examination of learning based on constructionist design principles to learners working with digital design fabrication in informal settings. Based on her earlier work, this new work explores both the affordances of digital design fabrication as a constructionist tool and the importance of an explicit pedagogical approach that scaffolds learners in bringing their social and culturally rooted epistemologies to their work. The Spiro Inquiry (Hooper & Freed, 2013) explores creating spirograph-like designs with a construction kit that is made up of pre-fabricated shapes for exploring physical construction of designs, an app for constructing designs on-screen, and features of the app that enable the fabrication of gears and designs on lasercutters or craft cutters. Workshops have been designed based on a structure for inquiry-based science (Institute for Inquiry, 2006) that is designed to engage learners in identifying and pursuing their personal interests related to the mathematical ideas embedded in the work.

This work has implications for examining the political contexts of making and tinkering environments. It emphasizes that pedagogical structure can be designed to accompany the process of making with computational tools as a way to insure that multiple pathways to learning are supported. There is a political connection between equity and learning when multiple pathways are explicitly supported. This work exemplifies critical engagement with discourse in emerging fields of making and tinkering and the political and economic tensions that exist when design for multiple pathways to learning are not made explicit.

Migrant Youth in Hybrid Settings

Vossoughi has been engaged in ethnographic research on learning in hybrid settings that seek to meaningfully connect the development of disciplinary understandings with students' lived experiences. These settings blend formal and informal elements (school-like academic tasks and modes of engagement with apprenticeship models, play, familial social relations, and everyday language practices). In one setting, high school age migrant students worked with graduate students to analyze social problems relevant to the migrant community (displacement, economic exploitation, educational inequity, gender/patriarchy) through complex social theoretical texts and written genres. Here, she studies the ways social analytic tools became *consequential* to students—serving to deepen intellectual engagement and reshape social relations.

In this case, studying learning as a political process helped *make visible* the social and political work students were engaged in when drawing on academic tools to analyze social problems (a male student working to discuss patriarchy in careful ways, a student who was dominant in English reconfiguring the leader of her reading group so that a student who was dominant in Spanish could contribute his own words). Resisting common dichotomies (theory/practice, intellectual/political activity) this perspective can help us see classroom discourse as a site of struggle and change. At the same time, treating political conversations as arenas of learning highlights the intellectual work students were engaged in as they analyzed particular problems and experiences. This view can be used to organize and assert the value of political education as a potentially rich context for academic development, and to identify the specific forms of guidance and expertise involved therein.

Vossoughi draws on Matusov's (1998) work to identify the forms of development valued in this setting. According to this approach, we must not only ask 'did development happen?' – a question that assumes normative interpretations of what constitutes valued forms of academic engagement. Rather, we must first ask: what are the valued forms of development in this setting? This question moves us towards a more situated and emic understanding of learning, but it can also be used to surface dimensions of intellectual activity that may be undervalued by dominant measures and argue for educational self-determination (Vossoughi, Under Review).

STEM Learning Through Tinkering and Making

In a second study, Vossoughi is working with a team of educators and researchers to look closely at STEM learning in an equity-oriented after-school setting organized around tinkering and making (Vossoughi, et. al., 2013). This setting serves as a prime example of the shift towards scientific practices emphasized by the Next Generation Science Standards. Children work with facilitators to design artifacts that embed STEM practices in purposeful and artistic activity. However, her analysis of learning in this setting explicitly pushes back on discourses that treat equity (only) as a matter of expanding access among communities historically underrepresented in STEM fields and professions (or Making spaces) without critically engaging with the nature

and purposes of scientific labor and with the larger “Maker Movement.”

These political concerns directly influence the ways she analyzes learning within the after-school setting: what counts as scientific? Are multiple ways of knowing supported or marginalized? When and how do STEM practices become meaningful tools? What measures are used to assess effectiveness—persistent participation in STEM fields and/or a palpably deeper curiosity and intellectual disposition towards tinkering that may manifest across a range of contexts? Here, studying learning as a political process allows us to include disciplinary content and expertise as objects of analysis, leading to distinct interpretations regarding young people’s engagement with STEM. For example, students’ complex and sometimes tension-filled relationships with “Science” as a discipline are not treated as misconceptions to be solved on the linear path towards diversifying STEM fields, but as legitimate efforts to make sense of a field replete with its own history and politics. In this way, critical questions about goals and assumptions of “Science” become seeds for imagining alternative possibilities with regard to who can engage in science, with what meanings, and towards what ends.

Advancing Learning Theory

What would effectiveness in this endeavor look like? Learning sciences can put to rest the remarkably persistent presence of deficit as a valid way to approach humans and learning. As a field, we already believe this and work toward it, but our work has not put this to rest. Rather it has responded to it. Here is a theoretical proposition: Where knowledge is depoliticized and deficit organizes the practices around learning, dehumanization is likely to occur through the reduction of what is seen to what is missing. The learning sciences can devise and investigate an alternative proposition because we are a field that merges cultural, cognitive, computational, and democratic design in service of deep reflection and understanding: Where knowledge is openly recognized as political and contested, and humans are always making sense through social and historical practices, humanization is likely to occur as we contemplate one another through disciplined reflection. The former proposition emphasizes what should be learned and whether it is effectively learned. The latter proposition emphasizes how learning occurs through practices of seeing and re-seeing each other.

This year’s conference theme invited dialogue about practice, and each of us was experiencing a rift between theory and practice in our own work. Practice draws us into the lived experience of people—the location of our learning and becoming. Yet, if those experiences draw us toward political conclusions rather than discipline based, epistemic conclusions or scalable designs, it is an open question whether the political path is an open one within learning sciences scholarship. The fields’ leadership in sociocultural theory opens the door, but it isn’t swung wide. As a field, how are we learning and becoming? Do our practices allow us to become openly political actors in relation to our scholarship? Certainly, foundational thinkers of the field were engaged in these ways. Where can we carry those commitments today?

References

- Bang, M., Warren, B., Rosebery, A. S., & Medin, D. (2012). Desettling expectations in science education. *Human Development*, 55(5-6), 302-318.
- Barab, S., Thomas, M. K., Dodge, T., Squire, K., & Newell, M. (2004). Critical design ethnography: Designing for change. *Anthropology and Education Quarterly* 35(2), 254-268.
- Barab, S. & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, 13(1), 1-14.
- Booker, A. (2010). Framing youth civic participation: Technical, pragmatic, and political learning. In L. Lin, H. Varenne, & E. W. Gordon. (Eds.) *Educating comprehensively: Varieties of educational experiences* (pp. 209-231). Lewiston, NY: Edwin Mellen Press.
- Booker, A., Montgomery-Block, K., Scott, Z., Reyes, b., & Onyewuenyi, A. (2011). Public scholarship within an urban school district: A community and university partnership approach to service-learning. *Journal of Public Scholarship in Higher Education*, 1, 67-87.
- Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Harvard University Press.
- Cole, M. (2012) Commentary on conceptualizing cultural and racialized process in learning. *Human Development*. 55(5-6), 340-347.
- Cole, M., Hood, L., & McDermott, R. P. (1997). Concepts of ecological validity: Their differing implications for comparative cognitive research. *Mind, culture and activity: Seminal papers from the Laboratory of Comparative Human Cognition*, 49-56.
- Dewey, J. (2012/1916). *Democracy and education*. New York: Dover Publications.
- Eckert, P. (1989). *Jocks & burnouts: Social categories and identity in the high school*. New York: Teachers College Press.
- Engeström, Y. (1991). Toward overcoming the encapsulation of school learning. *Learning and Instruction*, 1, 243-259.
- Esmonde, I. & Caswell, B. (2010). Teaching mathematics for social justice in multicultural, multilingual elementary classrooms. *Canadian Journal for Science, Mathematics, and Technology Education*, 10(3),

- Freire, P. (1985). *The politics of education: Culture, power, and liberation*. New York: Bergin & Garvey.
- Harris, A. P. & Schultz, M. M. (1993). "A(nother) critique of pure reason": Toward civic virtue in legal education. *Stanford Law Review*, 45, 1773-1805.
- Hawkins, J., & Pea, R. D. (1987). Tools for bridging the cultures of everyday and scientific thinking. *Journal for Research in Science Teaching*, 24, 291-307.
- Hooper, P. & Freed, N. (2013). "The Spiro Inquiry": Exploring the application of an inquiry structure for learning science content to teaching with a digital design fabrication tool. *Proceedings Interaction and Design for Children 2013* New York: ACM.
- Goldman, S., Booker, A., & McDermott, M. (2007). Mixing the digital, social, and cultural: Learning, identity and agency in youth participation. In D. Buckingham. (Ed.) *Youth, identity, and digital media* (pp. 185-206). MacArthur Foundation Series on Digital Media and Learning.
- Gonzalez, N. (2005). Beyond culture: The hybridity of funds of knowledge. In N. González, L. Moll, & C. Amanti (Eds.), *Funds of knowledge: Theorizing practices in households, communities and classrooms* (pp. 29-46). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Gutiérrez, K. D. (2008). Developing a sociocritical literacy in the third space. *Reading Research Quarterly*, 43(2), 148-164.
- Gutiérrez, K., Baquedano-Lopez, P., & Tejada, C. (1999). Rethinking diversity. Hybridity and hybrid language practices in the third space. *Mind, Culture, and Activity*, 6(4), 286-303.
- Gutiérrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher*, 32(5), 19-25.
- Institute for Inquiry. (2006). Inquiry structure for science learning. *Fundamentals of Inquiry Facilitator's Guide*. <http://www.exploratorium.edu/ifi/workshops/fundamentals/index.html> San Francisco: Exploratorium.
- Lee, C. D. (2001). Is October Brown Chinese?: A cultural modeling activity system for underachieving students. *American Educational Research Journal*, 38, 97-141.
- Lee, C. D. (2012). Conceptualizing cultural and racialized process in learning. *Human Development*, 55(5-6), 348-355.
- Lave, J. (1987). *Cognition in practice*. New York: Cambridge University Press.
- Matusov, E. (1998). When solo activity is not privileged: Participation and internalization models of development. *Human Development*, 41, 326-349.
- Nasir, N. S. & Bang, M. (2012). Conceptualizing cultural and racialized process in learning. *Human Development*, 55(5-6), 247-249.
- Nasir, N. S. & Bang, M. (2012). Conceptualizing cultural and racialized process in learning. *Human Development*, 55(5-6), 247-249.
- Nasir, N. S., Rosebery, A. S., Warren B., & Lee, C. D. (2006). Learning as a cultural process: Achieving equity through diversity. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 489-504). New York, NY: Cambridge University Press.
- Papert, S. (1993). *The Children's Machine: Rethinking School in the Age of the Computer*. New York: Basic Books.
- Paris, D., & Winn, M. T. (Eds.). (2014). *Humanizing research: Decolonizing qualitative inquiry with youth and communities*. Los Angeles: Sage.
- Sawyer, R. K. (2006). The new science of learning. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 1-16). New York, NY: Cambridge University Press.
- Scribner, S. & Cole, M. (1988). Unpackaging literacy. In E. Kintgen, B. Kroll, & M. Rose (Eds.), *Perspectives on literacy* (pp. 57-70). Carbondale: Southern Illinois Press.
- Smith, L. T. (1999). *Decolonizing methodologies: Research and indigenous peoples*. London: Zed Books.
- Tejada, C., Espinoza, M., Gutiérrez, K. (2003). Toward a decolonizing pedagogy: Social justice reconsidered. In P. Trifonas (Ed.), *Pedagogies of difference* (pp. 10-40). New York: RoutledgeFalmer Press.
- Varenne, H. & McDermott, R. P. (1998). *Successful failure: The school America Builds*. Boulder, CO: Westview Press.
- Vossoughi, S. (Under Review). Social analytic artifacts made concrete.
- Vossoughi, S., Escudé, M., Kong, F., & Hooper, P. (2013). *Tinkering, Learning & Equity in the After-School Setting*. Paper presented at the FabLearn Conference, Stanford University.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: MIT Press.
- Wertsch, J. (1998). *Mind as action*. New York: Oxford University Press

Acknowledgment

We thank Michael Cole for reading a draft of this paper and talking with us about the ideas presented here.

“So, I think I'm a Programmer Now”: Developing Connected Learning for Adults in a University Craft Technologies Course

Deborah A. Fields and Whitney L. King
Utah State University, 2830 Old Main Hill, Logan, UT 84321
deborah.fields@usu.edu, whitney.king@aggiemail.usu.edu

Abstract: In the past decade many conversations about learning have turned from primarily discipline-based or space-based settings to concerns about helping students make connections between settings, what some have popularly termed “connected learning” (Ito et al., 2013). Nowhere is connected learning more needed than in areas of programming and engineering with new digital technologies, specifically craft technologies, which holds new potential to reshape the ways that we think about these disciplines. Often, discussions of connected learning focus on children and youth with an emphasis on informal, out-of-school settings. We argue college students and adults may benefit just as much from connected learning, and formal settings of learning should not be neglected in conversations of connected learning. We describe and analyze a new university course called Craft Technologies, developed to engage non-engineering, non-computer science major students by creating projects that would change the way they thought of and used computers and electronics.

Introduction

In the past decade many conversations about learning have turned from primarily discipline-based or space-based settings to concerns about helping students make connections between settings, what some have popularly termed “connected learning” (Ito et al., 2013). Many studies note the ways that conflicts between values, discourses, roles, and knowledge between home and school, or home and workplaces result in students’ disengagement from academic learning (e.g. Holland et al, 1998; Nasir & Hand, 2008). In answer, connected learning purposefully seeks to help people develop positive relations between home, school, family, friends, hobbies, and institutional settings like school, work, and civic institutions. Thus connected learning is “socially embedded, interest-driven, and oriented toward educational, economic, or political” opportunities that have high stakes for equitable participation in our society (Ito et al., 2013, p. 4).

Nowhere is connected learning more needed than in areas of programming and engineering with new digital technologies. For several years, the disproportionate lack of women in engineering and technology sectors have been noticed by practitioners and researchers not only in these fields, but also in education and learning sciences (Margolis & Fisher 2002; Hill, Corbett & St. Rose, 2010). Cited reasons for this disparate pursuit of programming and engineering include the lack of supportive social connections that can help students better identify with those fields (Brickhouse & Potter, 2001), implicit bias about abstract and top-down models as the “best” programming or engineering (Turkle & Papert, 1991), and limited views of what computer science is and what roles are available in it (Margolis & Fisher, 2002). The perspective of connected learning provides a potential answer for these challenges, for instance with the goal of enabling better social connections in programming and engineering, providing multiple legitimate ways to engage in them, and developing broader views of what they are and who can do them. Yet how can we achieve these goals in meaningful ways?

The hobbyist-driven Maker and digital fabrication movement holds new potential to reshape the ways that we think about programming and engineering. The movement is largely interest-driven and focuses on tangible (versus abstract, screen-based) artifacts in ways that often integrate well with everyday life and personal interests (Blikstein, 2013). In particular, we argue that aspects of ‘making’ that draw on knowledge and skills that people build at home, in community and religious groups, and with friends may hold particular promise for creating spaces of connected learning related to programming and engineering. Craft technologies are a subset of the digital fabrication movement that bring together traditional hand crafts like sewing, knitting, and woodworking with electronics and programming with microcomputers. They are inclusive of but broader than electronic textiles (e-textiles) which focus largely on soft fabrics and sewing, but like e-textiles they can disrupt people’s notions of what can be made with computers and who can make it (Searle, Fields & Kafai, 2013) through the integration of hand crafts with wiring and computing.

In this paper we describe and analyze participation in a new university course called Craft Technologies, developed with the aim to engage non-engineering, non-computer science major students in a series of projects that would change the way they thought of and used computers and electronics. Over the semester, students created a series of five semi-structured projects targeted to teach them techniques for using and understanding conductive materials (thread, fabric, yarn, wire, etc.), basic programming, human-computer interaction, and electrical properties (e.g. resistance, short circuits, polarity, etc.). The course culminated in a sixth and final creative project of each student’s choice with an accompanying Instructables.com entry to

provide detailed instructions and pictures about their final project to a broader audience online. Each project was intended to support connected learning by encouraging self-expression and customization within constraints designed to teach specific techniques and concepts. Thus our primary questions in this paper regard if and how connected learning happened for individual students, and whether and how it mattered to them.

Background

Within a framework of connected learning we recognize that learners live in multiple, intersecting spaces that may reflect different values, repertoires of practices, and even languages (e.g., Fields, 2010; Gutiérrez & Rogoff, 2003). Thus connected learning should reach across multiple spheres of learners' lives, including family, school, hobbies, and other communities to which people belong (Ito et al, 2013). Building on the Connected Learning report (*Ibid*), we suggest that in the context of creating with computational media, such as craft technologies, connected learning can reach learners in four distinct ways. First, design tasks should encourage self-expression, allowing learners to draw on their interests and longstanding preferences (Azevedo, 2011). Second, there should be low barriers to access new knowledge; it should be approachable with multiple points of entry (Margolis & Fisher, 2002). Third, learners should be able to draw on new and existing social supports for learning. This builds on pre-existing relationships from multiple environments as well as new relationships in local communities and through broader networks afforded by digital media (Grimes & Fields, 2012). Last, connected learning environments should encourage learners to make meaningful connections to their cultural backgrounds and their prior expertise whether that expertise was developed at home, with friends, or in school (Moll, Tapia & Whitmore, 1993).

In general, discussions of connected learning focus primarily on children and youth with an emphasis on informal, out-of-school settings. In this paper we argue that college students and adults may benefit just as much from connected learning, especially with computational media, as children and youth, and that formal settings of learning should not be neglected in conversations of connected learning. Indeed, the very studies that helped to generate theories of learning as situated in specific contexts and that recognize that people participate in multiple social worlds were based on studies of *adults* in workplaces, local communities, and self-help groups (e.g., Lave & Wenger, 1991; Wenger, 1998). Indeed, college-age students and adults may have an even greater need for engaging connectively in designing with computational technologies than children and youth. Many adults generally have strong skill sets in their areas of expertise, whether at work or at home, providing rich potential to build on existing practices, knowledge, relationships and simple common sense (Rose, 2012). For adults who are not already experts with technology or engineering, creating with digital technology can seem intimidating. They may already be committed to non-science or tech focused majors or careers with less time to engage in new pursuits. Many see children's and youths' facility with tech gadgets as a sign that the younger generation are fluent "natives" with digital technologies (Prensky, 2001), implying that older generations are the outsiders. Yet in a time when workers change jobs frequently (U.S. Department of Labor, 2003) and we have need of effective citizens who have the knowledge and skills to live and act in a complex, diverse world where digital technology permeates businesses, homes, hobbies, and civic activism, adults may have just as much need of learning to design with digital technologies as children and youth.

Why do we so rarely consider college students and adults in relation to connected learning and creation with computational media? One reason may be that classic psychological theory holds identity as an achievement of adolescence (Erikson, 1968). However, if we take a situated view of identity (Holland et al, 1998), we are inherently compelled to admit that identities and learning emerge and change over time at all ages. Unfortunately, only a small set of researchers have attempted to study the existing practices of "working people" in their attempts to make sense of, manage, and learn about technological changes (Sawchuk, 2003, p. 3). Research that we do have demonstrates that adult workers learn about computers across a range of social settings (Sawchuk, 2003) and often do so through exchanges of hands-on skills (Golding, 2011), making them especially relevant audiences for connected learning with craft technologies based on hands-on skills. In the following sections we discuss ways that students in the Craft Technologies course engaged in connected learning through designing with novel computational media. We feature three students whose trajectories of participation reflected broader trends in the class in developing new views of themselves as programmers and designers with computational media. In doing so they drew on pre-existing skills and knowledge developed primarily with families, hobbies, and community groups, built new and strengthened existing relationships, and integrated their new learning with longstanding preferences.

Context and Methods

Craft Technologies" was a brand new university-level semester-long course taught by the first author (Lee & Fields, 2013). The course was opened to both graduate and undergraduate students from any department in the associated university. Students from Art, Communications, Music and Education fields enrolled. The initial and final enrollment for the course was 20 (12 undergraduate students, 8 graduate students; 16 women, 4 men), at least half of whom were in their mid-twenties to mid-forties in age.

The course focused on hand crafts as a particularly promising entry point for novices into computing and hardware. To this end two of the projects utilized the LilyPad Arduino (Buechley & Eisenberg, 2008), a sewable computer with inputs and outputs and an accompanying set of sensors (e.g. light sensors, accelerometers) and actuators (LEDs, sound buzzers, vibrator boards), and one project used the new Makey Makey. While many students focused on sewing as a primary crafting technique, many others successfully integrated prior expertise with wood working, origami, or sculpting. Students also researched the properties of novel conductive craft materials like different conductive yarns, threads, fabrics, tapes, and paints, as well as everyday materials like cooking sheets and aluminum foil. They utilized these in conjunction with traditional (generally non-conductive) craft materials in order to create sensors that could sense pressure, stretching, touching, etc. Thus the course introduced students to basic computing and circuit design as well as material properties. At the end of the semester students shared their projects in a Student Showcase, where members of the University and community were invited to attend.

Data come from the experiences and designs of 18 of the 20 undergraduate and graduate students enrolled in the course, focusing on eight women students who agreed to participate in pre/post interviews (approximately 20-40 minutes). The pre interviews focused on students' prior histories with crafting, programming, and circuitry, as well as overall feelings about personal efficacy in each of these areas. Post interviews focused on experiences with the course, students' projects, and identity-based questions such as "what/who is a programmer?" The interviews were recorded and transcribed. Students also completed a pre/post survey about interests and expertise in several key areas related to e-textiles, as well as weekly reflective blog posts about the processes of making their projects. We conducted a two-step open-coding analysis (Charmaz, 2000) of interviews; emergent coding focused on self-perceptions, interests, abilities, and relationships as exhibited in their projects.

Results

In this section we concentrate on the experiences and shifting identities of three participants that illustrate connected learning through designing with computational media. Although each student's case is unique, the narrative cases are illustrative of broader trends revealed across students in the course, especially in the ways students built on old and new relationships, pursued longstanding personal interests in and beyond the course, integrated new with older knowledge, and considered new possibilities in the future.

Connected Cultural Learning: Naomi

One goal of connected learning is to provide new educational opportunities by supporting interest-driven learning and self-expression. An outcome of this goal is to facilitate learners' creation of new identities and vision of expanded possibilities for what they can do, a goal that we believe is equally important for adults as well as children. Our first case study, Naomi, illustrates ways that she was able to engage in self-expressive, connective activities through her new learning to design with craft technologies in the course. In particular, Naomi used projects in the course to connect to her mother's knowledge, her children's interests, and her local church community's do-it-yourself priorities. As a result, she was able to develop a new identity as a burgeoning programmer who had confidence in her abilities to program and saw applications for that in her job as well as at with her family and community.

Naomi, a graduate student and single mother of three, initially expressed being intimidated by programming and more "technical" skills at the beginning of class. In her pre-interview, Naomi indicated she both enjoyed and felt comfortable with crafting, and mentioned how her mother had "always been hands-on with crafts" and passed those skills on to her children. Her mom, a seamstress, helped with her 4-H classes when Naomi was a child. Naomi fondly mentioned having regular girls' nights where the women of the family crafted and a traditional "big craft day" the day after Thanksgiving. From the onset of the class, Naomi's past experience in crafting gave her confidence in her ability to construct projects, even though she had no programming background. She slowly built her knowledge of programming, circuitry, and conductive materials, linking those new domains of knowledge with her prior knowledge of crafts. For her final project, Naomi decided to make an interactive wall hanging for her youngest daughter as a part of an ongoing process to re-do her children's rooms. She quilted the wall hanging backdrop from scratch, adding flowers which lit up upon touch, stars that played "Twinkle, Twinkle, Little Star," a fairy whose wings twinkled, and butterflies that lit up when someone triggered the sound sensor. Naomi fluidly integrated her new knowledge with prior skills, and also learned new knowledge beyond the course content, teaching herself how to use and program certain elements like a sound sensor that were not covered in class. She saw her newly developed hybrid knowledge as connecting to her mother, her local church, and her master's degree work, and planned to continue this heritage by teaching her children both traditional and new skills she had gathered together.



Figure 1. Various stages of Naomi's interactive quilt. She arranges the squares (left) to determine the layout. A close up of the fairy (middle) with twinkling wings. The final quilt (right) which plays music, and has sound and light sensors.

In a follow-up interview, after the Craft Technologies course had ended, Naomi expressed a new identity in that she now considered herself to be a programmer, because she had gained skills that she was able to transfer to another programming course she took. Naomi stated, "Remember, how I said I didn't consider myself a programmer? Well, maybe I do now...I was so much more confident...and I was the only one of the students who got a perfect grade on my final" She also indicated ways she would use her new-found skills for future crafting and programming projects, particularly as craft technologies opened "a whole new dimension" she can integrate when crafting with her daughters. Naomi also recognizes possibilities for crafting outside of her family. In a follow-up interview, she discussed how she would like to teach the light-up bracelet, the first activity in the course, to a group of 8-11 year old girls she mentors at her church. Additionally, Naomi's ability to program will be helpful in her career as a technology coordinator at her job, where coding often occurs. Naomi's transference of skills gained to other courses and to other practices outside of school was demonstrated in all interviewed students, and again, is indicative of how positive connected learning experiences can be for adults.

Connected Teaching: Fiona

Collaborative, peer support is another integral part of connected learning. This can occur as learners share and provide feedback to another, as well as when they provide guidance and expertise needed to accomplish shared and individual goals. An environment with strong peer support provides opportunities for learners to contribute in a variety of meaningful ways. In the space of the Craft Technologies course, students actively supported and shared expertise with each other. This occurred when a student skilled at sewing helped students with their stitch work, when students critiqued, discussed, and brainstormed about project designs, and as exhibited by this second case study, when an experienced community college teacher taught her classmates how to program.

Fiona, similar to Naomi, was a graduate student and mother of three. She had a background with computer programming, both as an undergrad and as a teacher at another college. She also enjoyed crafting activities such as sewing and cross-stitch, and drew from a family heritage of crafting, notably her mother and sisters who were skilled crafters. During the pre-interview, Fiona indicated she was not worried about either the crafting or the programming parts of the course, and hoped her prior experience in those areas would allow her to focus more on the electrical components, a subject was interested in.

In our pre-interview, and several times throughout the course, Fiona positioned herself away from a "programmer" or expert identity. She insisted she was not a programmer, despite her extensive experience with programming, both technically and in teaching. However, only a few weeks into the course, Fiona realized her programming and teaching background could be helpful for her classmates. She announced to her peers she was willing to help with any of their programming questions, and wrote her email on a class whiteboard with the note "Help with programming." Several students in the class did utilize her help and many acknowledged her in their blog posts and post interviews. During our post-interview, Fiona stated out of everything she experienced in the course, even over learning basic electrical engineering, she was most of proud of the classmates she taught to code. She stated she loved when she could see her classmates understand a programming principle and it gave her a "rush" to teach others how to program. She further went on to state how Craft Technologies had changed her view of how computer science can and should be taught. She stated,

So maybe I'm thinking you could take a CS 101 class and have the students build the circuit and tell it what to do. Instead of writing, "Hello World" in Python. I think in some ways, it gives a better understanding of what's behind in programming, and why you have to be so specific in programming.

This contrasted with her earlier view that teaching programming by building circuits was the “wrong way” to do it. Based on her experiences in the course, she acknowledged that craft technologies was a viable way to introduce programming, “... I've seen people who don't see themselves as a programmer, and don't seem interested in pursuing programming, and they've built and programmed circuits, and maybe you really can do it this way, and it works. It's been cool.” This shift in belief of how programming should be taught was one of the biggest takeaways Fiona took from the course. Because she was able to connect her identity as a teacher with her experience as a student in this course, she showed a marked change which could influence how she approaches teaching programming in the future. She indicated in future programming courses she taught, she would include a unit on building programmable circuits because she felt it was so valuable for novice programming students. Yet this striking change wasn't the only one Fiona had as a result of the course.

As mentioned earlier, several times through the semester Fiona denied she was a programmer, because she wasn't “paid” to write code for anyone. She thought that to be a programmer, one had to have corporate or institutional validation, removing any notion of programming as a hobby or for the love of it. However, by the time of our post-interview she had altered her definition of who and what a programmer is. She expressed that her love of programming, even if it was done only in a hobbyist setting, made her a programmer. This is a radical shift in her identity exemplifies how identity for adults isn't fixed, but can be fluid and dynamic. Fiona was both a beneficiary and an agent of connected learning. She helped manifest a collaborative, supportive environment for her peers, and in return, received a positive identity crystallization.

Connecting Art and Technology: Camille

Connected learning also attends to creativity and “creative work,” specifically as it relates to digital media. Ideal connected learning environments allow for learners to engage in interest-driven, creative, academically oriented endeavors which could eventually parlay into economic opportunities. For this reason, we argue adults, especially those in college, have great need for this type of learning experience. Our third case study centers on Camille, an undergraduate art student who was able to connect several, seemingly disparate, facets of her life into her course projects. As a result, she was able to learn new skills she could transition into making unique sculpture art both for school-based and outside artistic pursuits.

Camille identified strongly as an artist and a “maker,” and acknowledged her profound skills in these areas. An adult college student in her late twenties, she began college after several years in the Marine Corps and other jobs, and was pursuing an art major at the time of the course. Her interests focused on sculpture with ceramics, metal, and glass and she had a broad set of skills related to glass blowing and kiln making in addition to sewing, knitting, woodworking and crafting in general. On the side she enjoyed making creative projects at home with her husband, an electrical engineering student. Camille brought to the course a limited computing background that she associated both with her father, who was a computer programmer, and from her time in the Marine Corp, where she worked with small systems computers. Throughout the course Camille created projects which linked to each of these aspects of her life. She learned circuitry and programming skills which allowed her to forge connections with her husband's experience as an electrical engineer and to enhance the “making” they did together. Furthermore her experience in Craft Technologies gave her an opportunity to blend electronics and technology with her art making interests and skills.

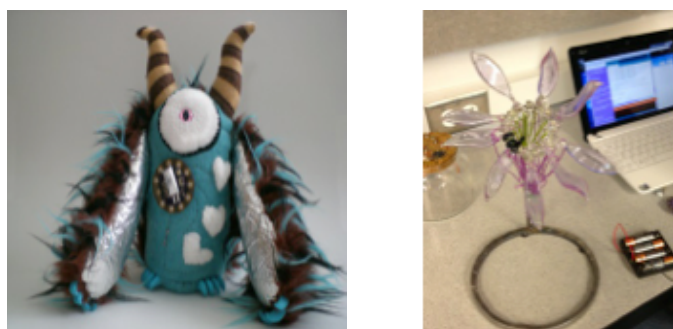


Figure 2. Two of Camille's projects. The art doll monster (left) uses conductive patches on the arms to turn on lights sewn into the hearts on its chest. The blooming flower (right) made using nitonel, which opens and closes.

In her pre-interview, she specifically mentioned her passion for creating “fun, goofy” art as a contrast to the typical fine art that is often produced by artists. This love of creating interesting, whimsical artifacts was actualized with the creation of many of her projects. For example, her project for the human sensor project enabled her to learn programming and circuitry techniques, and integrate those with sculpting and sewing skills to create an artifact that played to her interests in sculpture, art, and making (see Figure X). Her goal was to create a fine “art doll” monster, inspired by makers whose work she been studying on the Internet. She carefully selected the fuzzy fabric for the monster skin, hand-sewed every piece (including the striped horns) and sculpted

the claws. Camille took the project further than most students, studying persistence of vision in programming the LEDs to save battery power over time and adding a carefully curled resistor to improve the function of the project.

Camille's final project is a prime example of how she brought together past expertise and interests with new knowledge learned in the course to create new methods for artistic expression. Camille had long been interested in interactive art or moving sculpture, but had been unable to learn how to make it in her art courses. However, the connected learning environment of the Craft Technologies course allowed her to explore ways to bring her ideas to life. For her final project, she wanted to create a kinetic flower sculpture that bloomed. In order to do this, she taught herself to use nitinol, a dynamic material similar to glass, which can be programmed to change shape. In the post-interview, Camille indicated she would further explore this juxtaposition between art and technology, showing while craft technologies was a relatively new endeavor, it was also built upon a longstanding preference for pushing the boundaries with art and making (Azevedo, 2011).

Additionally, Camille received validation for her creations widely outside of the Craft Technologies course. She posted a tutorial for making the flower sculpture on Instructables, a do-it-yourself social networking forum where users share detailed tutorials of projects of all kinds from woodworking, robotics, and of course, craft technologies. Within days of posting her tutorial, Camille had several thousand views. Currently, the tutorial has over 40,000 hits. At a recent Maker Faire, Camille was recognized by the Instructables community for her tutorial's contribution to the website. This is a particularly striking example, as prior to the Craft Technologies course, Camille had often looked at tutorials on Instructables, but had never posted her own. She has since posted a second Instructables entry and has become a more visible and contributing member of both the Maker and Instructables communities. She will also be a TA for second offering of the Craft Technologies course in Spring 2014, giving her a chance to parlay her expertise into teaching and supporting new students. Further, Camille included her art doll monster in a portfolio for a scholarship, which she was awarded. Undoubtedly a connected learning environment gave Camille relevant, educational opportunities to learn and apply new skills by integrating them with existing practices.

Discussion

In this paper we have shown how college age and adult women changed the way they viewed themselves as well as how they took up new practices that integrated computational media with existing interests through connected learning in the Craft Technologies course. All interviewed students acknowledged that they considered themselves—if only as beginning or hobbyist—programmers by the end of the course. Further, over half of the interviewed students indicated they had future plans to continue programming, either by taking more courses, making or teaching. As none of these students, with the exception of Fiona, came from a computer science background, the course presented a unique chance for them to navigate unfamiliar academic terrain, and to “learn new crafting and programming practices of through creative and self-expressive designs. In other words, the college-age and adult students “crafted” new versions of themselves along with all the projects they made in class.

Amidst the many influences of parents, children, hobbies, and prior education on students' connected learning in the context of this course, local, cultural knowledge played a particular role in legitimizing students' prior expertise in the Craft Technologies course. In Utah and neighboring states, the setting of the course, there is a strong emphasis on do-it-yourself practices, largely influenced by the dominance of the Church of Jesus Christ of Latter-day Saints (LDS church). This emphasis on DIY crafting and related practices is not unique to the LDS church nor its members, but it has certainly become codified as a practice and spilled over into the milieu of making in Utah. For women, this spirit of do-it-yourself is predominantly exhibited as crafting with soft materials (paper, textiles), and takes place either at auxiliary church meetings or as a part of family practice. All of the participants interviewed were influenced by this making culture, whether overtly from being raised in the LDS church (the majority of students) or tacitly as members of the broader community. Several participants specifically mentioned learning to sew or craft because of church activities. Others discussed how they would participate in “crafting weekends” with their moms and sisters. The unique fusion of the “soft and tactile” with programming and circuitry concepts allowed students to build on their past crafting experience, while learning new knowledge. In a broader sense, the Craft Technologies course also validated students' crafting and DIY activities which are often pursued outside of formal learning environments.

Earlier we described the importance for engaging not just children but adults in connective learning and becoming through design practices with computational media. As many students in the class expressed, opportunities for self-expression and creativity can be very limited in the lives of busy working adults—kids are ‘allowed’ this kind of play (in many circles at least) but adults are often excluded from this through commitments to work and family. The course opened up a space for older students to pursue personal interests and learn new techniques. These opened up new identities as well as new skills that were relevant in workplaces, hobbies, and with families. Beyond this there may be other benefits to engaging older students and adults in learning with computational media. As hinted in the narratives about Naomi and Fiona, parents,

teachers, and community leaders have strong influences on the opportunities given to children and youth, be it as counselors, teachers, mentors or as parents. Their ideas about children have significant effects on how children view themselves in areas like science and math, with implications for their decisions about what careers and experiences to pursue (Yee & Eccles, 1988). Our research opens up possibilities of engaging adults in learning to design with new technologies in ways that may reshape not only their own identities and practices but also their views of younger students' capabilities and even their willingness to engage in new forms of teaching and mentoring.

References

- Azevedo, F. (2011). Lines of Practice: A Practice-Centered Theory of Interest Relationships. *Cognition and Instruction*, 29, 147-184.
- Banks, J.A., Au, K.A., Ball, A.F., Bell, P., Gordon, E.W., Gutiérrez, K., . . . & Zhou, M. (2007). *Learning in and out of school in diverse environments*. Seattle, WA: The LIFE Center, University of Washington.
- Blikstein, P. (2013). Digital Fabrication and 'Making' in Education: The Democratization of Invention. In J. Walter-Herrmann & C. Bü ching (Eds.), *FabLabs: Of Machines, Makers and Inventors*. Bielefeld: Transcript Publishers.
- Brickhouse, N. W., & Potter, J. T. (2001). Young women's scientific identity formation in an urban context. *Journal of research in science teaching*, 38(8), 965-980.
- Buechley, L., and Eisenberg, M. (2008). *The LilyPad Arduino: Toward wearable engineering for everyone*. *Wearable Computing Column in IEEE Pervasive*, 7(2), 12-15.
- Buechley, L. (2013, October). Thinking about making. Keynote address at the annual FabLearn 2013 Digital Fabrication at School conference. Palo Alto, CA
- Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 509-535). Thousand Oaks, CA: Sage Publications.
- Erikson, E. H. (1968) *Identity: Youth and crisis*. Norton: New York.
- Fields, D. A. (2010). *Trajectories of identification across social spaces: Intersections between home, school, and everyday settings*. Unpublished dissertation. University of California, Los Angeles.
- Fisher, A., & Margolis, J. (2002). Unlocking the clubhouse: the Carnegie Mellon experience. *ACM SIGCSE Bulletin*, 34(2), 79-83.
- Golding, B. G. (2011). Social, local, and situated: recent findings about the effectiveness of older men's informal learning in community contexts. *Adult Education Quarterly*, 61(2), 103-120.
- Grimes, S., & Fields, D.A. (2012). *Kids online: A new research agenda for understanding social networking forums*. The Joan Ganz Cooney Center at Sesame Workshop.
- Gutiérrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational researcher*, 32(5), 19-25.
- Hill, C., Corbett, A., & St. Rose, A. (2010). *Why So Few? Women in Science, Technology, Engineering, and Mathematics*. Washington, DC: AAUW.
- Holland, D. Lachiotte, W., Skinner, D., & Cain, C. (1998). *Identity and Agency in Cultural Worlds*. Harvard University Press, Cambridge, MA.
- Ito, M., Gutiérrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K., . . . & Watkins, S.C. (2013). *Connected Learning: An Agenda for Research and Design*. Irvine, CA: Digital Media and Learning Research Hub.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Lee, V. R., & Fields, D. A. (2013). A clinical interview for assessing student learning in a university-level craft technology course. Paper presented at the III Digital Fabrication in Education Conference (FabLearn 2013), Stanford, CA.
- Moll, L. C., Tapia, J., & Whitmore, K. F. (1993). Living knowledge: The social distribution of cultural resources for thinking. *Distributed cognitions: Psychological and educational considerations*, 139-163.
- Nasir, N. I. S., & Hand, V. (2008). From the court to the classroom: Opportunities for engagement, learning, and identity in basketball and classroom mathematics. *The Journal of the Learning Sciences*, 17(2), 143-179.
- Prensky, M. (2001). Digital Natives, Digital Immigrants. *On the Horizon*. NCB University Press, 9(5).
- Sawchuk, P. (2003). *Adult learning and technology in working-class life*. Cambridge: Cambridge University Press.
- Turkle, S., & Papert, S. (1991). Epistemological pluralism and the revaluation of the concrete. *Constructionism*, 161-192.
- Searle, K. A., Kafai, Y. B. & Fields, D. A. (2013, April). Building new clubhouses: Opening doors to computing for girls and boys with electronic textiles. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

- U.S. Department of Labor. (2003, September). FY 2003-2008 strategic plan. Retrieved February 13, 2007, from http://www.dol.gov/_sec/stratplan/strat_plan_2003-2008.pdf.
- Yee, D. K. & Eccles, J. S. (1988). Parent perceptions and attributions for children's math achievement. *Sex Roles*, 19(5/6), 317-333.

Report and Reflection Papers

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

Billie Eilam, University of Haifa, Israel, beilam@edu.haifa.ac.il

Abstract: A comprehensive daily-report instrument supplying metacognitive prompts was used to promote a high-achieving group of 5 ninth graders' collaborative planning, execution, monitoring, and adjustment of a long-term, open-ended ecological inquiry about a live ecosystem (greenhouse). Prompts guided the group's discussion and documentation of each laboratory session's planned and executed behaviors, aiming to trigger students' metacognitive cues about gaps and progress along the yearlong project. Fine-grain analysis of session transcripts, students' written documentation, and videotapes for this single group case yielded a recurrent pattern of within-session behavior and evidence for the group's use of metacognitive knowledge and skills over time as well as salient difficulties. Discussion focused on the role of specific components in this student centered self-regulated long-term learning environment – the daily reports, the teachers, and the live ecology inquiry context. Implications were discussed for future research and long-term inquiry-driven science education design.

Major Issues Addressed

The present study addresses metacognitive planning and monitoring, using a specially designed prompt for enhancing learners' Metacognitive Awareness (MA) while performing a long-term, authentic open inquiry of live complex ecosystems, in an environment of full personal and academic autonomy. Such an environment has the potential to activate metacognitive knowledge (Chinn & Hmelo-Silver, 2002), overcome some inherent difficulties and develop MA (Eilam, 2002; Eilam & Aharon, 2003). The present study examined learners' real-time on line MA while planning a complex long-term yearly ecosystems inquiry, executing these plans, monitoring outcomes in light of set goals and adjusting thoughts and actions. Metacognitive prompts (DPSRI – Daily Planning Self-Report Instrument) were provided for facilitating a case group of high-achieving 9th graders' metacognition as evidenced through their observed verbalizations, behaviors and written self-reports.

Potential Significance of the Study

Thus far, metacognition and monitoring have been mostly examined in shorter computer-based simulations. The present environment of a long-term inquiry enables the on line fine-grain studying of MA due to the continuous, dynamic, and interactive process of thinking and rethinking that involves flexibility, judgment, and metacognition, applied in response to changes occurring in the researched system. I also suggest a novel prompt/tool for promoting, scaffolding, and studying MA, planning and monitoring. Presented on paper, the Daily Planning Self-Report Instrument (DPSRI) scaffolds learners' MA. Last, although planning is a core element in metacognition, knowledge about students' planning is deficient.

Theoretical and Methodological Approaches

Rooted in the constructivist and sociocultural perspectives, a group case study method was selected to investigate students' MA. A case is a specific, complex, functioning, well-bounded, integrated system, characterized by patterned consistent behavior that is clearly differentiated in some of its features from other systems (Stake, 1995).

Participants

The single distinct group (one of eight in the classroom) of 5 high-achieving and motivated ninth graders (two girls and three boys), performed the ecological inquiry collaboratively with the final distant goal of producing a scientific report of their research. The group was unique in its composition and in its personal and social boundaries, distinct patterns of interactions and collaboration, and their different inquiry boundaries. No roles were assigned and no persistent “leaders” emerged. Because MA is related to domain knowledge (Alexander et al., 2011) this high achievers group was expected to provide a worthwhile case for analysis of students' MA.

The Learning Context

Laboratory Schedule and Physical Layout

The inquiry was performed over 30 sessions, in weekly session of 3 hours (including breaks), in the school science laboratory. The group occupied one table of the 10 contained in the lab. Students collaborated on the

DPSRI. Students built two greenhouse ecosystems (control and experiment). Members had to learn about their selected ecosystem and its components, formulate a broad topic of interest, focus gradually on possible research questions, raise hypotheses based on their knowledge of ecology and biology, choose one hypothesis for examination in their greenhouse, and select its biotic and abiotic components. The case group examined the "greenhouse effect" by increasing the CO₂ amounts and measuring the system varied variables (e.g., temperature, number and behaviors of organisms and related processes.) Initially, the group collected a set of all possible measures from both the experimental and control systems (e.g., organisms' size and weight; plants' number of leaves; bacteria samples from the soil, air, and leaves; temperature; pH) over a 1-month period. Students designed their own representations like tables, graphs, and drawings to document the lengthy data collection. Then, the group introduced CO₂ from a special container into the "experimental" system only. Students cared for their systems, collected and processed data, generated conclusion and submitted a written report.

Student Autonomy

In order to enable students' free planning and managing of the inquiry, the onus of responsibility was shifted entirely onto students to plan, perform, and manage their long-term inquiry as they saw fit, after considering possible alternatives and choosing among them according to self-determined goals and resources. Group members were granted complete autonomy over their behavior, including decisions regarding taking breaks, coming to sessions, or homework. Students were never reprehended for anything but disrupting a classmate's learning. They had to present their research in a school conference and submit a final written scientific report.

Support Mechanisms

Students have never experienced such a lengthy SR project, and teachers' external feedback about the quality of regulation and the inquiry performance were mostly avoided to enable their autonomic decisions. Therefore, support and scaffolding of MA were provided as: (a) The DPSRI (see below) including practice of its use and a textbook about inquiry and its skills; (b) Teacher guidance; (c) Theoretical courses in ecology, taught traditionally and concurrently with the 3 weekly laboratory hours; (d) Students' files.

Scaffolding Tools to Provide Metacognitive Prompts

Prompts were developed based on literature concerning SRL, prompts, cues and metacognition. It required learners' continuous real-time reporting about their performance in each session and enabled calibration of thoughts and behaviors by planning and monitoring plans execution. *Planning*. Planning is intentional, promotes goal attainment under specific circumstances, enhances performance, allows for anticipation of consequences by generating external representations of future behavior prior to enactment, which guides future planning in similar situations, prevents some mistakes, and involves the employment of metacognition to transform learning intentions into action plans. Planning is enhanced by a less structured environment (Jordan, Ruibal-Villasenor, Hmelo-Silver, & Etkina, 2011), but is particularly difficult in it because of the unknown variables and the uncertainty about actions' results (Allen, Hendler, & Tate, 1990). Concrete and attainable goals are set following the examination of the task's features (on-going inquiry) and serve as proximal regulators that emerge from the decomposition of distal goals. The initialization of relevant actions is ensured by selecting a set of activities, determining their enactment order, and coordinating these actions to achieve the stated goals. Planning involves awareness of one's own metacognitive knowledge and resource allocation (e.g., time, space, equipment) while considering the affordances and constraints of the specific circumstances, as well as the efforts that need to be invested to attain the goal (Gollwitzer, 1996; Prins, 2002; Ward & Morris, 2005). *Monitoring*. If execution of plans is not monitored, evaluated, and revised accordingly, no change in behavior would be possible. Metacognitive monitoring enables control and management over the effectiveness of learners' executed actions and cognition that may affect behavior. Monitoring activity relies on learners' ability to perceive internal and/or external cognitive, affective, and/or situational cues (Carver & Scheier, 1990). Perceived cues are triggered by comparisons between current and goal state and constitute feedback that initiates learners' MA, which in turn produces a behavior to reduce perceived gaps by changing plans or parts of them. *Time management*. Time management is a core component of planning and monitoring and is subjectively experienced; students' diverse time orientations may shape their mode of engagement with tasks (Duncheon & Tierney, 2013). Very little has been published on this issue as related to scientific inquiry. During a long-term, complex process of inquiry, time constraints call for increased efficiency and require making metacognitive decisions concerning choices among alternative actions and time management. Cues suggesting an unexpected rate of progression toward goals may influence planning.

The Daily Planning Self-Report Instrument – DPSRI

Completion of the DPSRI included (see Figure 1): (a) Reporting concrete, limited goals that could be achieved during the session; (b) Reporting suggested plans in the "suggested plan" column, by: (i) selecting and

describing context-dependent, accurately defined activities (e.g., to read about the frog) and categorizing them using the ready activity segment labels pool (e.g., gathering information); (ii) determining the sequence of activity execution; (iii) allocating time to each activity; and (iv) determining the setting for performing the activity (e.g., cooperatively, individually); (c). Reporting enacted performance: Students completed the “enacted plan” column separately for each planned segment, immediately after its execution to decrease cognitive load and increase accuracy of reporting. These “enacted” reports included the recording of (i) the performed activity’s description and segment label, (ii) their actual sequencing, (iii) their duration, and (iv) the work setting used; (d). Assigning homework.

DAILY PLANNING SELF-REPORT					Student Name: Tim Group: White Date: January 24 Session Goals: Complete the principles and raise a research question			
Hour	SUGGESTED PLAN				ENACTED PLAN			
	Activity description	Segment label	Time allocated	Setting	Activity segment	Segment label	Time used	Setting
8:15	Planning	Completing the DPSRI		Group	Ask the teacher about the Temp.	Information gathering		Group
8:25	Talk to the technician	Technical		Indiv.				
8:35	Measuring	Technical: Taking care of system		Group	Caring for the biotic,	Technical: Taking care of system		Group
8:45		Information gathering: Collecting data			Planting plants			

Figure 1. An Example of The Daily Planning Self-Report

The DPSRI self-report differs from other instruments in several key features that increase its validity and achieve additional benefits: (a) *Activity description and categorization using a pool of ready segment labels*, for increasing objectivity and creating a uniform “language” of self-reporting, which facilitated student comparisons between the planned and enacted columns within a session as well as comparisons of reports between sessions, thereby enabling cue input and monitoring of plan effectiveness, promoting MA. In addition, this uniformity and objectivity increased coders’ inter-judge agreement; (b) *Reporting on-line of enacted activities immediately following their completion*, rather than after the entire 3-hour session, thus reducing the effect of cognitive load and increasing reporting accuracy; (c) *Legitimacy of reporting any enacted activity, thus reducing social desirability*; (d) *Monitoring planned-enacted identified gaps, thus providing evidence of MA*. Because reasoning about gaps causes and making decisions regarding necessary future behavioral changes, is difficult and require MA, changes in planning along the year may be indicative of MA and enables deepened insight into students’ behaviors. In addition, successful enactment must account for past similar experiences and may even include monitoring through the examination of previous DPSRIs. The intrusiveness involved in externalizing students’ actual behaviors probably also influenced such behaviors by increasing MA of certain planning and time management elements. A 90% inter-rater reliability while rating 20% of the reports was accepted.

Data Collection and Analysis

Group's working modes and interactions over the year were recorded by a video camera and discourses were fully transcribed. Data were analyzed from these transcripts and members' completed DPSRI and videotapes were consulted in cases of ambiguity. Transcript analysis focused on utterances related to the DPSRI, which yielded ample evidence on students' MA.

Identifying the Group's Content and Sequence of Work Patterns

The aggregated selected utterances were repeatedly read by two experts in biology and in learning and were scrutinized for internal consistency regarding the group's work patterns over the year as related to completing the DPSRI. Central student activities were identified in their chronological order in each single session, and then activity sequences were compared between sessions (inter-judge agreement of 86% regarding the group's behavioral pattern). Work pattern showed very few deviations despite changing contexts.

Report Analysis

The written DPSRI materials were read carefully by the same two judges to identify explicit direct and indirect evidence of students' MA, as manifested by their planning, gap monitoring, reasoning about those gaps, and adjustments of subsequent plans. The DPSRIs were analyzed within and across the predetermined phases of self-reporting (i.e., goals, suggested plans, enacted performance, and assigning homework). Reports were

analyzed within each single session, and changes in behaviors were traced across sessions (83% inter-judge agreement).

Major Findings, Conclusions, and Implications

Group Behavioral Pattern

Overall, a consistent 4-part pattern emerged: (a) *Preparations for planning*: Browsing through the inquiry textbook and examination of the systems to decide what needed to be done (e.g., exercises) and be studied (e.g., ecology knowledge). (b) *Collaborative planning and recording plans in the DPSRI*. After planning, students copied from each other plans to yield identical reports. They filled the goals supported by the inquiry textbook (“We saw what we have to do. We have to gather materials to focus the topic onto a limited subject for inquiry”), and the suggested plan column by discussing activity segments and settings but rarely arguing about time allocation. Categorization of activities into label segment was acquired over few sessions. Initially, planning was mostly intuitive, based on past experiences and school habits and norms rather on monitoring. Around the fifth session students exhibited MA probably evolving from perceiving DPSRI cues. Time pressures were explicitly expressed and affected planning. (c) *Enacting plans*: Students reminded each other to fill in this column after completing each activity. (“G: Wait, before you go out, report the enacted. R: Right, let’s fill out the daily report.”) Most sessions’ discourse exhibited students’ implicit comparisons between the planned and enacted, evidencing students’ perception of metacognitive cues (D: “We’re behind, look at the yearly report. T: But the whole idea is that you move forward in the inquiry at your own pace.”). They verbally expressed their perceived results of such comparisons (time lags, gaps in activities), and tried to repair them by changing plans. From the eighth session on, they accepted gaps as inevitable but monitored their thinking and progress in attempt to improve future plans. Some activities (reading and exercises) were always enacted individually, whereas inquiry related activities were performed in collaboration. To save time students frequently performed different activities in parallel within the same activity segment.

Group Written Reports Over Time: Metacognitive Awareness

To reveal MA as manifested by students’ written reports, the DPSRIs were compared and analyzed over time. Because their reports were almost identical only one of them was examined. Findings are presented according to the DPSRI phases, indicating what can be construed as MA for each.

(a) *Setting proximal goals*. MA was evidenced in realistic, relevant, logically sequenced goals that could be operationalized. These were usually related to structured and linear tasks (e.g., to gather specific information). Ill-structured tasks resulted in general goals (e.g., design the experiments) that impeded students’ ability to enact plan successfully. From midyear on, goals demonstrated increased MA of situational factors (e.g., time) and future group need (e.g., devoting a lesson for fixing the system). (b) *Suggested and enacted plans*. MA was evidenced in the accurate definition of easily applicable and limited scope activities that were almost always categorized correctly to the general segment labels and by students’ added new customized segments according to identified needs. Planned-enacted gaps initiated MA monitoring, and identification of the problem that caused it expressed in the enactment of an unplanned segment; changing break time to be after a discussion was completed because “a discussion after we come back is never the same”; or inserted unplanned “teacher consultations” after being aware of their lack of understanding during a planned discussion. (c) *Allocating time resources*. MA was evidenced in students’ comments about a temporal gap and along time. For example, when too many activities were previously planned for a particular time unit, they reduced them in the following plans, when plans did not include break to save time and students felt hungry and tired, they introduced breaks in the following session plan after discussing the problem; acknowledging that they have learned that a new unfamiliar activity requires more time than familiar ones; administering homework and meeting after school to compensate for time lags; consciously allocating session time for devising the suggested plan. (d) *Setting*: A preference for a collaborative work was found probably due to the challenging task. However, MA was evidenced when students flexibly shifted to individual work when time constraints cues were perceived. (e) *Homework were Assigned* initially, when no gaps were perceived, due to habits and norms. MA was evidenced when time gaps have been perceived and homework were expected to save time. In addition to developing MA students demonstrated several difficulties such as a deficient ability to identify missing domain and procedural knowledge in their own cognition, evidenced by the longer-than-anticipated time required for processing information and by the help-seeking segments that were introduced only later, while enacting the plan; difficulty in operationalizing perceived cues into actions that would improve performance; unsuccessful attempts to strike a balance between different dichotomous factors like, time constraints and activities to be carried out, deep narrow understanding and superficial broader understanding, individuals’ needs and group needs, or between personal preferences and more efficient options. Such balance can only develop through diverse experiences of learning to navigate in complex learning environments and to negotiate among its components.

Conclusions and Implications

A deficient knowledge exists about planning and monitoring as expressing MA applied in a long-term inquiry. Evidence of MA are difficult to observe. The present instrument explicated such evidence and hence promoted understanding of these phenomena. Researchers' efforts should be continued and directed toward the development of refined instruments that may capture the application of different metacognitive components. The current findings demonstrate the possibility of providing students with opportunities recommended by many researchers of science education: namely, to experience an authentic self-regulated inquiry that will promote the development of students' MA and science. The findings suggest that in spite of the many difficulties involved, high-achieving students at this age can cope with the complexity of the environment, utilize its sources, and gain academically and personally from their experiences. However, such a project requires teachers' expertise, students' full autonomy, and length of time. To achieve better planning abilities, and to construct accurate and flexible knowledge representations of planning, students have to experience and train in learning environments that promote their ability to orchestrate diverse activities while performing complex, long-term tasks.

References

- Alexander, P. A., Winters, F., Dinsmore, D. L., & Parkinson, M. (2011). The role of domain knowledge in self-regulated learning. In B. Zimmerman, & D. Schunk (Ed.), *Handbook of self-regulation of learning and performance*. New York: Routledge.
- Allen, J. F., Hendler, J., & Tate, A. (1990). *Readings in planning*. San Francisco: Kaufman.
- Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, 97, 19-35.
- Chinn, C. A., & Hmelo-Silver, C. E. (2002). Authentic inquiry: Introduction to the special section. *Science Education*, 86, 171-174.
- Duncheon, J. C., & Tierney, W. G. (2013). Changing conceptions of time: Implications for educational research and practice. *Review of Educational Research*, 83(2), 236-272.
- Eilam, B.. (2002). Strata of comprehending ecology: Looking through the prism of feeding relations. *Science Education*, 86, 645-671.
- Eilam, B. (2012). System thinking and feeding relations: Learning with a live ecosystem model. *Instructional Science*, 40(2), 213-239.
- Eilam, B., & Aharon, I. (2003). Students' planning in the process of self-regulated learning. *Contemporary Educational Psychology*, 28, 304-334.
- Ge, X., Chen, C. H., & Davis, K. A. (2005). Scaffolding novice instructional designers' problem-solving processes using question prompts in a web-based learning environment. *Journal of Educational Computing Research*, 33(2), 219-248.
- Gollwitzer, P. M. (1996). The volitional benefits of planning. In P. M. Gollwitzer, & J. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 287-312). New York: Guilford Press.
- Hmelo-Silver, C. E., Surabhi, M., & Lei, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *Journal of the Learning Sciences*, 16(3), 307-331.
- Jordan, R. C., Ruibal-Villasenor, M., Hmelo-Silver, C. E., & Etkina, E. (2011). Laboratory materials: Affordances or constraints? *Journal of Research in Science Teaching*, 48(9), 1010-1025.
- Pieschl, S. (2009). Metacognitive calibration: An extended conceptualization and potential applications. *Metacognition Learning*, 6, 205-211.
- Prins, F. J. (2002). Search & see: The roles of metacognitive skillfulness and intellectual ability during novice inductive learning in a complex computer-simulated environment (pp. 15-49). Leiden, The Netherlands: Print Partners Ipskamp.
- Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, 36, 111-139.
- White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition & Instruction*, 16(1), 3-118.

Investigating the Effect of Curricular Scaffolds on 3rd-Grade Students' Model-Based Explanations for Hydrologic Cycling

Laura Zangori, University of Nebraska-Lincoln, Lincoln, NE 68583, laura.zangori@huskers.unl.edu

Cory T. Forbes, University of Nebraska-Lincoln, Lincoln, NE 68583, cforbes3@unl.edu

Christina V. Schwarz, Michigan State University, East Lansing, MI 48824, cschwarz@msu.edu

Abstract: Opportunities to generate model-based explanations are crucial for elementary students yet are rarely foregrounded in elementary science learning environments, despite evidence that early learners can reason from models when provided scaffolding. We use a mixed methods design to investigate the comparative impact of embedded visual and written prompts (e.g. curricular scaffolds) on 3rd-grade students' formulation of model-based explanations for the water cycle. Students from six 3rd-grade classrooms generated models of the water cycle during an 8-week water unit. Findings suggest that students in the scaffolded group (n = 120) more frequently represented sequences of water underground than the unscaffolded group (n = 112). However qualitative results indicate the scaffolded group was also less likely to generate model-based explanations from their sequences of 'hidden' water processes. We conclude that embedded curricular scaffolds may support students to consider 'hidden' components but, alone, may be insufficient for theory building.

Objectives and Potential Significance

Model-based explanation-construction in science learning environments involves students constructing an external representation to explain *how* and *why* a systems works (Bechtel & Abrahamsen, 2005; Lehrer & Schauble, 2010; Schwarz et al., 2009). The external representation is an abstraction of a process that represents system components, interactions, and connections. Students are then able to operationalize *how* and *why* through the underlying unobservable cause – the mechanism - of key processes that interconnect to define the system (Braaten & Windschitl, 2011; Machamer, Darden & Craver, 2000). Though generating model-based explanations is a powerful practice, it is rarely foregrounded in elementary science instruction or widely available science curriculum materials, despite evidence that students are capable of engaging in these practices *when supported to do so* (Gunckel, Covitt, Salinas, & Anderson, 2012.; Lehrer & Schauble, 2010; Manz, 2012; Schwarz et al., 2009). Most often when students engage with models in elementary science learning environments, they serve as illustrations, demonstrations, or summaries of processes, but are not used or considered as a *way of learning* (Abell & Roth, 1995; Windschitl, Thompson, & Braaten, 2008).

In this study, we draw upon Sherin and colleagues' emphasis on 'learning artifacts' (2004, pg. 406) to explore the comparative impact of curricular, task-based scaffolds on 3rd-grade students' model-based explanations about sequences underlying the water cycle, a complex system comprised of many constituent processes. The *Next Generation Science Standards* (NRC, 2012) identify the hydrologic cycle as a core conceptual strand across the elementary grades that anchors future learning about water-related phenomena (Gunckel et al., 2012). However, water-related phenomena are challenging for students (Henriques, 2002), particularly process sequences that underlie system dynamics. Yet, little research has been conducted on elementary students' reasoning about hydrologic cycling and, as a result, there is much to learn about how to optimally support early learners to use models to reason scientifically about the water cycle. This study is grounded in a learning performances framework (Forbes, Zangori, & Schwarz, 2014; Schwarz et al., 2009) that highlights mechanism and sequence as core elements of model-based explanations. Our questions are:

1. Do embedded curricular scaffolds impact students' model-based explanations for the water cycle?
2. If so, how do 3rd-grade students formulate model-based explanations for the water cycle when provided embedded curricular scaffolds?

Background and Theoretical Framework

The hydrologic cycle is a particularly challenging area for students because many underlying processes that comprise this complex geosystem, such as phase change and groundwater flow, do not lend themselves to unaided observation (Gunckel et al., 2012; Henriques, 2002). Early learners struggle to postulate unseen mechanisms for observable phenomena. Both evaporation and subsurface water flow are 'hidden' processes that even older students have difficulties conceptualizing (Gunckel et al., 2012; Henriques, 2002). Rich, learner-centered science learning environments should therefore be designed around multiple scaffolds, both curriculum-embedded and provided by the teacher through instruction, that work synergistically to address cognitive and practice-oriented learning outcomes (Sherin, Reiser, & Edelson 2004). Curricular scaffolds, here defined as visual and written prompts embedded within curriculum materials (McNeill & Krajcik, 2009), are an

important mode of providing cognitive supports to students. However, little research has been conducted to explore how to optimally scaffold early learners' use of models to reason about water systems.

To support both the design AND empirical study of model-centric elementary science learning environments, we have developed a comprehensive, domain-specific learning performance framework to account for students' model-based explanation-construction and conceptual understanding about *water in motion* (i.e. 'big ideas') and the scientific practice of modeling (Forbes et al., 2014; Schwarz et al., 2009). A core component of the framework is process sequences. Sequences represent the connections and the continuity between components of the process that comprise a complex system (Machamer et al., 2000). For example, when scientists attempt to understand or explain complex systems through modeling, the components of the system are connected with arrows (e.g., condensation → precipitation → evaporation). The arrows represent the activity for *how* and *why*, for example, the condensation component 'became' the precipitation component and so on. It is within the arrows that the mechanism – the underlying unobservable cause – is articulated. It is only in the understanding of what the arrows represent that the mechanism is understood (Bechtel & Abrahamsen, 2005; Machamer et al., 2000). Sequences are apparent in geosystems which are defined by overlapping and integrated processes creating dynamic complex systems. These processes are a function of sequential cause and effect, both observable and unobservable, in which one phenomenon impacts another. In order to understand dynamic systems, including the water cycle, early learners must be provided opportunities to both identify the sequences within this system of occurrences and postulate mechanisms that connect them to provide a foundation for a more sophisticated understanding of systems (i.e., systems thinking) that they will develop in later grades (Gunckel et al., 2012; Schwarz et al., 2009).

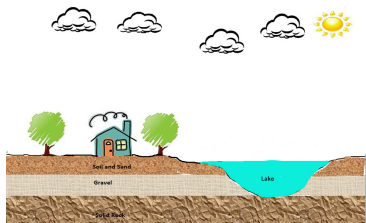
Research Design and Methodological Approach

This concurrent mixed methods study is situated within a broader design-based research program conducted as part of a 3-year, NSF-funded project designed to a) explore and promote 3rd-grade students' formulation of model-based explanations for hydrologic cycling through curriculum materials enhancement and instruction, and b) empirically investigate associated instructional and student learning outcomes. Six 3rd-grade classrooms were purposefully selected (Patton, 2001) based on teachers' teaching experience, student demographics, and their use of an existing, commercially-available, kit-based curriculum module about water.

Design

To afford student opportunities to formulate model-based explanations, two supplemental lessons were integrated into the curriculum module, one each at the beginning and end. Each lesson afforded students the opportunity to complete a modeling task that involved constructing a 2-D diagrammatic process model of the water cycle and responding to a series of prompts designed to elicit students' mechanism-based explanations for answers to the following question: '*where does the rain go when it reaches the ground?*'. Two different versions of the modeling task were employed. Students in three classrooms completed modeling tasks that included no written or visual prompts while students in the other three classrooms completed tasks with both verbal and visual prompts (Table 1). The modeling tasks were otherwise identical.

Table 1: Elements of scaffolded and unscaffolded student modeling tasks

	Unscaffolded	Scaffolded
Student Instructions	Use the box on the next page to draw a model of what you think happens to rain after it reaches the ground.	Use the box on the next page to draw a model of what you think happens to rain after it reaches the ground. <ul style="list-style-type: none"> • Include what you think are the very <u>most important things</u> that happen to rain when it reaches the ground • Include what you think happens <u>on top of</u> and <u>under</u> the ground when it rains • Show <u>why</u> these things happen to rain when it reaches the ground If helpful, use <u>words</u> and/or <u>numbers</u> to label parts of your model
Model Template	(Empty box)	

Data Collection and Analysis

All pre- and postunit student modeling artifacts were collected ($n=120$ scaffolded; $n=112$ unscaffolded) and scored using a rubric developed from an empirically-tested learning performances framework for elementary students' sensemaking about the hydrologic cycle (Forbes et al., 2014; Schwarz et al., 2009). We examined and scored the students' modeling tasks for the level of sophistication in their representations of sequences (e.g. water in the sky \rightarrow falls to the ground) for the big idea of *water in motion*. The scoring rubric is shown in Table 2. Scoring levels were identified and validated through empirical development of the learning performance framework (see Forbes et al., 2014). The unit of analysis for scoring was the students' models and written responses to task prompts. The individual scores for students' pre- and postunit modeling tasks were imported into SAS for statistical analysis using double-factor repeated measures mixed model ANOVA. The dependent variable was the students' scores on the postunit modeling task and the independent variable was the scaffolded and unscaffolded modeling groups, controlling for student scores on preunit modeling task.

Table 2: Sequence scoring rubric

Level	Sequence representation
0	No sequence represented
1	Only describes 1 change (water in the sky \rightarrow falls to the ground)
2	Describes 2 or more changes but the sequence only goes in one direction
3	Describes 2 or more changes and the sequence goes in two directions

Five students from each classroom were also selected to participate in reflective grounded interviews about their pre- and postunit water cycle models and written responses ($n=60$). All student interviews were imported into ATLAS.ti and analyzed qualitatively using classical content analysis (Patton, 2001) for *a priori* code of sequence and mechanism. Qualitative analysis involved an iterative process of data coding, displaying and verification (Miles & Huberman, 1994) to identify themes within the interviews that provided insight into the students' articulation of mechanisms generated from their modeled water cycle sequences.

Empirical Results

In research question 1, we examined the impact of embedded curricular scaffolds on students' representation of water cycle sequences and found that embedded curricular visual and written prompts supported students' representation of the direction and quantity of process sequences underlying the water cycle in their postunit models, $F(1, 90) = 0.05$, $p = 0.01$, as shown in Figure 1.

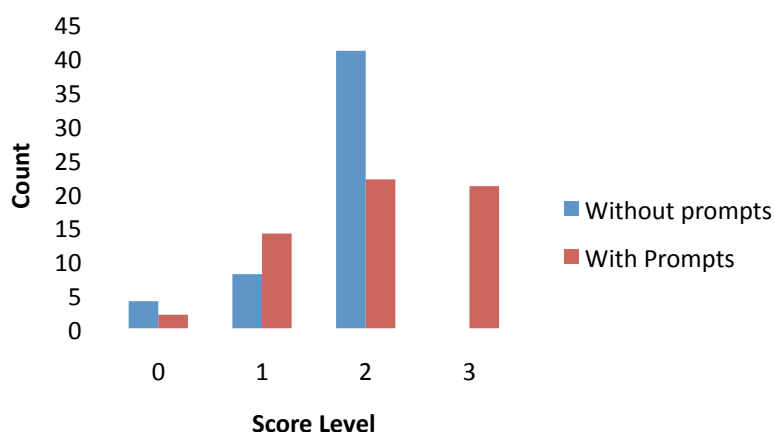


Figure 1. Process sequences in students' scaffolded ($n=60$) and unscaffolded ($n=55$) water cycle models

As shown in Figure 1, the most substantial difference between the two conditions was the prevalence of level '3' scores in the scaffolded group as compared to the unscaffolded group, in which there were no level '3' scores. More students from the scaffolded group were able to incorporate two or more process sequences into their models and illustrate multidirectionality of these process sequences in their explanations.

In research question 2, we qualitatively analyzed students' modeling tasks and interviews to explore differences in students' model-based explanations between the two conditions. We found two dominant themes. First, students in the scaffolded modeling group more frequently represented 'hidden' sequences of groundwater and water movement underground prior to returning to the sky (Figure 2). The students in the scaffolded group used both words and arrows to show sequences of rain moving through the underground layers. They frequently

articulated that “water get trapet [trapped]” (A.CM1) in the gravel layer and was unable to move any further underground. They drew and traced with their finger sequences of water moving vertically until it reached the gravel layer then moving horizontally into the lake (Figure 2). As Jackie stated, since “[rain] can’t go under the solid rock...it goes back into the lake” (P1:2713:2714). While the unscaffolded group also drew sequences of rain reaching the ground, their representations did not go further into the ground than just under the surface. They articulated that vertical water movement stops at ground surface because once rain reaches the ground, it does not have “anywhere else to go” (P1: 2904:2907) under the ground. We found that both groups included evaporation in their sequences (Figure 2); however, only the scaffolded group connected evaporation to their underground sequences creating a continuous loop and thereby representing the water cycle. The unscaffolded group most frequently only considered evaporation occurring from puddles and did not connect this sequence to their representations of water falling to the ground. While evaporation in both groups was drawn as arrows moving from bodies of water back to the sky, the scaffolded group most frequently drew arrows of water returning to clouds while the unscaffolded group drew arrows of water returning to the sun (Figure 2).

Second, even though the scaffolded group more frequently represented arrows connecting and indicating ‘hidden’ sequences underground, they were unable to articulate the mechanism - the *how* and *why* – for how water moves underground and returns to the sky. When we asked students in the scaffolded group why water flows vertically through the gravel layer, they responded with a description of what occurs rather than *how* or *why* it occurs. For example Nancy stated “...it [rain] comes from the sky...into the grass, and through the soil, sand and gravel, and...it makes the lake bigger...and the thing [evaporation] is taking the water from the lake” (P1:2836:2841). In this manner, Nancy has described her model to us, but not generated an explanation for the process sequences. However the unscaffolded group used their models to begin to attribute mechanisms for *how* and *why* water moves both in their representations of water falling to the ground and then water returning to the sky. They discussed that water reaches the ground because plant matter “acts like a magnet” for pulling water from the sky (N.CM2) and the “sun is so hot” it ‘takes’ (P1:2981) or ‘soaks’ (P1:423) water up from the ground. Overall, the unscaffolded group more frequently generated model-based explanations models for process sequences while the scaffolded group typically did not.



Figure 2. Students sequence in embedded scaffold (S3.AM2) and unscaffolded (T.NM2) postunit models

Discussion and Relevance to the Conference Theme

This work foregrounds elementary students’ learning within a discipline-specific (i.e. hydrologic cycle) epistemic practice (i.e. modeling) and therefore exemplifies the ICLS 2014 conference theme of ‘learning and becoming in practice’. Scientific modeling is a core scientific practice (NRC, 2012) that remains underemphasized in K-12 science, particularly in the elementary grades. Too often, when models are used in science classrooms, they are provided as static illustrations rather than to engage students in model development, use, and refinement as an active, sustained practice. As a result, little research exists to guide efforts to engage students in epistemically-rich, model-centric elementary science learning environments that foster and promote modeling as a *way of learning* (Abell & Roth, 1995; Windschitl et al., 2008). Here, we engaged students in co-development of modeling practices and content knowledge to support their formulation of model-based explanations (Forbes et al., 2014; Lehrer & Schauble, 2010; Manz, 2012; Schwarz et al., 2009).

The water cycle is a complex geosystem fundamental to understanding biotic and geospheric phenomena and, as such, is central in K-12 science standards (Gunckel et al., 2012; Henriques, 2002; NRC, 2012). However, students often struggle with the epistemic and cognitive demands placed upon them by the dynamic processes that comprise complex systems such as the hydrologic cycle where major components are largely ‘hidden’ from view. In order to support students in considering hidden process sequences, we provided visual and written prompts (McNeill & Krajcik, 2009) to explore the ways in which students engaged with these prompts to conceptualize water cycle sequences, and reason about groundwater. Study findings indicate that

visual and written prompts embedded in curricular tasks supported students to consider some important ‘hidden’ components of water in motion that the unscaffolded group typically did not include in their models. This may imply that embedded scaffolds provide students with additional representational space to add elements they would not otherwise consider (Sherin et al., 2004). However, results also indicate that the students who were prompted to model underground water processes seemed unable to articulate conceptual understanding about the phenomenon and did not generate explanations for sequences they represented within the visual prompts. We hypothesize that even though the invisible components were present, the phenomena may become meaningless to the learner if they cannot generate an explanation. In these instances, models became illustrations to describe rather than representations for use in sensemaking (Bechtel & Abrahamsen, 2005; Braaten & Windschitl, 2011). In contrast, the unscaffold group did not represent water underground, but did articulate model-based explanations for the sequences they included in their models. Even though the explanations from this group were based on naïve mechanisms of evaporation, their reasoning indicates that they were able to generate scientifically accepted explanations from their representations (Machamer et al., 2000). Taken together, these results suggest that engaging elementary students in generating model-based explanations about the water cycle may require multi-modal representations and varied scaffolds to support students in conceptualizing hidden components and how they may function within a system to afford mechanistic outcomes (Sherin et al., 2004). However, caution must be followed using embedded visual prompts within the modeling task as they may inhibit student opportunities for sensemaking (e.g., Abell & Roth, 1995).

References

- Abell, S. & Roth, M. (1995). Reflections on a fifth-grade life science lesson: Making sense of children’s understanding of scientific models. *International Journal of Science Education*, 17(1), 59-74.
- Bechtel, W. & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Science*, 36, 421-441.
- 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
- Forbes, C.T., Schwarz, C., & Zangori, L. (in press). Development of an empirically-based learning performances framework for 3rd-grade students’ model-based explanations about hydrologic cycling. In Proceedings of the Eleventh International Conference of the Learning Sciences – ICLS 2014. Boulder, CO: International Society of the Learning Sciences.
- Gunckel, K. L., Covitt, B. A., Salinas, I., & Anderson, C. W. (2012). A learning progression for water in socio-ecological systems. *Journal of Research in Science Teaching*, 49(7), 843-868. doi:10.1002/tea.21024
- Henriques, L. (2002). Children’s ideas about weather: A review of the literature. *School Science and Mathematics*, 102(5), 202-215. doi: 10.1111/j.1949-8594.2002.tb18143.x
- Lehrer, R. & Schauble, L. (2010). What kind of explanation is a model? In M. K Stein & L. Kucan (Eds.) *Instructional explanations in the disciplines*, 9-22, New York, NY: Springer.
- Machamer, P., Darden, L. & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(1), 1-25.
- Manz, E. (2012) Understanding the codevelopment of modeling practice and ecological knowledge. *Science Education*, 96(6), 1071-1105. doi: 10.1002/sce.21030
- McNeill, K. L. & Krajcik, J. (2009). Synergy between teacher practices and curricular scaffolds to support students in using domain specific and domain general knowledge in writing arguments to explain phenomena. *The Journal of the Learning Sciences*, 18(3), 416-460.
- Miles, M.B. & Huberman, A.M. (1994). *Qualitative data analysis*. Newbury Park, CA: Sage.
- NRC, (2012). *A framework for K-12 science education: Practices, cross-cutting concepts, and core ideas*. Washington, DC: National Academies Press
- Patton, M. Q. (2001). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
- Schwarz, C. V., Reiser, B. J., 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 learner. *Journal of Research in Science Teaching*, 46(6), 632-654.
- Sherin, B. Reiser, B. J., Edelson, D. (2004). Scaffolding analysis: Extending the scaffolding metaphor to learning artifacts. *The Journal of Learning Sciences*, 13(3), 387-421. doi: 10.1207/s15327809jls1303_5
- Windschitl, M., Thompson, J. & Braaten, M. (2008). How novice science teachers appropriate epistemic discourses around model-based inquiry for use in classrooms. *Cognition and Instruction*, 26(3), 310-378. doi: 10.1080/07370000802177193

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.

Exploring How Mobile Technology Provides Inquiry Supports for Middle School Students in Conducting Scientific Practices in a Ubiquitous Learning Context

Wan-Tzu Lo, Alex Kuhn and Chris Quintana, University of Michigan
Ibrahim Delen, Michigan State University

Steven McGee and Jennifer Duck, The Learning Partnership, Western Springs, IL

Email: {loclaire; kuhnalex; quintana}@umich.edu; delenibrahim@gmail.com, {mcgee; jenn}@lponline.net

Abstract: In a ubiquitous learning context, students can conduct scientific practices by using mobile and wireless technologies inside and/or outside of school. This particular learning context provides learning opportunities for students to collect and evaluate first-hand data anytime anywhere. However, learning challenges emerge in ubiquitous learning context when students are situated in both non-traditional classroom setting and information-rich out-of-class setting. To identify and address students' challenges in ubiquitous learning context, this study takes a longitudinal view to examine students' scientific practices throughout a school year. This paper explores how mobile and wireless technology can provide students with inquiry supports in conducting scientific practices, which include collecting qualitative multimedia data, evaluating and using first-hand collected qualitative data to construct scientific explanations in three different ubiquitous learning environments.

Introduction

Ubiquitous learning refers to a new learning paradigm in which students can learn anytime, anywhere, using tools such as portable and wireless communication technologies (Burbules, 2009; Yahya, Ahmad, Jalil, & Mara, 2010). Traditional technology-based learning environments allow students to use computers and the Internet in a fixed classroom setting. However, in a ubiquitous learning environment, students have more agency when their tools become mobile and personal. In other words, learning is no longer limited to a classroom setting; instead, students can move around between learning environments (Ogata & Yano, 2004) and conduct inquiry activities in different environments (Chen, Kao, Sheu, & Chiang, 2002, August; Chen, Kao, Yu, & Sheu, 2004, March; Hsi, 2003). Moreover, students are able to carry out investigations by collecting data inside or/and outside of school via mobile devices. They can further access and retrieve those data for later use through mobile technology and cloud services (Quintana, 2012). Hence, the distinction between formal and informal learning has become blurred (Burbules, 2009) as learning can take place seamlessly inside and outside of school. In addition, science educators have called for the need to expand students' learning opportunities in informal environments (e.g. everyday settings and family activities, designed settings, afterschool programs) (National Research Council, 2009), which are also characterized as learner-motivated, guided by learner interests, personal, and ongoing learning (Falk & Dierking, 2000). More importantly, informal science learning experiences are believed to lead to learners' further inquiry and enjoyment (National Research Council, 2009). To respond to the need for greater integration and connection of informal environments and learning experience, ubiquitous learning environments has become a unique context in terms of potential learning opportunities.

For science learning, a ubiquitous learning context allows students to utilize mobile devices and wireless networks to engage in scientific practices in a real world environment. A set of practices including "asking questions and defining problems," "developing and using models," "planning and carrying out investigations," "analyzing and interpreting data," "using mathematics and computational thinking," "constructing explanations and designing solutions," "engaging in argument from evidence," and "obtaining, evaluating, and communicating information" have been defined and emphasized in the Next Generation Science Standards (NGSS Lead States, 2013). These practices such as collecting data, analyzing data, and using data to construct explanations are commonly seen in a ubiquitous learning context. However, students may also face inquiry challenges when situated in this unique context. Essentially, inquiry comprises a series of hands-on and minds-on activities that students need to be engaged in. But, some studies have shown that students face challenges when conducting data collection, including mindless observation and careless recording (Driver, 1983; Eberbach & Crowley, 2009; Smith & Reiser, 2005). In addition, more challenges may emerge when students move from a familiar classroom environment to a novel and unfamiliar environment outside of school. This unfamiliar learning context such as a museum may cause cognitive load when students conduct scientific practices on their own (Falk & Dierking, 2000). Students may need different learning supports within these learning settings from a highly structured setting (e.g., museum) to a less structured one (e.g., a park). Furthermore, teachers' real-time and in-person feedback is limited especially when students are moving around outside of school. Second, students may face difficulties in choosing appropriate data to collect when they are situated in an information-rich learning environment (Vavoula, Sharples, Rudman, Meek, & Lonsdale, 2009).

Lastly, students may struggle to evaluate their first-hand collected data and to select appropriate data as scientific evidence from a data pool. As discussed, these challenges introduce complexity for ubiquitous learning in science contexts. Unfortunately, few studies have addressed the complexity of ubiquitous learning and discussed how educators can scaffold learners during inquiry processes in ubiquitous learning environments.

Since ubiquitous learning can extend students' science inquiry experiences, there is a need to understand how students perform in these environments and how we can support students to mindfully engage in scientific practices in a ubiquitous learning context, specifically collecting data, evaluating the qualitative data they collected, and using those data to construct well-supported scientific explanations. Hence, we designed this year-long study to examine how a mobile computing tool can provide inquiry supports in a ubiquitous learning context. The main research question drives this study: *"How do students use mobile-based inquiry tools to facilitate their scientific practices in a ubiquitous learning context?"* We examine students' performance when conducting scientific practices in different environments, and discuss if the technical support of mobile technology helps students to conduct scientific practices. This ongoing study has been enacted in a middle school in the Midwestern USA over the past year. Currently, we have completed data collection and we are now in the process of analyzing research data. Here, we will discuss the design of this longitudinal study, preliminary findings, and how this study may provide more information about how to support students in science-oriented ubiquitous learning contexts.

Study Overview

Participants included 35 eighth grade students and one science teacher recruited from a middle school. The teacher chosen for this study was comfortable using technologies and inquiry-based pedagogy in her class. In order to answer our research question, we situated students in a ubiquitous learning context where they used the Zydeco mobile-based system (Quintana, 2012) to conduct scientific practices. Each student was assigned an iPad loaded with the Zydeco program, and students could work in a group that consists of three persons. Over one school year, students conducted three different investigations in three environments: a classroom lab setting, a museum, and an outdoor river site. For the first investigation, students were asked to conduct a project related to plant growth. Students carried out experiments in the classroom lab and collected data during their class periods. The whole plant project lasted for two months. In the second investigation, students were introduced to a vehicle design project in which students needed to collect relevant data from the exhibits to figure out how to make a vehicle move from a one-day museum field trip. In the last investigation, students conducted an investigation related to the water quality of a local river, where they collected data during one-day outdoor field trip. When students returned from their fieldtrips, their collected data was accessed and used for later analyses. Both the museum and the river investigations lasted for one month. All the lesson plans were developed collaboratively with the teacher, and the other Zydeco project researchers to make sure the plans were aligned with the teacher's curricular plan, Grade Level Content Expectations, and the Next Generation Science Standards (NGSS Lead States, 2013.)

Zydeco, a combined mobile and web-based learning system (Quintana, 2012), has been developed to support scientific inquiry activities across different learning environments such as in a museum (Cahill et al., 2011; Cahill, Kuhn, Schmoll, Pompe, & Quintana, 2010) or in a science center (Lo et al., 2012). The Zydeco system consists of two parts: a tablet program for iPads and a website. Students can conduct a set of scientific practices via different phases on Zydeco: "Plan", "Collect", "Review", and "Explain" (Figures 1-4.) Figure 1 shows the "Plan" workspace where students can add sub-questions (or "helper questions") to refine the driving question, or hypotheses. Figure 2 shows the "Collect" workspace with a variety of data collection methods including capturing photos, text notes, audio and video. When collecting data, students can link specific helper questions or hypotheses to collected data, or use the "Other stuff" category for unsorted data. Additionally, Zydeco has a labeling system that allows students to attach a short textual label to a specific piece of data to easily annotate and search for data. With this annotation support, students are encouraged to reflect on their data collection and collect more meaningful data. Figure 3 shows the "Review" workspace where students can review and filter data collected by all users within the same investigation. The filter system allows students to sort data by the data collector, the data type, labels attached to the data, helper questions, and hypotheses associated with the data. Figure 4 shows the "Explain" workspace where students construct their scientific explanations based on the "Claim, Evidence, Reasoning" model to address the driving question (McNeill & Krajcik, 2011). According to this model, a "claim" can be defined as an answer to a question, "evidence" is the scientific data collected by students that supports their claims, and "reasoning" is the justification that explains the link between the evidence and the claim. With Zydeco, students have easy access to personal and peer data, which give students opportunities to think over the data more carefully when making an explanation.

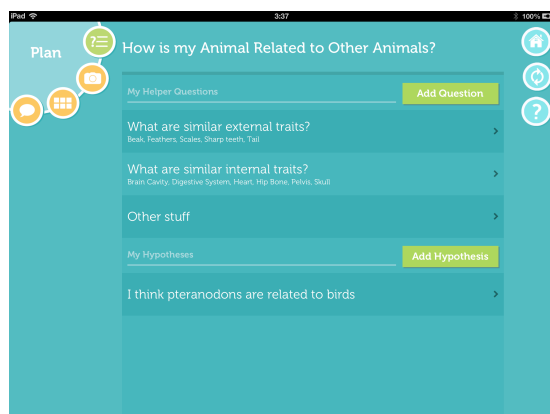


Figure 1. Zydeco plan workspace

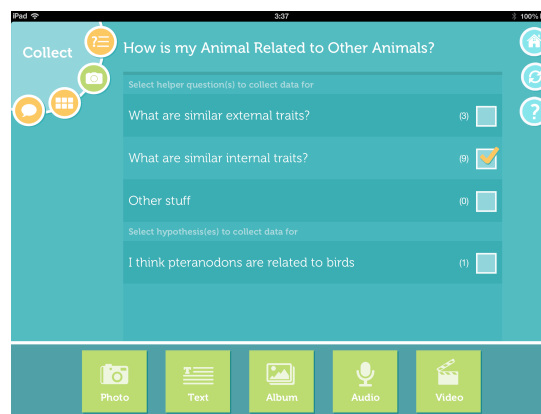


Figure 2. Zydeco collect workspace

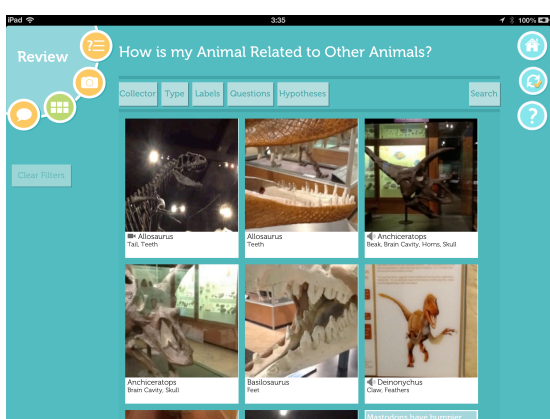


Figure 3. Zydeco review workspace

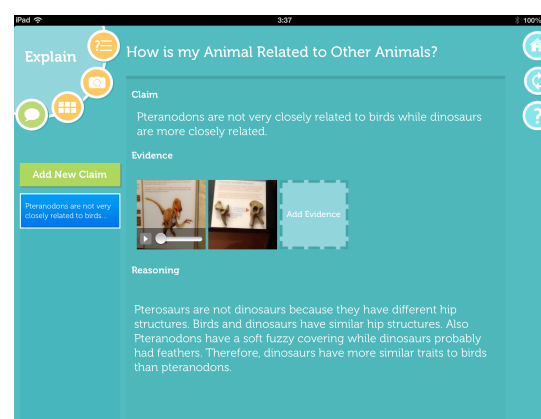


Figure 4. Zydeco explain workspace

Data Collection and Analysis

In order to answer the research question, students' iPad data were collected through the Zydeco web server. We examined whether students used the annotation system to help them collect more meaningful data on site. Based on the coding rubric that was tested and modified from previous Zydeco studies (Cahill et al., 2011; Lo et al., 2012; Lo et al., 2013), we determined whether the data were attached with labels, titles, and other additional information. We further examined the quality of that annotated information (whether labels or other information in relation to the object and the investigative question), and if those data and assigned to hypotheses or helper questions were related. This analysis is helpful to determine whether students make mindful observation and collect appropriate data. The coding rubric can be seen in Table 1.

Table 1: Rubric for analyzing students' on-site collected data

Coding	Data Type (Video, Audio, Photo, Text)	Data Quality			Data Relatedness	
		Object	Title Accuracy	Label Accuracy	Helper Question Relatedness	Hypothesis Relatedness
0	no data	the object is unclear to be defined (e.g., blurred photo)	no title	no label	no question attached	no hypothesis attached
1	with data	the object is mismatched with the investigative question	the title is mismatched with the object	the label is mismatched with the object	the object is mismatched with the question	the object is mismatched with the hypothesis
2	-	the object is matched with the investigative question	the title is matched with the object	the label is matched with the object	the object is matched with the question	the object is matched with the hypothesis

Another supportive feature of Zydeco is the embedded "Claim-Evidence-Reasoning" template that is designed to help students construct evidence-based explanations. To analyze students' explanations constructed on iPads, we adopted the rubric from McNeill et al. (2006) as shown in Table 2. Based on students'

explanations created on iPads, it is useful to determine whether the embedded “Claim, Evidence, Reasoning” template helps students construct better supportive explanations throughout a school year.

Table 2: Rubric for analyzing students’ explanations (modified from McNeill, 2006)

Scale	Claim A conclusion to a question.	Evidence Scientific data that supports the claim.	Reasoning A justification that explains the link between the claim and the evidence.
0	No claim	No evidence	No reasoning
1	Make inaccurate or incomplete claim	Provide appropriate but only one piece of evidence.	Link evidence to claim, without including scientific principles.
2	Make accurate and complete claim.	Provide appropriate and multiple pieces of evidence.	Link evidence to claim, including scientific principles.

Preliminary Findings

On average, every student collected 11 pieces of data from the lab investigation, 15 pieces of data from the museum investigation, and 6 pieces of data from the river investigation. In Figure 5, it indicates that in both lab and river investigations, the object and the title accuracy of the data is over ninety percent. However, in the museum investigation, students collected less accurate data and attached less appropriate title. In terms of data relatedness, Figure 6 shows that when students were situated in the museum investigation, students tended not to attach any helper question to their data. In comparison of three investigations, the river investigation shows the highest percentage of the data relatedness (50.68%) in terms of the accurate associated helper questions, and there was only 36% in the museum investigation.

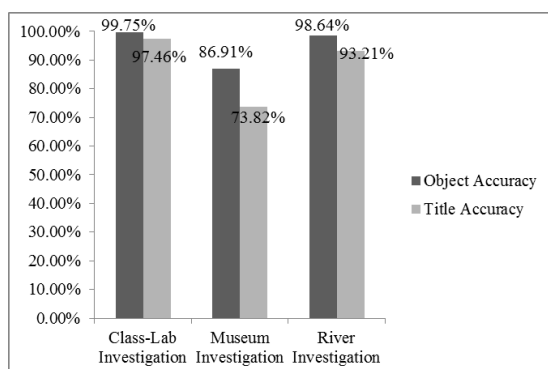


Figure 5. Accuracy of students’ collected data

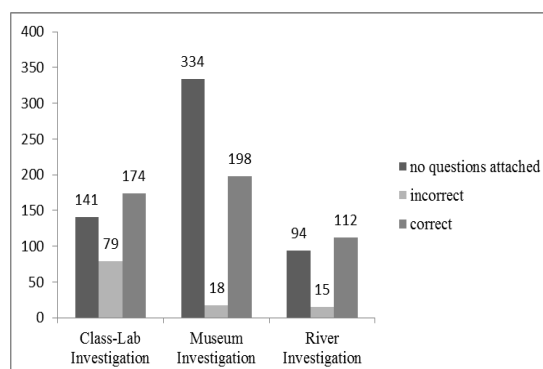


Figure 6. Helper question relatedness

Concluding Remarks

To expand students’ science learning opportunities in K-12, researchers and school teachers have tried to seek a way to integrate mobile technology into their curricula. By situating students within a ubiquitous learning context across environments, this study is helping us examine whether one single tool with embedded supportive features (i.e., the data annotation feature, accessible class data sets, the embedded “Claim, Evidence, Reasoning” template) can help students conduct scientific practices across environments. Our preliminary findings show that students may collect less accurate data and tended not to attach helper questions in the museum investigation. This indicates that when students are situated in a less-structured environment, students may collect less appropriate data. This suggests that educators may need to provide more inquiry supports before and during museum visits. Although in this report paper, we only provide some preliminary findings. However, we believe that this longitudinal study can help us understand whether students’ practices change overtime through the use of mobile tools in this ubiquitous learning context.

With the increasing use of mobile devices in educational fields, more educators are aware of the educational potential that mobile technology and cloud computing may bring into the school (Johnson et al., 2013). Therefore, before making an investment in mobile technology at school, district, or state level, researchers need to understand how students engage in complex practices like inquiry in a variety of environments, what types of challenges students may face, and how educators or program designers can provide support accordingly. Unlike previous studies that are constrained by a particular learning environment, this study will provide a more generalized and holistic view since students’ performances were evaluated based on three science projects in multiple settings. Furthermore, the findings from this study will be useful for educators from informal learning institutions (e.g., museums, zoos), and possibly formal classroom educators, to design more ubiquitous learning activities. Additionally, the learning scenarios used in this study can be adopted and further modified to broaden and connect visitors’ learning experiences.

References

- Burbules, N. C. (2009). Meanings of ubiquitous learning. In B. Cope & M. Kalantzis (Eds.), *Ubiquitous Learning* (pp. 15-20). Urbana, IL: University of Illinois Press.
- Cahill, C., Kuhn, A., Schmoll, S., Lo, W.-T., McNally, B., & Quintana, C. (2011). Mobile learning in museums: how mobile supports for learning influence student behavior. In *Proceedings of the 10th International Conference on Interaction Design and Children* (pp. 21-28). doi:10.1145/1999030.1999033
- Cahill, C., Kuhn, A., Schmoll, S., Pompe, A., & Quintana, C. (2010). Zydeco: using mobile and web technologies to support seamless inquiry between museum and school contexts. In *Proceedings of the International Conference on Interaction Design and Children* (pp. 174-177). doi:10.1145/1810543.1810564
- Chen, Y.-S., Kao, T.-C., Sheu, J.-P., & Chiang, C.-Y. (2002, August). A mobile scaffolding-aid-based bird-watching learning system. In *Proceedings of the International workshop on wireless and mobile technologies in education* (pp. 15-22).
- Chen, Y.-S., Kao, T.-C., Yu, G.-J., & Sheu, J.-P. (2004, March). A mobile butterfly-watching learning system for supporting independent learning. In *Proceedings of the International Workshop on Wireless and Mobile Technologies in Education* (pp. 11-18).
- Driver, R. (1983). *The pupil as scientist?* Milton Keynes, England: Open University 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.
- Falk, J., & Dierking, L. (2000). *Learning from museums: Visitor experiences and the making of meaning*. Walnut Creek, CA: AltaMira Press.
- 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.
- Johnson, L., Adams, S., Cummins, M., Estrada, V., Freeman, A., & Ludgate, H. (2013). *The NMC Horizon Report: 2013 K-12 Edition*. Austin, Texas: The New Media Consortium.
- Lo, W.-T., Delen, I., Cahill, C., Kuhn, A., Schmoll, S., & Quintana, C. (2012). A New Type of Learning Experience in Nomadic Inquiry: Use of Zydeco in the Science Center. In *Proceedings of the 7th IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education* (pp. 57-61). doi:10.1109/WMUTE.2012.16
- Lo, W.-T., Delen, I., Kuhn, A., McGee, S., Witters, J. L., & Quintana, C. (2013, April). Zydeco: A mobile-based inquiry learning system to support project-based learning. Paper presented at the American Educational Research Association, San Francisco, CA.
- McNeill, K., & Krajcik, J. (2011). Supporting grade 5-8 students in constructing explanations in science: The claim, evidence, and reasoning framework for talk and writing. New York: Allyn and Bacon.
- 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.
- National Research Council. (2009). *Learning science in informal environments: People, places, and pursuits* (P. Bell, B. Lewenstein, A. W. Shouse & M. A. Feder Eds.). Washington, DC: The National Academies Press.
- NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.
- Ogata, H., & Yano, Y. (2004). Context-aware support for computer-supported ubiquitous learning. In *Proceedings of the Wireless and Mobile Technologies in Education, 2004*. (pp. 27-34).
- Quintana, C. (2012). Pervasive science: using mobile devices and the cloud to support science education. *interactions*, 19(4), 76-80.
- Smith, B. K., & Reiser, B. J. (2005). Explaining behavior through observational investigation and theory articulation. *Journal of the Learning Sciences*, 14(3), 315-360.
- Vavoula, G., Sharples, M., Rudman, P., Meek, J., & Lonsdale, P. (2009). Myartspace: Design and evaluation of support for learning with multimedia phones between classrooms and museums. *Computers & education*, 53(2), 286-299.
- Yahya, S., Ahmad, E. A., Jalil, K. A., & Mara, U. T. (2010). The definition and characteristics of ubiquitous learning: A discussion. *International Journal of Education and Development using Information and Communication Technology* 6(1), 117-127

Acknowledgments

We would like to thank Joy Reynolds for her support and guidance. This study is supported in part by the National Science Foundation under grant number DRL 1020027. Any opinions and findings expressed in this study are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Empowering Under-Represented Middle School Youth in Engineering Knowledge and Productive Identity Work

Angela Calabrese Barton, Michigan State University, acb@msu.edu
Daniel Birmingham, Loyola University Chicago, dannybirmingham@gmail.com
Edna Tan, University of North Carolina at Greensboro, tane@msu.edu
Takumi Sato, Virginia Tech, takumisato74@gmail.com

Abstract: Drawing from critical sociocultural perspectives, we report on how urban middle school youth from non-dominant backgrounds engage in productive identity work in engineering in an after-school program, as well as the relationship between identity work and their participation and learning in engineering design. Longitudinal, ethnographic data of 14 youth were analyzed with our conceptual framework using constant comparative analysis. Findings indicate different pathways taken by youth. Data also revealed the iterative and on-going movement between a) investigating and b) communicating with and taking action in the community supports powerful learning in conjunction with productive identity work in engineering. The process of figuring out what and how to communicate to others about one's investigation in scientifically rigorous and culturally resonant ways, and in accounting for multiple outside perspectives, is pivotal to both *learning* and *identity development*.

Objectives

Studies reveal that student success in school science is not well correlated with the pursuit of engineering trajectories. Even when students are successful in learning science, many still see the subject as uninteresting and disconnected from their lives and pursuits. Identity gaps continue for students from non-dominant backgrounds. Long-term studies in engineering indicate that identity development is critical to how and why individuals pursue engineering trajectories (Eccles 2007). If one does not see oneself as scientific, or as a member of the engineering communities, this can negatively impact opportunities to learn (NRC, 2012).

Our research has shown that interest in science and engineering through the middle grades is sustained by opportunities to *be an expert in practice in community* and to be recognized by others as having relevant expertise (Tan, Calabrese Barton, Kang, & O'Neill, 2013). These opportunities to be such an *expert* supports the authoring of productive identities. We have learned that when youth positioned themselves as community science experts – or someone with knowledge of science and community needs and how to leverage that knowledge towards taking action – their learning gains were stronger and they more positively identified with science (Calabrese Barton & Tan, 2010). However, we do not have evidence for how to scaffold this kind of identity work as a part of the process of conducting community-based investigations.

Research Questions

Our research focuses on how and why youth construct identities in engineering in ways that bridge community expertise with engineering expertise: 1) In what ways do urban middle school youth from non-dominant backgrounds engage in identity work in engineering through participation in an after-school engineering program? 2) What ideas, resources, and identities youth move across spaces (investigation, action in the community), and how this movement matters in identity work and engagement in engineering? And 3) what is the relationship between identity work and their participation and learning in engineering design?

Theoretical Framework

Our study is grounded in social practice theories, which suggest that identities reflect one's on-going social existence in the world. As individuals move through the world, they are exposed to, positioned by, and react to a range of people as well as institutional and cultural structures and forces. As individuals join new communities of practice, they call upon salient practices and ways of being that are learned from other places, creating new hybrid practices that can position one as either central or marginal to their new community. Such acts of identity work are complex, for how one is recognized within new communities is an artifact of the power dynamics that operate there (Nasir, 2011) and which reflect the cultural norms of "local practice" and "historically institutionalized struggles" (Holland & Lave, 2009).

We use the phrase *identity work* to capture the idea that authoring oneself in engineering or in any domain involves an on-going effort, and positions the author with agency. It is through the actions people take and the relationships they form that they position themselves as particular kinds of people over time and space. Because identity work happens within and against local norms and expectations and as a part of longer standing sociocultural and historical narratives, its outcomes are always uncertain and gain new meaning as they get traced in time. The practices of the engineering learning community or the peer culture as informed by dominant

norms and routines position youth in particular ways and they react to those positionings. This identity work brings together “two forms of history” – the personal and the institutional – and as Holland and Lave argue, what emerges is a sort of “local contentious practice” grounded in “cultural-historical conjuncture” (p 13).

In our paper we discuss “productive identity work”, which references movement towards seeing oneself as an important producer and critic of engineering. We argue that productive identity work in engineering reflects 1) one’s developing *knowledge and practice* within a community of practice, and an ability to gain membership through knowledge and practice; 2) an ability to navigate the dialectic between *structure and agency* (Do students see themselves as someone who is capable, and can leverage an array of resources to gain understanding and take action?) and 3) *recognition/positioning* by others (Is one accepted by others as the person they desire to be and with the salient expertise?). This last point is salient for youth who experience a disjuncture between their home worlds and the world of engineering.

Research Methods

To build generalized claims about the youths’ identity work we employed longitudinal multi-sited ethnography case study, paying particular attention to the power dynamics involved. Multi-sited ethnography involves the study of learners across contexts, and inquires into the ways that “people, ideas, tools, artifacts and questions, move and become reconstituted across the boundaries of school, home, and community spaces *and* across multiple contexts and environments” (Vossoughi & Gutierrez, in press, p. 5). This approach requires the collection of multiple forms of data using different strategies such that the corpus of data results in complementary strengths and non-overlapping weaknesses, which is useful in expanding understanding and informing theory and practice. The longitudinal component maximizes opportunities to refine theory related to how decisions developed overtime by assuring a steady stream of data at different time points in the process.

Context

The study takes place primarily at the Boys and Girls Club of Great Lakes City, MI, a mid-sized midwestern city. The Club has served the community for 50 years, and welcomes over 2,400 youth annually between the ages of 7-17 from low income backgrounds. The club provides a safe place for youth to engage in many activities allowing opportunities to play, learn and have fun. One of the programs offered to members of the club is an informal science learning program, “Green Club.”

Green Club meets for 3 hours each week during the school year and 2 full weeks during the summer. The program has been designed on the premise that meaningful learning happens when youth engage in authentic investigations with scaffolded opportunities to communicate the findings of those investigations to others. It does so by providing a year-round after-school program that emphasizes youth development into science and engineering *experts* and *citizens* by using technology to take on relevant green energy issues and communicate findings to their community. An explicit goal of Green Club is to support youth in becoming community science and engineering experts – or youth who are uniquely positioned to draw upon their expansive knowledge of community and science/engineering to engage in meaningful, local practice.

For example, during the 2011-2012 school years, one of the two main years in which this study took place, Green Club youth investigated the design question of whether the Green Club could “get off the power grid.” Students engaged in activities investigating the electrical production system and its alternatives, with a particular focus on energy transformations in these various pathways and the feasibility of implementing one or more of these approaches into the current Club. Youth also studied the impact of these different transformation pathways on the local environment and economy. The youth designed a system where they generated the electricity for the lighting and mobile computer lab by riding bicycles attached to a generator and battery, and conservation techniques to accommodate the difficulty of producing all of their own electricity. They developed a plan to involve as many youth from outside Green Club in riding the bicycle during the club’s open hours, although instituting this broad participation proved a barrier. Youth recommendations for modifying their system included adding parallel bicycles to allow greater production during shorter time windows, and designing a solar system or wind system.

Participants

The target audience for Green Club are Great Lakes City area youth from underrepresented backgrounds. Child poverty in Great Lakes City has increased over 40% since 2000. 27% of Great Lakes City children live below the poverty line, with the rate jumping to over 40% for youth from African American backgrounds. Green Club strives to recruit a variety of youth with a range of skills and interests. Youth who have not performed well in school or in school science, or who are not interested in science or engineering, are encouraged to participate in Green Club. Within the program, youth are encouraged to leverage the various forms of expertise and interests they bring to studying energy and their environment by incorporating and valuing art, technology, and community concerns. The data presented in this paper is focused on the 14 youth participants from the 2011-2012 school year. However, most youth participate in Green Club for 2 to 3 years,

allowing us to draw upon multiple years of data. For 8 of the 14 youth, we draw upon 2 or more years of data. All of the youth were between ages ten and 13. Seven of the eight case study youth are African American and seven of the eight are also female (one African American boy, and one white female). The youth were in 5th, 6th or 7th grade dispersed in several local schools during the school year.

Data Generation

The findings of this paper are based on data collected from multiple sources during the 2010-2011 and 2011-2012 school years and included: 1) *Collaborative Conversations*. A researcher met with 6-8 youth weekly to debrief and develop a stronger sense of what they cared about with respect to green energy in their community. These conversations were held for fifteen weeks each year lasting between 60-90 minutes. 2) *Interviews*. Interviews were conducted with all participating youth. Interviews focused on the artifacts generated during the Green Club units and youth and researcher identified key moments, their role in creating these artifacts/moments, and the role these artifacts played in making change in their community. 3) *Transcripts of Green Club sessions*. Each week, Green Club was video and audio recorded and sessions were transcribed. Green Club held 20 regular sessions each year between October and April that lasted approximately 90 minutes each. We turned to these videos to situate and to make sense of students developing understandings and decisions. 4) *Student work*. We collected the artifacts youth generated as they investigated various green energy issues primarily from the “Getting off the Grid” unit (i.e., movies and raps about green energy and climate change, posters on alternative energy, PowerPoints and web pages, blog posts, and key data representations used in support of these, such as tables, graphs, and maps).

Data Analysis

We developed portraits of each case study youth that included background information, stories about current and future self(s), family, and science, participation over time in Green Club, and participation in Green Club work outside of Green Club. We focused on several focal events where the youth actively appropriated resources and positioned themselves in ways that supported engagement in engineering. The identified focal events were the ones that stood out retrospectively because of how they were referenced in future activities over time or how they appeared to reflect a shift in a youth’s perception of self and/or how others perceived them. We wrote descriptions of focal events, and analyzed the events using our figured worlds framework, including noting the (a) rules and norms, (b) tools and resources, (c) practices, (d) division of labor, and (e) object of work. Descriptions were shared at weekly meetings and debated until consensus was reached on the interpretation of descriptions and the emergent claims. We then analyzed the roles the youth played in each of the events, the ways in which the youth drew upon resources, and the produced identity artifacts. We paid attention to both resources and identity artifacts mediated the youth’s engagement in the activities, and the meanings produced about youth’s self(s) in engineering. These “role, resources, and identity artifact maps” were then shared at group meetings over the course of three months for further group analysis, discussion and debate.

Findings

We present two main findings. First, the youth in our study all engaged in forms of practice that positioned them as community engineering experts – as youth who are expert in engineering design and in the needs of their community, and uniquely positioned to leverage both towards taking action. However, how and why youth did so looked different across the youth, and had implications for how they positioned themselves as experts. Below we describe three pathways that frame their identity work in engineering along these lines:

Horizontal Movement as Critical Expertise

The six youth, whose identity work we would categorize in this way, privilege their horizontal expertise in engaging in engineering design (e.g., expertise gained in home and community, such as their funds of knowledge), and they leverage it towards advancing their opportunities to gain access and status within the engineering design process. Consider Quentin who writes a letter to his teacher explaining that he is not the C or D science student that he is viewed of in science class. He is really an A or B student who does “things out of school and out of Get City that involve science. I went to door to door and ask adults if they use CFL lights. The majority of the adults did NOT use CFL lights, I will try to decrease the amount of people who use incandescent lights. I did it on Wain Wright Ave and I did it because people’s bills are up because they use just Incandescent lights.” From this perspective, the work youth do to become expert in engineering, involves re-inventing that practice in ways that frame their cultural knowledge and expertise as critical to doing engineering.

Humanizing Engineering

For the five youth whose identity work falls into this category, we mean more than emphasizing the human elements of the practice. Rather we draw from critically oriented sociocultural studies which frame humanizing as relationships of dignity and care for both researchers and participants.” (Paris, 2011, p. 137).

While not applied to engineering practices or identity work (such ideals have been applied primarily to research and how participants are constructed & positioned through the process) we believe that this captures how youth frame their engagement with engineering – what they know and care about, what they do, how they view themselves, and want to be recognized by others. As Vossoughi & Gutierrez write, (2014), “actively cultivating new forms of perception can open up new ways of imagining and organizing environments for transformative and consequential forms of learning—a fundamental premise of our work on social design experiments” (p. 13). We present a discussion of the case of Hannah later to portray this pathway.

Distributed and Snowballing Expertise

This pathway, which helps to explain the identity work of three of the youth in our study, involve youth who iteratively play with ideas and designs as an important component of their engineering knowledge. This playfulness of ideas is accompanied by the on-going need to involve others and their relevant expertise. A critical piece of this is drawing upon their influence to recruit new insights, ideas, and participants from across communities to make their design work successful and accepted across their many communities of participation.

Our second finding focuses on the relationship between productive identity work and the design of the learning environment. Our data revealed the iterative and on-going movement between a) investigating and b) communicating with and taking action in the community supports learning in conjunction with productive identity work in engineering. The process of figuring out what and how to communicate to others about one’s investigation in scientifically rigorous and culturally resonant ways, and in accounting for multiple outside perspectives, is pivotal to both *learning* and *identity development*. The process involves on-going engagement with the design process, the embedded engineering content knowledge, and the needs/concerns of the community or end user. This iterative process approximates the work of engineers as they make on-going design decisions in light of ever-refining understandings of the problem space and design constraints.

We now use the case of Hannah below to illustrate the humanizing pathways and its implications for her identity work. Hannah has been an active participant of Green Club for the past 4 years. She first joined Green club because her friends were participating. Hannah is not deemed a strong student in school, and is on an individualized education program (IEP). Her grades are poor, in the range of Cs and Ds. Hannah articulates a future career as a hair dresser. When she first joined Green club, she did not come with an interest in science or engineering. Her motivation for being part of Green club was to socialize and spend time with friends. Through her Green Club participation, Hannah started to see herself as a “make a difference expert”. Her desires to make a difference in her community caused her to find ways to tackle the hard job of learning science and math.

When asked about her high point or a really significant moment with science and engineering, Hannah chose to talk about bringing what she had learned about energy efficiency and the design of different types of light bulbs to younger students at her school. With a small group of other middle grades girls, Hannah took what they had learned about energy efficient technologies and behaviors from Green Club and adapted the content to teach some younger students at her school. They prepared a Powerpoint and then led the class through an experiment looking at the difference between CFL and incandescent light bulbs. Hannah was to co-teach this class with her friend, a girl who is also a Green club member and who is an academically much stronger student. However, when her co-presenter was suddenly beset with acute stage fright, Hannah stepped up admirably. A typically dead silent girl in science class, Hannah presented the mini-lesson and hands on experiment with aplomb. Hannah points to this episode as a significant moment for her because she showed everyone she can do this kind of work, and she was helping her peers make a difference in their community.

During one Green Club session, Hannah and her friends, in their effort to explain the difference between Green Club science and school science, came up with what they termed “Fscience”: “Fscience \ˈfi-ən(t)s\ - “Fscience is what describes Green Club. It’s science that’s fun. Green Club knows how to make science fun without getting bored. Instead of a bored face, you will have a happy face” (Summer & Hannah). The girls also stated at Green Club, they did “science that matters.” When we asked Hannah to tell some of the other youth at one of the first Green Club sessions of the year, she stated that science that matters is: “Doing things that are good for the community because of what we know. We know a lot of science and we also know a lot about our community. Who else can put these ideas together?”

In a following conversation group where the similarities and differences of Green Club science and school science were again debated, the researcher asked the group (which included Hannah), to elaborate on “Fscience” and how “the science in Green club matters.”

- Jayah: “I think it matters because science in school we just sit there and read a book and that is not doing anything. All we do is sit there and read a book about doing something. And when we do something, it is like an experiment maybe that doesn’t really matter.
- Quentin: “And the teacher does it.”
- Jayah: “Yeah, and the teacher just shows you. Then he said, well you guys are going to get to do this, no – never mind, I can just show you. So all we did was make water drip into a

- bucket through a straw. When we are in Green Club we actually do something. We don't just sit there and read a textbook and watch our teacher drip water through a straw.
- Hannah: Is watching a straw and water drip through it even science? That is not even science.
- Researcher: Anyone else want to disagree or agree with the statement that the science in Green Club matters? So does it matter?
- Kat: It matters to the future.
- Quentin: And if you want to do engineering like Brittany.
- Researcher: Ok, so it is important to the future of the Earth and your own future. Usually when we say something matters we say it matters to someone. So who does this science matter to?
- Hannah: Me.
- Jayah: Our community. They don't know it yet, but our community. Watch, when we save the Earth from all of the disastrous stuff that is going to happen to us, they are going to be like, oh I should have, yeah.

Hannah had consistent opportunities to explore and engage in engineering practices, while also engaging in educative actions, directly applying her new knowledge. That Hannah still wants to be a hair dresser when she grows up is important. We do not see her future career choice as problematic or incompatible with her growing interest and facility with science and engineering knowledge and practices. Hannah views her engineering self as a part of being a good citizen. It is important to her that as a good citizen she has the expertise in some areas of science and engineering that is relevant to everyday life, and that she can use her expertise to make a difference now. In the future, she also wants to use green energy techniques at her hair studio that she will one day own. Hannah thinks that making a difference happens in the little things, but that you have to work hard to figure it out. Even though Hannah does not profess to “love engineering” nor aspire to a STEM career, she has engaged in productive identity work in engineering through her performances as a “make a difference expert” identity. And, this identity has supported her in tackling difficult learning moments.

Significance

In the US, youth from low-income and non-dominant backgrounds express interest in or opt to move into engineering at extremely low rates. For example, African Americans make up only 5% of the engineering workforce, with a majority holding technician rather than managerial or leadership positions. This statistic has changed little despite reform efforts. Our research sheds light on mechanisms that support productive identity work in engineering, how that might be scaffolded, and how such scaffolds can transport across context. We also offer insight into approaches for breaking the cultural barrier to engineering for underrepresented youth.

References

- Calabrese Barton, A. & Tan, E. (2010). We be burnin: Agency, Identity and Learning in a Green Energy Program. *Journal of the Learning Sciences*. 19(2), 187-229.
- Eccles, J. (2007). Where Are All the Women? Gender Differences in Participation in Physical Science and Engineering. In Ceci, S. & Williams, W. (Eds). *Why aren't more women in science?: Top researchers debate the evidence.*, (pp. 199-210). Washington, DC, US: American Psychological Association, 254.
- Holland, D. and J. Lave, eds. (2009). *History in person: Enduring struggles, contentious practice, intimate identities*. Albuquerque: School of American Research Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Nasir, N. i. S. (2011). *Racialized identities*. Palo Alto: Stanford University Press.
- National Research Council [NRC]. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press.
- Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational Researcher*, 41(3), 93-97.
- Tan, E., Kang, H. O'Neill, T. & Calabrese Barton, A. (2013). Desiring a career in STEM-related fields. *Journal of Research in Science Teaching* 50(10), 1143-1179.
- Vossoughi, S., & Gutiérrez, K. (in press). Toward a multi-sited ethnographic sensibility. In J. Vadeboncoeur (Ed.), *NSEE yearbook*. New York: Teachers College Press.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants No. DRL-0737642 and HRD- 0936692. Any opinions, findings, and conclusions, or recommendations expressed in this material are our own and do not necessarily reflect the views of the National Science Foundation.

Becoming a Youth Worker in a Classroom Community of Practice

Laurie Ross, Clark University, 950 Main Street, Worcester, MA, lross@clarku.edu

Abstract: Traditional university classrooms are more conducive to learning about youth work than they are learning how to become a youth worker. In this paper, I explore how a university classroom can function as a community of practice (CoP) in which actionable youth worker expertise is transmitted. Through narrative analysis of two youth worker dilemma stories, I show how a classroom-based CoP facilitates the development of three youth work ‘abilities.’ These abilities include: how to frame complex and ambiguous youth work problems; how to bring personal knowledge into practice; and how to reflect-on and -in practice.

The fundamental aim of youth work is to build trusting and mutually respectful relationships with young people. Youth workers create safe environments for young people to connect with supportive adults and peers and guide those harmed by oppressive community conditions through a process of healing. Youth workers help young people to develop knowledge and skills in a variety of areas including: academic, athletic, leadership/civic, the arts, health and wellbeing, and career exploration. In short, youth workers create transformative experiences for young people in formal and informal spaces outside of homes and schools.

Efforts to professionalize youth work in the United States largely have focused on establishing youth worker competencies (Akiva, 2005; Vance, 2010). Starr, Yohalem and Gannett (2009) define competencies as basic knowledge, skills, or attitudes in a specific domain. Advocates for competencies argue that they allow for clear standards of practice, consistent job requirements, reliable evaluation procedures, and identifiable career pathways (Vance, 2010). Yet, others question whether the acquisition of competencies ensures expertise. Walker and Gran (2010) argue that competency models set benchmarks, but do not guide novices to advanced status. More problematic is that a competency focus tends “to reduce practice to the most measurable, reducing youth work to a technical skill,” (Ibid, p. 2). An unintended implication is that youth workers could demonstrate competence in child and youth development on a test, but be unable to handle a child-related challenge at work.

An *expertise* frame, on the other hand, focuses on the successful application of knowledge and experience in context. In the face of dilemmas, expert youth workers “orchestrate multiple competencies into a full range of behaviors necessary for effective practice.” (Walker & Gran, 2010, p. 3). The reference to ‘orchestrate’ implies that expertise involves a complex process of combining and blending different types of knowledge and skills in context specific ways. This dynamic process allows youth workers to read and understand people and situations in order to resolve the everyday and extraordinary dilemmas of practice.

Three ‘abilities’ that facilitate the development of youth work expertise are (1) how to frame complex and ambiguous youth work problems (Larson & Walker, 2010); (2) how to bring personal knowledge into practice or ‘self-in-action’ (Krueger, 1997); and (3) how to reflect-on and -in practice (Emslie, 2009; Schön, 1987). A competency approach to youth worker professional education is not conducive to learning these abilities. I argue that novice and experienced youth workers can develop and/or deepen these abilities through classroom-based communities of practice (CoP) if the CoP is informed by specific design principles.

Cultivating Youth Work Expertise in a Classroom Community of Practice

Lave and Wenger (1991) argue that learning is not the acquisition of knowledge by individuals so much as a process of social participation and a movement from periphery to center in a community of practice. This process involves identity development in which one learns the language, actions, and practices that constitute the community. While it may seem contradictory for a classroom to be a community of practice, I have been working with three design principles that allow novices to learn from the situated knowledge of experienced youth workers and for experienced youth workers’ learning to deepen through reflection (Ross, 2012).

First, having experienced youth workers from the local community learning alongside traditional college students is a core feature of this classroom-based youth work community of practice. The youth workers tend to be more diverse racially, span a much wider age range, and are more balanced in terms of gender than the college students. The youth workers have many years of work experience, but tend not to have formal training or degrees in youth work; in fact many do not have college education.

Second, holding the course in ‘practice-space’ (i.e. at a community-based youth development organization) creates a teaching and learning environment that is more comfortable for the youth workers than the college students. The students have to leave their campus bubble to get to class. While unfamiliar to most of the students, the youth workers know the space, having worked, run programs, and/or attended events there. The class location privileges youth workers’ knowledge and de-centers expertise away from me as the professor.

Third, the class focuses on the practice of youth work. I focus on practice by organizing the course around case studies of youth work dilemmas. The dilemmas come from two sources: 1) interviews I conduct

with experienced youth workers and 2) from the students' required dilemma journals. In addition to describing the dilemma the youth worker faces, the case studies provide some biographical information. Each case study is accompanied by one or two scholarly readings and a set of questions to guide class discussion and written responses. In the dilemma journals, students are asked to provide a deep description of a challenge they are facing, to include the outcomes of their actions, and to reflect on how they handled the dilemmas. Students share journal entries in class to hear how others would handle the problem as well as to provide an opportunity to analyze the causes and structure of the problems. Students who are not youth workers are required to 'apprentice' with experienced youth workers so that they will have a source of dilemmas to draw on. I call them apprenticeships so that students realize they are meant to learn from the youth workers and not 'help' the youth.

To facilitate students' ability to contextualize and analyze the dilemmas, I introduce several conceptual frameworks (e.g. ecology of human development; Positive Youth Development; and Social Justice Youth Development). The class-based dialogues and written reactions to dilemmas cultivate youth workers' ability to reflect-on-action and ultimately encourage them to reflect-in-action. I obtained IRB approval and student consent to audio-record classroom conversations and use the dilemma journals for research purposes.

These three design principles create an environment in which novice and experienced youth workers learn what it means to 'be' a youth worker and not just learn about youth work. I demonstrate the teaching and learning that occurs in this classroom-based community of practice by presenting excerpts from two youth worker dilemma narratives; both youth workers were also students in the class. Narrative inquiry is a form of qualitative research in which people's stories are the unit of analysis (Bleakley, 2005; Bruner, 1991). Names of people and gangs have been changed to maintain confidentiality.

The first is Ricardo's story about an incident involving a young person who brought a gun into his youth program (Ross, 2013). Student reactions to his story demonstrate how a dilemma focus allows novice youth workers to learn how to frame ambiguous and complex dilemmas. The second is Jessica's story about managing the aftermath of a gang-related shooting that involved members of her organization. In this story I show how guided questions allow an experienced youth worker to develop a more reflective stance toward youth work. Ricardo's and Jessica's dilemmas are written from their point of view, mostly in their own words.

Case One: Framing Complex and Ambiguous Youth Work Dilemmas

Ricardo is thirty-one and has been in the youth work field since he was 16. He works at a drop-in youth center. He is a physically large man but soft spoken and approachable. When he was fourteen, Ricardo started going to the YMCA to play basketball. His neighborhood was tough and his mom had to work a lot of hours to keep the family going. He said going to the Y, "Gave me a mental break from having to watch out for where we were walking, and what kids we would run into." Even though he had the Y, he admitted he still did some, "stupid stuff" like stealing cars and going joy-riding. Many of his close friends carried guns, "sometimes it was like the Fourth of July, kids would just let off shots." He saw his friends graduating from selling weed to cocaine. They would invite him along asking him if he wanted to make some money. But by that time he had already started working with youth. He wondered, "Why would I do something that isn't going to show me love back? Especially when these little kids are looking up to me?" A lot of his friends thought he was a "cornball" for choosing work and school over them. Ricardo, however, completed college with a degree in human services.

Ricardo's Dilemma

Not long ago I dealt with something that I have never had to deal with. The day started normally, but soon Nicki, the GED teacher approached me, looking worried. She told me that Lisa, one of her students, came into her office during a break, wanting to go home. When asked why, Lisa broke down and told her that while she (Nicki) had stepped out to answer a phone call, two students started talking about a recent fight. Anthony, one of the students, said that he didn't feel safe in the streets and had to do something about it. Lisa pretended to not look, but out of the corner of her eye she saw that Anthony had a gun tucked under his shirt.

I quickly realized we didn't have a protocol for this. I informed the Executive Director and we talked about calling the police. But then we thought about Anthony. I knew he wouldn't be a kid that would be pulling out a gun to shoot me. He's tough; he's gone through stuff. But I've been around some pretty crazy people, and he didn't come off that way. He had been coming to GED, he had a kid on the way. He was trying to do the right thing. I thought he is probably scared. I thought it was probably not even a real gun and if it was real, I thought it might not be loaded. I thought that he's probably trying to prevent people from beating him up-versus the 'I want to go hurt somebody'. And I thought about him bringing it here was more about him getting here safely, versus him thinking 'I'm looking for somebody to shoot'. I didn't think that Anthony would hurt somebody here. He loves this place. If it was a kid I didn't know, then I might have thought differently.

We decided to pull Anthony into another room to talk rather than call the police. I was quiet at first, but soon I found myself telling Anthony, "some students have left, we think maybe because they saw a gun on you. Is this true?" Anthony admitted it. He said it was a BB gun. I said, "I'll believe it when I see it." Anthony handed it to me. I have to say, it did look like a gun. Anthony had taken the safety off, the orange part that

makes it noticeable that it was not real. When I held it, I knew by the weight that it wasn't real. While I started to breathe easier, I knew because he took the safety off that it would have been considered a concealed weapon.

We talked to Anthony about how dangerous it is to carry a gun and the implications of him carrying a concealed weapon given his criminal background. During this conversation, Anthony admitted that he was having trouble with youth in his neighborhood. We wrapped up by saying that we would give the BB-gun back and we would not alert police but next time our hands would be tied and we would be forced to call. We called his mother to confirm his story; we also connected him with the onsite mental health counselor at the center.

Thinking back on this incident, I know it's better to find out what's going on than it is to jump to conclusions. If we had called the cops we could have easily solved the problem or we could have exacerbated the issue. If the gun was real someone could have been hurt upon police arrival. Since the gun wasn't real we could have sent a youth to prison for an issue we could have resolved internally. We could have sent him on his way with the BB gun, but, if somebody was going to beat him up and he pulled that out and they had a real gun, then what? The question is, are you putting yourself and others in serious danger? The relationships you have with your youth determine the kind of actions you will take when they make a mistake. I am thankful that the humanitarian in us allowed us to make the right decision.

Student Learning in the Community of Practice

The class spent a long time discussing this case, eventually coming to the conclusion that Ricardo and the other staff did the correct thing by not calling the police. Students then wrote responses to a series of questions. Aside from being in awe of how Ricardo was able to stay calm, two themes emerged from an analysis of the students' responses. The first is that many students began to have a deeper understanding of the importance of 'personal' knowledge in youth work. These excerpts from students' case study write-ups exemplify this theme:

Ricardo grew up involved in youth organizations that helped people who were struggling. Because of this, I believe he is much more understanding of youth like Anthony who appear to be "bad" or a lost cause. Rather than jump to assumptions and take legal action, he knew he could connect on a different level with Anthony and find out what was really going on.

Another student wrote:

I don't know that I would have been able to make the same kind of choices that Ricardo made. I don't bring the same kind of background knowledge to youth work, and would not be able to distinguish between a BB gun and a real gun, or be able to "read" a youth the way Ricardo did and determine he was not trigger-happy. I would probably just have called the police. My priority would have been to keep everyone safe.

Some of the students began to recognize gaps in their knowledge and realized that they would need to be able to execute a different type of strategy if they were to encounter a similar situation.

The second theme relates to the intangible skill of how to appraise and respond to complex problems of practice. Again, drawing on evidence from the students' case study write ups: "Ricardo took into account all the factors affecting the situation, utilized his knowledge and networks, and came up with an appropriate solution." Another student wrote, "Ricardo guided Anthony through the thought processes that youth should be having on their own before making a choice (e.g. the potential consequences of getting caught carrying a concealed weapon, etc.) so that they can see how to arrive at a good decision on their own."

Ricardo's narrative shows how dilemma stories make youth workers' thought process and problem analysis visible to novice and experienced youth workers. Ricardo's participation in the class facilitated the students' learning in that they were able to dialogue directly with him and hear the reactions and questions of youth workers and other students in the class. Their learning was reinforced by being able to read the case and respond to questions in writing. While students read case studies and provide written responses in a typical university course, the fact that Ricardo was a student along with other experienced youth workers allowed for a more authentic form of dialogue in which novices could begin to develop the language and thought process needed to handle pressing dilemmas.

Case Two: Reflection On- and In-Practice Deepens Youth Work Expertise

In this section, I present a dilemma Jessica shared with the class. Jessica, a Latina youth worker in her mid-twenties started doing youth work as a peer leader in her mid-teens. The arc of her thinking about an incident involving a gang-related shooting demonstrates how reflection-on-practice can deepen youth work expertise. In this case, it is not so much that upon reflection she came to a new conclusion. Rather, Jessica began to consider the benefits of intentionally bringing more of herself into her practice.

Jessica's Dilemma

In the summer of 2010, the city experienced an increase in youth violence that affected our organization deeply. Members of the different sites were at war with each other, because their geographic locations in the city predestined them as lifelong enemies. One night two members were involved in a shooting; here is their story.

Ariel, sixteen-years old, had been a Club member for about two years. He really only came to play basketball. Ariel had been one of the trouble makers in the Club, talking back to staff and refusing to follow the rules. The staff had reached out to his parents and quickly realized that they would be of no help since they were gangbanging drug dealers with few positive aspirations. Samuel, a seventeen-year old attended another one of our sites. A member since age twelve, his participation in recent years had diminished as he got more interested in the street life than what the Club had to offer. His grades dropped and his drive for sports had fallen.

Fast forward to the past summer. Ariel, who claims to be a UGE gangster, was on his way home on the city bus when he noticed Samuel, a rival Y-Block gangster. They got off at the same bus stop and after a verbal altercation Samuel pulled out a gun and shot Ariel. Fortunately, the bullet hit his lower leg and he survived.

After Ariel recovered and committed to making changes in his life, he was allowed back into the Club. For a few weeks he did his homework and participated regularly. Unfortunately this was short lived. Ariel was imbedded into the negative lifestyle and his family supported his criminal behavior. Samuel was moved into a relative's home outside of the city. The staff believed that this was the best move for this young person. As for the Club, we took a multistep approach to address this incident. I spoke with the OGs (Original Gangsters), the schools, and other youth. I spoke to the entire Police Gang Unit to discuss the possibility of further retaliation that could happen near or at our organization. I arranged to have a police officer that had been working closely with the Club to have informal discussions with the staff and the youth about tensions between the rival gangs.

Now that this incident is over and we did what we could, I still wonder why what we do in the Club doesn't provide the youth with a bubble of protection outside our walls? So many times when I hear about a particular youth that we have worked with who gets into serious trouble, I shake my head and think, "they are just not those people when they are with us."

From Describing Events to Reflecting on Dilemmas in a Community of Practice

Jessica's dilemma is powerful. Yet, hearing this story left the class with a lot of questions. We wondered how she knew Ariel's family was involved in the thug life, assuming they didn't come out and tell her. We wondered how she was able to ask Original Gangsters about the likelihood of more retaliation. We pushed her to say more about her final statement, "they are just not those people when they are with us." These questions led Jessica to do more thinking and writing that she later shared with the class orally and in a written case study. She wrote about herself, her family, and her insights into these young men's lives. Below is an excerpt of that writing.

Jessica's Personal Reflection

My mom's from a big family and some of my relatives are into the thug life. My house was raided when I was five years old. My father was in jail for drug related charges for most of my early childhood. I'm sure those things come into how I think about the police and violence; I understand where the youth are coming from. My mom realized that she wanted a different life for her kids. She broke away. So, the family dynamics are interesting; those who are heavily involved and those who are not do not spend much time together. There might be a summer BBQ, but when we see each other we don't talk about what we are doing. That's how we can all be together. But I learned early on how to tell when someone is doing something they shouldn't. It's not something that I was told or taught as much as it is something I picked up.

They're Just Not Those People When They're With Us...

Samuel is a perfect example. He is a Hispanic male living in a poor neighborhood; I understand that he needs to defend himself. But, to be the kid who I found out he was after the shooting? No. I just couldn't believe it. All summer when he worked in our camp he spoke professionally and was well mannered. But the cops knew him and they didn't want anything to do with him. We vouched for him and convinced them Samuel had changed. And to his credit, Samuel presented himself that way, so the police gave him the chance. And it worked; he made it through the summer. But then, about a month and a half later, he's shooting a kid.

When expectations are set for young people, that's what they follow. If they are expected to be bad, then they're going to be bad. Here at the Club, we get to know the youth and what they like to do. There is a sense of comfort and safety here and we become like another family. But when something like this happens, I always have to ask myself, why doesn't we do stay with them when they leave our building? I know the answer. They go back into the real world and everyone is expecting them to be this other person, so they just fall into it. Although a lot of the staff have worked here a number of years, we are not their family. Their family will be with them forever. At some point, there is a line drawn between us and them.

Again, we can look at Samuel about how this plays out. His father has been in and out of prison. He is gang involved. So yes, Samuel can be a great kid here with us, but when he goes back home, there's certain

expectations that his family has of him. I think that given the opportunity Samuel would want a different life. But how does he explain that to his father without tarnishing their relationship? Ariel's situation was similar. There weren't many rules set for him and there were rumors that his father had given him a gun to protect himself. You have to look at the whole picture. I don't know at what point we become strong enough, as a youth development movement to break that kind of cycle for our participants.

Like Ricardo's story, Jessica's narrative provides novice youth workers access to her thinking process. Jessica's story also shows how experienced youth workers can benefit from participating in a classroom community of practice. Jessica handled her dilemma competently and compassionately. Yet, when she initially shared her dilemma in class, she was more descriptive than reflective. She told the class what she did and a little about why she did it, but we didn't understand her deeper motivations nor how she knew what to do in the situation. It appeared that she had not interrogated those aspects of her practice before. The students posed a series of questions to her that had her reflect more deeply on the dilemma and her own background. While she was still left with the question about how to have more significant impact with gang-involved young men, her opportunity to reflect more deeply with novice and experienced youth workers in the class prepared her to act more intentionally the next time she worked with this population.

Becoming a Youth Worker in a Classroom-Based Community of Practice

The three abilities I focus on in this paper—how to construct complex and ambiguous youth work problems in a way that they become actionable; how to integrate personal knowledge into practice; and how to reflect in- and on-practice—are three of many complex and somewhat intangible aspects of youth worker expertise. By working together on everyday and extraordinary youth work dilemmas in a classroom-based community of practice these aspects of youth work expertise were communicated. As the experienced youth workers responded to dilemmas, their thought processes became more transparent and tricks of the trade were revealed. The college students were relatively quiet through many of these discussions, feeling a combination of awe, intimidation, and inexperience. Their early participation was indeed 'peripheral.' Yet, as the college students developed relationships with the youth workers and worked through their own dilemmas, they began to develop language and abilities associated with effective youth work practice.

Students were able to acquire these abilities due to the design of our classroom-based community of practice. Having students and youth workers enrolled in the class, holding the class in "practice space" rather than on campus, and focusing teaching and learning on everyday and extraordinary youth work dilemmas helped students become youth workers who understand how to think about and respond to violence-related youth issues and not just about the topic of youth violence. More generally, their stories show how both novices and those with more experience can learn to 'be' a youth worker rather than learn about youth work in a classroom-based community of practice.

References

- Akiva, T. (2005). Turning training into results: The New Youth Program Quality Assessment. *High/Scope ReSource, Fall/Winter*, 21-24.
- Bleakley, A. (2005). Stories as data, data as stories: Making sense of narrative inquiry in clinical education. *Medical Education*, 39, 534-540.
- Bruner, J. (1991). The narrative construction of reality. *Critical inquiry*, 18(1), 1-21.
- Emslie, M. (2009). Researching reflective practice: A case study of youth work education. *Reflective Practice*, 10(4), 417-427.
- Krueger, M. (1997). Using self, story, and intuition to understand child and youth care work. *Child & Youth Care Forum*, 26(3), 153-161.
- Larson, R. & Walker, K. (2010). Dilemmas of practice: Challenges to program quality encountered by youth program leaders. *American Journal of Community Psychology*, 45:338-349.
- Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge, England, Cambridge University Press.
- Ross, L. (2013) Urban Youth Workers' Use of 'Personal Knowledge' in Resolving Complex Dilemmas of Practice. *Child & Youth Services*. 34(3), 267-289.
- Ross, L.(2012). Disrupting Borders: A Case Study in Engaged Pedagogy. *Michigan Journal of Community Service Learning*.19(1), 58-68.
- Schön, D. (1987). *Educating the Reflective Practitioner*. San Francisco: Jossey Bass.
- Starr, B., Yohalem, N., & Gannett, E. (2009). Youth work core competencies: A review of existing frameworks and purposes. Seattle: School's Out Washington & Next Generation Youth Work Coalition.
- Vance, F. (2010). A comparative analysis of competency frameworks for youth workers in the out-of-school time field. *Child Youth Care Forum*, 39(6), 421-441.
- Walker, K. & Gran, C. (2010). Beyond core competencies: Practitioner expertise as a critical component of quality. University of Minnesota Extension Center for Youth Development.

Designing for Engagement in Environmental Science: Becoming "Environmental Citizens"

Susan Bobbitt Nolen, Gavin Tierney, Alexandra Goodell, Nathanie Lee and Robert D. Abbott
Box 36300, University of Washington, Seattle, WA 98195-3600
sunolen@uw.edu; gtierney@uw.edu, alexg79@uw.edu, nalee@uw.edu, abbottr@uw.edu

Abstract: We report Design-Based Implementation Research (DBIR) on a year-long project-based AP Environmental Science curriculum in 11 classrooms in two urban districts. We report its impact on students' engagement, practice-linked identities as environmental citizens and performance on a complex transfer task. Results from the initial design-redesign phase in a suburban district were replicated. Implementation data provided new information about how two design features (positioning students as change agents in their own lives and gradually widening from local to global issues) contributed to engagement and identity development.

Major Issues Addressed

Environmental Science courses naturally fall at the boundary between science and citizenship. Students learn not only the scientific concepts and processes needed to understand the world around them, they learn their roles and responsibilities as citizens in the ongoing interaction between humans and that environment – their “practice-linked identities (Nasir & Hand, 2008).” An unfortunate byproduct of students’ increased knowledge about current environmental problems might be a sense of “doom and gloom.” In making sense of issues of sustainability and human impact on the environment, students may learn to be passive, coming to believe that environmental problems are so big that an individual can have no meaningful impact. In a project-based curriculum, this could lead to student disengagement with real-world and simulated activities that are the spine of the course (Parker, et al, 2013). We designed and tested an experimental, project-based, Advanced Placement Environmental Science (APES) course focused on increasing engagement with scientifically-informed practices by (1) emphasizing students' agency in making decisions that affect the environment and (2) supporting the development of practice-linked identities as environmental citizens. In this PBL curriculum, projects are the spine of the course, meant to provide a context and purpose for all learning activities. After a two-year design-test-redesign-test cycle in a suburban school (Goodell et al 2014), we used DBIR to study the implementation of the curriculum in 11 classrooms in two urban districts and its impact on students' practice-linked identities as environmental citizens, self-reported interest in environmental issues, and performance on a transfer task requiring hypothesis generation, requesting additional information, supporting/refuting hypotheses and proposing solutions to a real-world environmental problem.

Theoretical Framework

Practice-Linked Identities

To understand the connections between students' engagement in project-based environmental science activities and their sense of their role in addressing environmental issues in the world, we turned to Nasir and Hand's (2008) concept of "practice-linked identities." Nasir and Hand define practice-linked identities as "the identities that people come to take on, construct, and embrace that are linked to participation in particular social and cultural practices" in which there is a "sense of connection between the self and the practice" (p. 147). The kind of project-based learning embodied in the APES curriculum is intended to provide opportunities for students to learn how specific practices can have positive or negative effects on the environment. By expansively framing content as transferrable to out-of-school contexts, the designers hoped students would identify with or take up those practices in their daily decision-making (Engle et al., 2012). Considering the environmental impact of everyday decisions and using science to inform one's opinions and actions as a citizen in a democracy we call "environmental citizenship." To the extent that students took up practices and used concepts from the course in this way, we considered them to have begun to develop practice-linked identities as environmental citizens.

Nasir and Hand (2008) describe three dimensions of learning contexts that might support the development of practice-linked identities: "(a) access to the domain as a whole, as well as to specific skills and concepts within it; (b) integral roles and accountability for carrying out those roles; and (c) opportunities to engage in self-expression, to make a unique contribution, and to feel valued and competent in the setting." (p. 248). Contexts that supply all three dimensions would be expected to support higher levels of student engagement as students begin to identify with the practice of particular contexts. Students in the PBL-APES course participated in real-world or simulation projects. As designed, the project-based APES course could provide access to the domains of environmental science and environmental citizenship through framing

environmental science concepts and principles in the context of complex real-world (reduce your family's ecological impact) or simulated real-world (design a sustainable farm, participate in a global energy summit) projects. Students were also clearly assigned roles in each project, and evidence from the year two redesign suggested they saw these roles as "integral" and "accountable" to the project context. The first cycle, Eco-Footprint, cast students in the role of collecting data at home and making proposals to their own families, providing what Barton and Tan have called "hybrid spaces" and opportunity to integrate home and environmental-science related identities and discourses (see also Tzou et al, 2010). In each project, students had latitude to make decisions and express personal interests and choice, and interview data provided evidence that students felt "valued and competent in the setting." Survey data from year two also supported the connection between engagement in the project tasks and end of year interest in the environment, after accounting for initial interest in environmental science. Interview data suggested students had (further) developed their identities as "environmental citizens," able to suggest more specific solutions to problems, reporting more specific instances of transfer to life out of school, and reporting less pessimism and more agency in addressing environmental problems than students in year 1 (Goodell et al, 2014).

Expansive Framing for Transfer and the Development of Interest

The approach to project-based learning used in the APES course is consistent with what Engle and her colleagues have called "expansive framing" to promote transfer (Engle, 2006; Engle et al., 2012; Engle, Nguyen, & Mendelson, 2011). By assigning roles in projects at the beginning of each cycle (e.g., sustainable farmer, natural resource manager), rather than using projects as a way to *apply* learned knowledge, students should learn with the expectation that the knowledge will be useful in the (at least immediate) future. To the extent that students found the projects to be authentic, that is, that they reflected the ways that real people in the world used the concepts and practices they were learning, the projects also should also increase the extent to which students see future value for transfer, increasing engagement and interest. Continuous participation in projects was expected to trigger and sustain student interest in the content, providing an opportunity for students to develop a more stable identification with and interest in environmental issues. This, in turn, could be expected to lead to students seeking out additional information and opportunities to use their environmental science knowledge in other contexts (Renninger, Bacharach, & Posey, 2008).

Context

Design-Based Implementation Research

As we moved our curriculum from a well-resourced suburban district to urban, poverty-impacted districts, we were concerned with a variety of contextual and institutional characteristics that could impact implementation. In our larger project, we have adopted a DBIR approach to understanding and innovating around issues of scaling (Penuel, Fishman, Cheng & Sabelli, 2011). In this brief report, we focus primarily on the effects of implementation differences, but it is important to note that the reasons for those differences represent persistent problems of practice to which the APES curriculum might be adapted, or which districts need to address in some way if implementation in urban contexts is to be successful. For example, the districts in this study differed in the availability of time to regularly meet across sites to discuss and plan implementation. District 1 had little or no history of such collaboration and researcher-organized meetings were not successful in establishing it. District 2 had provided released time and consistent district-level encouragement to collaborate monthly. Constraints of scheduling (block vs. modified block) and differential access to computers and other resources required curriculum flexibility. Other common issues complicated implementation, including teacher and administrator turnover, student absences, and differences in historic achievement levels and expectations for students. The analyses presented here were made possible by the impact of these issues on implementation.

Initial Design and Theoretical Framework

The project-based, Advanced Placement Environmental Science (APES) course provided students with opportunities to prepare for adaptive transfer through multiple, quasi-repetitive opportunities to learn and apply scientific concepts and processes in the context of real-life projects and simulations (Parker et al., 2011; 2013 Schwartz & Bransford, 1998). All instruction occurred in the context of 6 multi-week simulation projects that provided reasons for learning from a variety of sources and experiences, and that attempted to cast students in roles as active problem-solvers in settings beyond the classroom (e.g., as "green" event planners, as environmental consultants, as representatives of various countries). This expansive framing (Engle, Lam, Meyer, & Nix, 2012), was not sufficient to overcome students' push-back based on the enormity of environmental problems and their own lack of power in addressing them. Year 1 feedback from students and teachers led to a significant redesign of the curriculum focused on two main fronts. *First*, we re-ordered the project cycles to help students develop active practice-linked identities by beginning with a local, real-world project (reducing their family's Eco-footprint) and then gradually broadening their perspective across 5 project

cycles, ending with a Global Energy Summit simulation. *Second*, we redesigned several project cycles to enhance the expansive framing of tasks and instruction to improve adaptive transfer to out-of-school settings (Engle et al, 2012). This redesign resulted in students being more likely to report an increased sense of personal responsibility for the environment and adoption of specific sustainable practices, along with demonstrating increased specificity in proposed solutions to environmental problems. This suggests that the design modifications (changing the order of cycles to start with students as agents in their own families, redesigning cycles to emphasize agency and frame for transfer) had an impact on the extent to which suburban students developed practice-linked identities as environmental citizens. Student interview data suggested that beginning with the Eco-Footprint cycle was particularly important in helping students see the importance and environmental impact of their own decisions and practices. The curriculum was modified slightly in year 3 to emphasize the gradual widening of spheres of influence. The third-year project cycles as designed were, in order, Eco-Footprint, My Community Ecology (as state resource managers), Food Systems (as sustainable farm designers), Ocean in Action (as citizens debating the introduction of aquaculture to their island ecosystem), and Negotiation of Nations (representing countries in environmental negotiations).

Implementation in Urban Schools

Suburban students in a relatively "green" region of the US might have brought significant prior knowledge and sustainable values into the course. The teacher, a curriculum co-designer, might have been particularly effective in implementation, especially in year 2. In the third year, the expansion to two urban districts provided an opportunity to test the robustness of the curriculum while investigating local differences in context and implementation that could affect its success. One new district was in the same "green" region, where it might be easier to establish hybrid spaces because of the similar concepts and values of in- and out-of-school contexts. The second was in the middle of the US, in an area of concentrated and long-standing economic difficulty. Both districts served a mix of students but included significant numbers of immigrant families and similarly high rates of free-or-reduced lunch qualification. Teachers differed in the extent to which they remained committed to implementation of the whole curriculum, resulting in differences in the number of cycles (curricular units) taught and the extent to which teachers curtailed project-based learning and supplemented with lecture-based presentations of information. All teachers were new to the curriculum, although some teachers in both districts had experience teaching environmental science or environmental studies courses.

Research Questions

Scaling

We were interested in whether the curriculum design was robust when implementation was expanded to districts that differed from the design environment. Specifically, did urban students develop practice-linked identities as environmental citizens as indicated by their reports of specific practices adopted and a positive rather than pessimistic outlook for citizen action? In addition to collecting interview data on students' self-reported environmental practices, we also assessed end-of-course *Environmental Citizen Identity* with items including interest in environmental issues, feeling that they knew enough to make good environmental decisions and a belief that "people like me" can make a difference. We also investigated processes theoretically involved in the development of practice-linked environmental identities. The extent to which students had opportunities for integral roles and accountability and opportunities for self-expression was measured by class mean levels of *Agentic Involvement*, a scale with items indicating engagement in the project tasks ("I actively participated") and the perception of an integral role in the group's learning ("I usually felt like I contributed to our learning"). Scores on the end-of-course transfer task, an indicator that students had had access to the discipline and *could* use the practices of environmental science to address a complex environmental problem, were examined for links to two engagement measures, Agentic Involvement and Flow (reported concentration, losing track of time), given initial interest in environmental issues (*Initial Environmental Interest*).

Implementation

Two main design changes were of interest: (1) engaging students at the beginning the year with a real-world project aimed at informing and influencing family members, and (2) sustaining engagement and a sense of agency by positioning students as decision-makers in larger and larger spheres of influence over the course of the year-long curriculum. Although all classes implemented the first cycle, only about half implemented all or most of the curriculum. Comparing students in these two groups of classes provided a means of examining whether the curriculum-length design change was necessary or whether similar results could be obtained by "jump-starting" engagement and identification by implementing the Eco-Footprint cycle alone.

Substantiation

Methods

Participants

We analyzed data from the classrooms of four teachers in three schools with a total of six APES sections from District 1 recruited to teach the experimental curriculum. All five APES teachers (five sections, five schools) from District 2 participated. District 2 adopted the experimental APES curriculum as a district with the consent of the teachers. Data from 217 students who completed pre- and post-course surveys were analyzed.

Data Collection

Students completed surveys at the beginning and end of the school year. Scales included Initial Environmental Interest (Fall), and Agentic Involvement, Flow, and Environmental Citizen Identity (Spring). Students also completed a written transfer task, the Complex Scenario Test (CST), which presented a real-world ecological problem (e.g., flooding in Cambodia) and asked students to generate hypotheses, request additional information, support or refute hypotheses, and propose and justify solutions. Semi-structured interviews were conducted with approximately 10% of the students to gather self-reports of transfer, along with other experiences in the course.

Results

Scaling

Analysis of the end of year interviews replicated year 2 (suburban) results, indicating that many students were developing practice-linked identities as environmental citizens. Most students interviewed (69%) provided both specific descriptions of changed personal practices and expressed a belief in their own agency and responsibility in contributing to sustainability. The following example combines these characteristics:

So just, like I said before, like, you know, I just changed. It changed me as a person... learning the facts and how many gallons of water get wasted a day and how it's possible that we have another Tragedy of the Commons. I was like, "You guys can't be showering for like 40 minutes each. And don't leave the water running." So that was really helpful as well. We definitely recycle now a lot. And we have a little separate thing for compost as well (Student 2970).

Survey data were used to test the relationship between being in a class with higher mean levels of Agentic Involvement and individual interest in the environment at the end of the course. The HLM analysis is summarized in Table 1.

Table 1. Effects of classroom-level Agentic Involvement on end-of-year Environmental Identity

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	3.667756	0.038887	94.318	21	<0.001
AGENTICI, γ_{01}	0.549041	0.092206	5.955	21	<0.001
INTRCPT2, γ_{10}	0.231301	0.053339	4.336	136	<0.001

Implementation

Several analyses compared full implementation (at least 4 of 5 cycles) with partial implementation (Eco-Footprint plus 1-2 additional complete cycles). Evidence of impact on practice-linked identities as environmental citizens was more limited in partial-implementation classrooms. All of the students who could not provide specific transfer examples or expressed a pessimistic, non-agentic position in post-course interviews were in classes where the curriculum was only partially implemented. Five of the six full-implementation classes were taught by teachers in District 2, where teachers planned and adapted as a group through the year. This is likely to have supported commitment to a challenging new set of teaching tools.

Full implementation appeared to support engagement and interest more effectively than partial implementation. HLM analysis of survey data, with beginning of year interest as a covariate, revealed that partial-implementation classrooms had lower reported levels of Agentic Involvement, Flow, and end-of-year Environmental Citizen Identity (all $p < .001$). Since all classes implemented Eco-Footprint, these findings suggest that an initial experience being positioned as environmental change agents, though powerful, was not sufficient. Positioning students as agentic in gradually widening of spheres of influence over the course of APES appeared to provide the support necessary to increase practice-linked identification with environmental

issues. Finally, reported engagement was positively related to scores on the Complex Scenario Test across the entire sample. Specifically, HLM analyses found that Agentic Involvement predicted CST scores for Hypothesis Generation, Supporting/Refuting Hypotheses, and Proposing Solutions (all $p < .001$). Flow (task immersion) positively predicted scores for Hypothesis Generation and Proposing Solutions (both $p < .05$). These results suggest that, in addition to supporting identification with environmental citizenship, engagement in project tasks promoted student learning of the practices of environmental science.

Limitations

With the exception of the CST scores, the data reported here come primarily from students' self-reports. Video observation data in several of the project classes indicates that while many students were consistently engaged in project tasks, they may need additional supports in order to engage in the knowledge practices of environmental science. The connections between engagement and performance on the CST may indicate deeper learning for engaged students, but could also mean that engaged students were more likely to exert effort on the CST. These and other issues of implementation are the subject of ongoing research in the larger project.

Relevance to Conference Theme

The study reported here embodies the conference theme is "Learning and Becoming in Practice." In the project-based APES curriculum development project, providing spaces for students to develop as environmental citizens and as environmental scientists has been a key aim. The data presented in this brief paper suggest that the curriculum and its implementation did have an impact on students' practice-linked identities as environmental citizens. Students had opportunities to engage in the practice of both environmental science and citizenship through the project. The developmental structure of the curriculum, moving from real-world hybrid spaces focused on students' personal and family environmental practices through increasingly broader frames of reference, seemed an important contributor to the curriculum's impact. The evidence suggests that student engagement and identity development occurred alongside the knowledge practice of environmental scientists. By purposefully positioning students as decision-makers throughout the curriculum and by focusing on ways to address difficult environmental problems, we were able to reduce the disidentification, pessimism and resistance to deep engagement often seen in environmental science courses.

References

- Barton, A. C., & Tan, E. (2009). Funds of knowledge and discourses and hybrid space. *Journal of Research in Science Teaching*, 46(1), 50-73. doi: 10.1002/tea.20269
- Engle, R. A. (2006). Framing Interactions to Foster Generative Learning: A Situative Explanation of Transfer in a Community of Learners Classroom. *Journal of the Learning Sciences*, 15(4), 451-498.
- Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist*, 47(3), 215-231.
- Engle, R. A., Nguyen, P. D., & Mendelson, A. (2011). The influence of framing on transfer: Initial evidence from a tutoring experiment. *Instructional Science*, 39(5), 603-628. doi: 10.1007/s11251-010-9145-2
- Goodell, A., Nolen, S., Tierney, G., Whitfield, L., Nguyen, D., Becherer, K., Cooper, S. E., Pak, C.-f. (2014). Designing for engagement in environmental science: Developing practice-linked identities. Paper presented at the American Educational Research Association, Philadelphia, PA.
- 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.
- Parker, W., Lo, J., Yeo, A. J., Valencia, S. W., Nguyen, D., Abbott, R., Nolen, S. B., Vye, N. L. (2013, online). Beyond breadth-speed-test: Toward deeper knowing and engagement in an advanced placement course. *American Educational Research Journal*.
- Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing Research and Development at the Intersection of Learning, Implementation, and Design. *Educational Researcher*, 40(7), 331-337.
- Renninger, K. A., Bacharach, J. E., & Posey, S. K. E. (2008). Learner interest and achievement motivation. In M. L. Maehr, S. Karabenick & T. Urdan (Eds.), *Social Psychological Perspective on Motivation and Achievement* (Vol. 15, pp. 425-460). London: Emerald Press.
- Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, 16(4), 475-522.
- Tzou, C., Scalone, G., & Bell, P. (2010). The role of environmental narratives and social positioning in how place gets constructed for and by youth. *Equity & Excellence in Education*, 43(1), 105-119.

Acknowledgements

This work was supported by grants 1261892 and 1251595 (Susan Nolen, PI) from the National Science Foundation, and grants from the George Lucas Educational Foundation (John Bransford, PI).

Student Regulation of Collaborative Learning in Multiple Document Integration

Jun Oshima, Ritsuko Oshima and Keita Fujii

Graduate School of Informatics, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu-shi, Japan
joshima@inf.shizuoka.ac.jp, roshima@inf.shizuoka.ac.jp, kfujii@kai.cs.inf.shizuoka.ac.jp

Abstract: We designed a learning environment for facilitating students' regulation of their collaborative learning in a pre-service teacher course, based on a theoretical model of learning regulation and the jigsaw method. Each student in a group read a different document related to learning environment principles (expert group work). Then, students worked together to integrate the ideas in the documents through discussion (jigsaw group work). We examined the proposed design by analyzing the types of regulation that students used and by interpreting their means of solving social conflicts. Furthermore, the relationship between group characteristics and learning outcomes was explored. Results show that stable engagement in socially shared regulation of learning in the jigsaw group was related to better learning outcomes. In addition, students failed to maintain their socially shared regulation when members had different regulation strategies for solving their social conflicts in groups.

Theoretical Background and Research Aims

When contributing to a collaborative task, learners have to regulate themselves, others, and the group as a whole (Winne, Hadwin, & Perry, 2013). In self-regulated learning, learners regulate their own learning in order to contribute to group performance, based on their individual perception of tasks and their strategic knowledge. In another layer of metacognition, namely, co-regulated learning, learners also regulate themselves in relation to others. Each learner in a group monitors the task perception, goals, and standards of other group members and considers ways their actions and interactions influence one another and the task. In the final layer of metacognition, learners engaged in a collaborative task collectively regulate their group cognition: this is socially shared regulation of learning (SSRL). In Hadwin, Jävelä, and Miller (2011), SSRL is defined as “interdependent or collectively shared regulatory processes, beliefs, and knowledge orchestrated in the service of a co-constructed or shared outcome/product” (p. 69). In SSRL, learners are collaboratively involved in the planning, monitoring, evaluation, and regulation of social, cognitive, and behavioral aspects of their learning. The importance of social regulatory processes for learning in small group settings has been preliminarily but empirically supported by several studies in the last decade. For instance, Järvenoja and Järvelä (2009) found that individual, shared, and other forms of regulation during group interactions can be differentiated and that these different types of regulation are used to maintain group work when students encounter a challenge. Moreover, social regulation of learning is associated with the use of deep-level learning strategies and learning transfer.

By drawing on recent findings on SSRL, it is possible to conduct a more systematic study of how collaboration should be designed. Here, we employ design-based research methodology to investigate how collaboration can lead learners to more successful learning. Based on the reported mechanism of regulation of collaborative learning (Hadwin et al., 2011; Winne et al., 2013), we designed a learning environment to facilitate university students' regulation of collaborative learning in a task called “collaborative reading comprehension” (Oshima & Oshima, 2011); we then analyzed how our design elements helped learners to regulate their collaboration for deeper conceptual understanding.

In collaborative reading comprehension, there are two types of group work. The first type is expert group work, in which learners in a group share the same content (here, a document) and attempt to construct their *own* understanding of it through their interaction. In the expert group, therefore, learners are required to engage appropriately in reciprocal co-regulated learning (named as two-way Co-RL later). A learner must understand the document in order to explain it to others who are unfamiliar with it in the subsequent jigsaw group; consequently, each individual learner in the expert group is more oriented toward their individual perception and goal. Their interaction with other members who study the same document facilitates reciprocal co-regulated learning where learners monitor each other's understanding by expressing their ideas and receiving feedback from others. The second type is jigsaw group work, in which members who studied different documents share and collaboratively integrate ideas from different documents. Each learner is required of contributing to the construction of shared understanding by referring to others' ideas as well as their own ideas. In this group work, each learner is required to engage appropriately in SSRL.

In this paper, we report our first year attempt to design collaborative reading comprehension based on the theoretical framework of learning regulation. We focus our attention on design conjecture rather than on theoretical conjecture (Sandoval, in press). In the early stage of design-based research, the main issue is to examine whether the implemented design would work as expected rather than whether the design would lead to

successful outcomes. To evaluate our design, we take up the following research questions. First, how do students engage in regulation of learning in the expert group and the jigsaw group, and are there any differences in learning regulation between groups? We examined regulation of learning by identifying which level of regulation learners engaged in. In addition, we conducted finer-grained discourse analysis by applying conversation analysis (Schegloff, 2007) to conversation segments that were representative of unique group characteristics. Second, if there are differences between the groups, what learning outcomes result from group differences in regulation of learning? We evaluated the following learning outcomes: learners' written discourse summarizing their ideas in a computer-supported collaborative learning (CSCL) system after collaborative reading comprehension.

Learning Context of This Study

Collaborative reading comprehension was implemented as a part of an intensive course during the summer term of a teacher certification program at a Japanese public university. The course was four consecutive days, and the activity was conducted during the first day and a half. The second author was the instructor. Toward fulfilling the course requirements, 7 third-year undergraduate students and 1 graduate student participated in this study. The goal of the course for students was to understand basic concepts of CSCL in order to apply lesson plans appropriately. The goal for students in the collaborative reading comprehension was to understand basic principles of how to design learning environments (Bransford, Brown, & Cocking, 1999).

Collaborative reading comprehension is an activity structure designed based on the jigsaw method (Aronson & Patnoe, 2011). It encourages learners to engage in collaborative knowledge construction through building an understanding of multiple document resources (see left in Figure 1). Students were first placed in expert groups after listening to an instructor's brief lecture on "the learning environment," the target concept. In each expert group, four students collaboratively read and constructed an understanding of one document, which they explained to others afterward in jigsaw groups. Through expert group collaboration, each student produced a summary by using a prepared worksheet, which would be used as a handout for the explanation in jigsaw groups. After the expert groups finished, jigsaw groups were formed; these groups consisted of one student from each expert group. Students in the jigsaw groups worked to integrate the ideas contained in the four different documents from explanations by the student expert for each document. After discussing the four documents, the students reported how ideas from the documents were related to one another and interpreted them with reference to the basic framework of learning environments in the Knowledge Forum CSCL system.

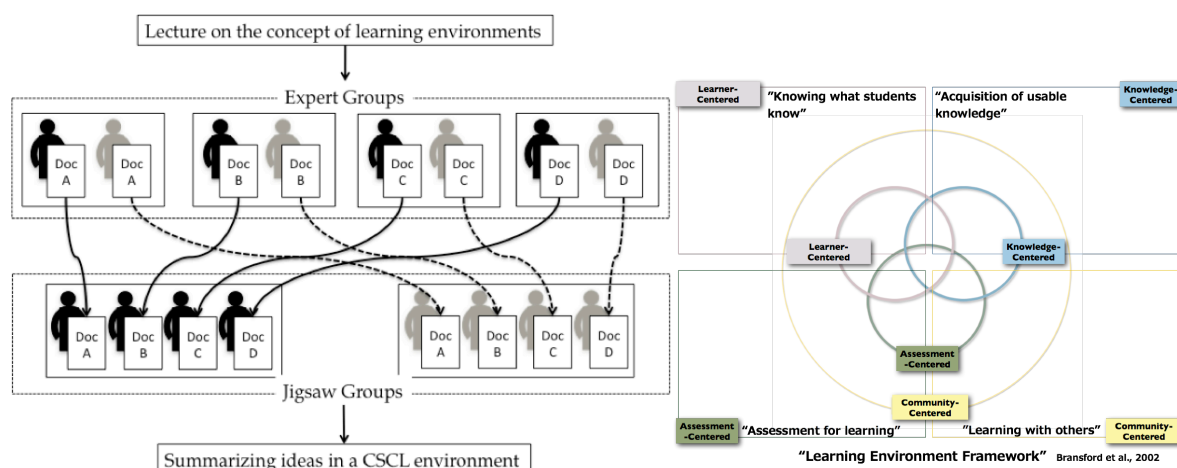


Figure 1. Participatory structure of collaborative reading comprehension (left) and worksheet for students to organize their ideas from documents (right).

Documents used in the activity came from a book on how people learn in Japanese (Inagaki & Hatano, 1989). We selected four chapters of the book, covering four basic principles of learning environments: learner-centered, knowledge-centered, assessment-centered, and community-centered. Each chapter was transcribed to electronic form and reprinted as a standalone document. These documents were given as reading assignments to the expert groups, and we provided students with a worksheet to organize ideas from their document in relation to the learning environments concept (see right panel in Figure 1). In jigsaw groups, students brought the ideas summarized from the documents on their worksheets and discussed how the ideas from different documents could be integrated for advancing their understanding of the principles of learning environments. To facilitate SSRL, we further provided students with a whiteboard, on which a large Venn diagram of the learning environment was drawn and sticky notes so that they could externalize and manipulate their ideas on the shared space.

Results

Data Collection

To examine student regulation in collaborative learning, we collected data in the following way. First, we collected each student worksheet and made an electronic copy (in Portable Document Format). Second, we took pictures of the whiteboard, where they had recorded their ideas by placing sticky notes on the learning environment diagram. Third, we video-recorded the conversations in the expert and jigsaw groups and transcribed them. Finally, students' written thoughts on integrating their ideas from documents in the CSCL system were included. In this study, the transcription of student conversations and the writings in the CSCL system were used for discourse analysis.

Regulation of Collaborative Learning for Students in Expert and Jigsaw Groups: Group Differences

Student conversation was first divided into segments by tasks that learners engaged in. Each segment of conversation was then categorized as one-way co-regulated learning, two-way co-regulated learning, or SSRL. Whenever a specific student intended to regulate collaborative learning and the others followed (whether or not they were willing to), the segment was categorized as one-way co-regulated learning. If the regulation was reciprocal in constructing understanding, the segment was categorized as two-way co-regulated learning. When students shared problems and collaboratively regulated their learning, this was categorized as SSRL. The first and third authors collaboratively conducted the segmentation of transcripts and independently categorized segments. The authors agreed on 70% of categorizations, with disagreements resolved through discussion.

We calculated how much time students spent in regulating their collaborative learning and found the following group differences: (1) two expert groups were more likely to engage in co-regulated learning (98% of total time for expert group 1 and 84% for expert group 4), and the other two were more likely to engage in SSRL (87% for expert group 2 and 79% for expert group 3) and (2) both jigsaw groups spent more than a half of the time on engaging in SSRL (58% in jigsaw group A and 63% in B). These results suggest that our design of participatory structure appropriately promoted learners' regulation in their collaborative learning. On the other hand, we also found that every expert group and jigsaw group A engaged in one-way Co-RL. To further investigate why and how learners performed the unexpected regulation of their collaborative learning, we focused on how regulation of collaboration was related to social conflict. As Järvenoja and Järvelä (2009) demonstrated in their work, students apply different levels of regulation strategies when confronted with social challenges within their groups. In the next section, we analyze segments of conversation in which learners had social conflicts and a learner exerted her/his one-way Co-RL strategy to terminate sequences of conversation turns.

How Students Regulated Social Interaction

We identified segments of conversation where social conflicts were found. The first example was from expert group 1, which was oriented to co-regulated learning. In the following segment, they discussed how a description in their document was related to the learning environment principles. The original conversation was in Japanese; we have provided an English translation below the original.

- 1 A1: だからその、そこ-そういう見方をすればそれが学習...者の特性で
/Well, so...if we take the perspective we discussed, this might be related to learner characteristics.../
- 2 B1: =ああ(1.0)なるほど。() それは学習者単独で[いいかな?評価も入ってる?]
/Oh, I see. So is it OK to say that this is directly related to the learner-centered idea? Or do you think that it might be closer to the assessment-centered idea?
- 3 A1: [う.....ん]
/Hmmm/
- 4 : (8.0)((B1は話を振った後メモを取っている。A1は枠組みの資料を見ながら考えている))
/B1 is taking notes after uttering line #2. A1 is thinking while looking at his framework worksheet./
- 5 A1: いいよ、そんな細かく分けなくて
/No, you (B1) need not think in such a strict way./
- 6 B1: =まあね
/Oh, yeah./
- 7 : (1.0)
- 8 A1: 主張がぶれる
/We would be confused./
- 9 B2: うん
/OK./

- 10 : (23.0)((B1はメモを取っている。A1もメモを取っている様子))
 /B1 is taking notes. A1 is also taking notes./
- 11 A1: ただこの視点は(1.0)え::この本文全体の大筋からはずれてるか
 /This perspective does not seem to be related to the main message in the document, does it?/
- 12 B1: う::ん(。)まあ、でもある。(2.0)°どうしよ。°
 /Ummm, yeah, but what should we do about it?/
- 13 : (4.0)
- 14 A1: まあ、いいよ
 /Well, it's alright./

In the conversation above, the two students had conflicting learning goals. B1 was attempting to understand deeply how the ideas in their document were related to learning environment principles; A1 was focused on task completion. In lines 2 and 12, B1 expressed concern about the relationship between the document ideas and the learning environment principles, but A1 ignored B1's requests for discussion and tried to end the discussion (lines 5, 8, and 14). Although these segments were categorized as one-way co-regulated learning during the previous analysis, here we found through fine-grained conversation analysis that these two students applied different regulation strategies to the conflict in goals. B1 challenged the task-completion goal held by A1 several times, trying an ultimately unsuccessful SSRL strategy. We found similar conversation segments across expert groups oriented to co-regulated learning.

Another example of conversation is given below. In this segment, students in jigsaw group A were examining each other's ideas in front of the whiteboard; these ideas had been written on sticky notes and placed on the whiteboard. A1 explained his idea by referring to his sticky note, and A2 raised a question about A1's original idea and proposed a different interpretation (lines 8 and 9). A2's proposal was then shared with A3 and A4, who responded (lines 10, 11, and 12). However, A1 did not discuss A2's proposal, simply moving his sticky note to the section that A2 had suggested. In the video of this conversation, the other three members seem surprised by A1's act of moving his sticky note. A3 expressed his surprise immediately after noticing A1's act (line 13). This segment of conversation was also categorized as one-way co-regulated learning because A1 was regulated by other group members (particularly by A2). Conversation analysis, however, presents a more detailed picture of what was happening. Three group members used the SSRL strategy to scrutinize A1's original idea and A2's proposed a new interpretation (lines 8–12), A1 applied the co-regulated learning strategy by accepting A2's proposal without discussion (line 13). Thus, we see conflicting regulation strategies, with different goals among group members. Through conversation analysis of expert group and jigsaw group activities, we found that conflicts in the learning goals of members might lead group members to failing to pursue the SSRL strategy.

- 1 A1: 見方によれば、自分の英語能力の低さを::人に見せないようにする、(本人の見えない壁)
 /In a sense, she does not like it when others notice her low level of English literacy./
- 2 : そういう見方もできるわけで、それは学習者によって違うから、学習者の特性による
- 3 え::
- 4 : (1.0)よって起こった()
 /We may usually think so. But, she may have a different perspective, too. It's totally dependent on her learner characteristics.
- 5 A3: (。)あ:::
- 6 A1: で::((自分で貼った付箋紙を確認する))
 /So, ((A1 is looking at his sticky note to confirm what he said.))/
 7 : (3.0)
- 8 A2: えっ、でもそれって結局自分で、自分のことをなんか評価してるんで::、ここ((評価中心))らへんじゃ
- 9 ないんですか?
 /Oh, but in this example the person assesses herself. Isn't this close to the assessment-centered idea?/
- 10 A3: =あ::なるほどね、評価が入ってるか
 /I see. I also think that it might be related to the assessment-centered idea.
- 11 A4: (。)う::ん
 /Let me see.../
- 12 A3: =あ::はいはいはいはい
 /Uh, yeah, I got you (A2)./
- 13 : (6.0)((A1は何も言わず付箋紙を動かす))
 /((A1 quickly moved his sticky note, which had been placed on the learner-centered section, to

the assessment-centered section without consulting other members.))/
 14 A3: あ:::
 /Oh.../

How Group Differences in Regulation Were Related to Learning Outcome

Learning outcome was measured by evaluating the quality of written discourse in the CSCL system after collaborative learning. The reasoning in the written discourse was categorized as either best-fit strategy or knowledge transformation (Bereiter & Scardamalia, 1993). When groups simply wrote how their ideas from documents were fitted to one or more of the four learning environment principles, their reasoning was categorized as a best-fit strategy. When a group attempted to describe the four principles in their own words based on their ideas from the documents, their reasoning was categorized as knowledge transformation. We found that jigsaw group B outperformed A in their number of knowledge transformation ($\chi^2 = 3.07$, $df = 1$, $p < .10$).

Discussion

We designed a collaborative reading comprehension exercise to facilitate knowledge integration in reading documents; this was done on the basis of cognitive models of collaborative learning regulation. Our preliminary analysis for the first year of implementation suggested that students were engaged in socially shared regulation within the designed learning environment (jigsaw participatory structure with shared worksheets and whiteboard), and that they were able to integrate their knowledge from multiple documents when they regulated collaborative learning in the socially shared way. Although we found positive effect from the designed learning environments, our analysis also revealed a problem that kept learners from productive collaboration. As suggested in previous research (e.g., Järvenoja & Järvelä, 2009), learners held and attempted to resolve social conflicts in their group work by exerting their different regulation strategies. The new finding in our study is a description of how group members attempted to solve conflicts but could not maintain the SSRL when members had different regulation strategies or different goals in collaborative learning. Thus, to improve collaborative reading comprehension, a socially shared goal in collaborative learning could be further promoted, and students could be instructed on how to learn collaboratively in a more successful way. In future studies on implementation, we plan to design goal instructions and scaffolds in expert and jigsaw groups.

References

- Aronson, E. & Patnoe, S. (2011). *Cooperation in the classroom: The jigsaw method (3rd ed.)*. London: Pinter & Martin, Ltd.
- Bereiter, C. and Scardamalia, M (1993). *Surpassing Ourselves: An Inquiry into the Nature and Implications of Expertise*. Chicago, IL: Open Court.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies Press.
- 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* (pp. 65-84). New York, NY: Routledge.
- Inagaki, K. & Hatano, G. (1989). How people learn—the world of everyday cognition. Tokyo: Chuo-koronsha. [in Japanese].
- Järvenoja, H. & Järvelä, S. (2009). Emotion control in collaborative learning situations? Do students regulate emotions evoked from social challenges? *British Journal of Educational Psychology*, 79, 463–481.
- Oshima, R. & Oshima, J. (2011, April). Knowledge building for pre-service teachers through collaborative reading comprehension. *Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA*.
- Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *The Journal of the Learning Sciences*, 23, 18–36.
- Schegloff, E. (2007). *Sequence organization in interaction: A primer in Conversation Analysis*. New York: Cambridge University Press.
- Winne, P. H., Hadwin, A. F., & Perry, N. E. (2013). Metacognition and computer-supported collaborative learning. In C. Hmelo-Silver, C. Chinn, C. Chan, & A. O'Donnell (Eds.). (2013). *The international handbook of collaborative learning* (pp. 462–479). New York, NY: Routledge.

Evaluating Lesson Design and Implementation within the ICAP Framework

Rod D. Roscoe (rod.roscoe@asu.edu) and Pedro J. Gutierrez (pgutier1@asu.edu)
Arizona State University, 7271 E. Sonoran Mall, Santa Catalina Hall, Mesa AZ 85212

Ruth Wylie (ruth.wylie@asu.edu) and Michelene T. H. Chi (mtchi@asu.edu)
Arizona State University, Learning Sciences Institute, PO Box 872111, Tempe AZ 85287

Abstract: The ICAP framework provides a theory of cognitive engagement based on overt learning activities that may inform instructional design. In ongoing work to investigate ICAP as a theoretically-grounded instructional design system, classroom teachers participated in a workshop to learn about the framework and design lessons at varying levels. In this paper, we consider how the ICAP framework can be used in the evaluative stages of the instructional design cycle to assess learning tasks and implementation. Using an example lesson from a language arts classroom, we describe the evaluation methodology, consider how analyses revealed potential disconnects between intention and implementation, and discuss how ICAP-based evaluations may advance the design cycle by informing future lesson plans.

Instructional Design within the ICAP Framework

The ICAP framework (Chi, 2009) was initially proposed as a theory for interpreting learning research findings by examining students' *overt learning activities*. In *passive* (P) tasks, learners *receive* information without acting with or upon it (e.g., listening to a lecture). *Active* (A) tasks focus learners' *attention* or *activate* relevant knowledge (e.g., highlighting a text while reading). *Constructive* (C) tasks require *generating* ideas that go beyond the presented information (e.g., self-explaining). Finally, *interactive* (I) tasks involve dialogues in which ideas are *jointly produced* by multiple participants (e.g., revising based on peer feedback). Chi (2009) found that interactive learning tasks tended to support deeper learning than individual constructive tasks; constructive tasks tended to outperform active tasks; and active tasks outperformed passive tasks.

ICAP has recently been applied to instructional design. Although the value of constructive and interactive activities are well known, it is not easy to develop appropriate curricula that foster high levels of engagement (e.g., Allen & Tanner, 2005; Armbruster & Patel, 2009). By focusing on overt learning behaviors, ICAP offers a tractable way for teachers to design lessons that target desired levels of cognitive engagement. In this paper, we consider how ICAP can be applied to the *evaluative stages* of instructional design to assess whether lesson designs and implementation successfully matched pedagogical intentions. In the following sections, we briefly summarize instructional design, describe our evaluation methodology, and discuss a case study using this methodology drawn from a language arts lesson on complex sentences.

Stages of Instructional Design

Instructional design seeks to create lessons and experiences that build learners' understanding of the domain. Many design models incorporate planning, conducting, and evaluating processes. For example, ADDIE (e.g., Chan, 2010) describes a five-stage process of analysis, design, development, implementation, and evaluation. In *analysis*, designers gather data about the domain and learners' needs. In *design* and *development*, designers set learning objectives and create the instructional activities and learning materials. The *implementation* stage puts the lesson plan into action. Teachers deliver lectures, engage students in discussion, assign worksheets or other activities. Finally, *evaluation* examines whether the instruction was delivered as planned and identifies changes needed to improve the design. A key principle is that lesson evaluations inform the next iteration of design.

For successful instruction, the design process must also be paired with a guiding theoretical framework. Such theories provide a perspective for analyzing learner needs and specifying effective learning activities. For example, cognitive load theories posit that learners are constrained by cognitive capacity limitations, and thus instructional materials must be designed to balance sources of productive or harmful load (van Merriënboer, Kirschner, & Kester, 2003). In contrast, Keller's ARCS model (Keller, 2010) emphasizes motivational factors of attention, relevance, confidence, and satisfaction. Instructional materials must be designed to grab attention, make the value of the subject matter salient, and support learners' confidence.

ICAP offers a theoretical perspective for instructional design that focuses on overt learning behaviors for deeper understanding (Chi, 2009). Learners' needs are analyzed with respect to desired levels of cognitive engagement and then overt learning activities can be specified to suit those aims. For example, when recall of key terms is the goal, then passive or active tasks may be sufficient. However, if students must develop a deep and flexible understanding, then constructive or interactive tasks may be better. This paper extends work on ICAP as an instructional design tool by considering how the framework can be used to evaluate lesson design.

Participating teachers created and implemented ICAP-based lessons during their school year and later submitted their materials and examples of students' work to the research team. If ICAP is to be viable for instructional design, then the framework must demonstrate utility for critically reviewing lesson implementation. The evaluation process must be able to reveal adherence to (or departures from) curricular goals and inform ways to improve future lesson plans (i.e., complete the design cycle). In this paper, we apply the ICAP framework to consider two key questions: *Did teachers' lessons fulfill their intentions to provide passive, active, constructive, or interactive learning experiences? Did students' overt activities and responses manifest the target levels of cognitive engagement?*

Method

Teacher Professional Development

Ten middle school and high school teachers participated in a workshop introducing ICAP. Teachers represented a diverse selection of domains, including history, language arts, general science, physics, chemistry, and earth science. Participating teachers reported an average of 9.7 years of teaching experience, including new and veteran teachers. Four teachers had completed or were enrolled in a master's degree program. The workshop began with an introduction to the ICAP framework, including a detailed description of each level, examples of overt activities within each level, and hypothesized cognitive processes associated with each level. In addition, we also discussed logical and practical issues involved with implementing ICAP lessons into a classroom and gave two full days to develop two lessons based on ICAP. Teacher learning and understanding was assessed through pretests, posttests, surveys, and evaluating the lessons they created during the workshop.

The ten teachers, with one exception, each created two lessons within their areas (19 lessons total). For example, an earth science teacher created and contrasted active and interactive lessons on (a) the science of decay and (b) earth systems. *Science of Decay* discussed principles of decomposition, such as the role of key microorganisms (e.g., bacteria and fungi). *Earth Systems* discussed various spheres (e.g., lithosphere) and the effects of natural and man-made events (e.g., volcanoes and fossil fuels). Similarly, one language arts teacher created and contrasted active, constructive, and interactive lessons for (a) analyzing sentence structure and (b) complex sentences. *Analyzing Sentence Structure* taught students to recognize fluency problems, such as choppy sentences and repetitive syntax. *Complex Sentences* taught students about independent and dependent clauses and how to combine two simple sentences into a single sentence using a subordinating conjunction. Such lesson examples demonstrate the rich and challenging concepts covered by the participating teachers. For this paper, and to demonstrate the ICAP evaluation methodology, we focus only on the Complex Sentences lesson.

Evaluation Methodology

In the first phase of evaluation, *Lesson Design Analysis*, we consider how various tasks within a lesson support the intended engagement level. This process has three steps:

1. *Review lesson plans and materials to determine the sequence of tasks that comprise the lesson.* Each task within a lesson may support a different level of engagement, and thus it is desirable to analyze them separately. Multiple tasks might be presented in one assignment (e.g., a workbook with word matching and diagramming tasks) but the level of analysis should be the constituent tasks.
2. *Categorize each task based on the level of cognitive engagement required.* Each task can be evaluated for whether the required overt behaviors are passive, active, constructive, or interactive. For example, matching terms and definitions is "active" whereas drawing a diagram is "constructive."
3. *Categorize the overall lesson based on the pattern of constituent tasks.* The holistic categorization is based on the *highest level of engagement* demonstrated by *at least one-third* of the tasks. For instance, if a lesson has 10 tasks with 3 active (30%), 5 constructive (50%), and 2 interactive (20%), then the lesson is labeled "constructive" because interactive tasks comprise less than 33% of the lesson.

Even well-designed lessons cannot guarantee student compliance. In the second phase, *Student Work Analysis*, we consider how students' overt responses exhibit a given engagement level. This process has three steps:

1. *Obtain and review available records of student work and products.* Student work may include notes taken while watching a video, answers to worksheet problems and questions, video recordings of students' interactions during a lesson, assessment tests, and other materials created by students.
2. *For each activity or task, develop a coding scheme to assess students' exhibited level of cognitive engagement.* The coding process should be specific to the assignment and engagement level. For example, active tasks might be coded based on completeness and constructive tasks might be coded

based on the number of elements generated. In some cases, the analysis may reveal gaps in the data (e.g., no record of students' interactions), which may imply ways to improve future lesson activities.

3. *Implement the coding scheme(s).* Code the data and establish expectations for the range of scores that indicate the target engagement level. For instance, if a worksheet is intended to be “constructive,” what range of scores would support the interpretation that students were indeed constructive?

Table 1: Complex Sentences lesson task list and ICAP categorization (*italics indicate coded actions*).

Task Order and Description	Category
Active Version (<i>n</i> = 20 students)	
1 <i>Copy</i> definitions of “independent clause” and “dependent clause” in a notebook	Active
2 <i>Read</i> pre-made flashcards containing example clauses	Passive
3 Mix and match flashcards to <i>create</i> complex sentences	Constructive
4 <i>Write</i> the generated sentences in a notebook	Active
5 Use flashcards to <i>create</i> groupings based on clause types	Constructive
6 <i>Write</i> grouped clauses in a notebook	Active
7 <i>Glue</i> the notes into a notebook to use as examples	Active
8 Practice <i>combining</i> two short sentences into one complex sentence	Constructive
Constructive Version (<i>n</i> = 29 students)	
1 <i>Generate</i> definitions of “independent” and “dependent”	Constructive
2 <i>Check</i> the definitions using the dictionary	Active
3 Answer <i>comparison</i> questions about their definitions versus the dictionary	Constructive
4 <i>Write</i> the correct definitions in a notebook	Active
5 <i>Generate</i> definitions of “independent clause” and “dependent clause”	Constructive
6 <i>Check</i> the definitions using the dictionary	Active
7 Answer <i>comparison</i> questions about their definitions versus the dictionary	Constructive
8 <i>Glue</i> the notes into a notebook to use as examples	Active
9 <i>Identify</i> correctly punctuated complex sentences and <i>justify</i> the choice	Constructive
10 Teacher <i>tells</i> the students which answers are correct	Passive
11 <i>Create</i> a general rule for using commas in complex sentences	Constructive
12 Practice <i>combining</i> two short sentences into one complex sentence	Constructive
Interactive Version (<i>n</i> = 30 students)	
1 <i>Generate</i> definitions of “independent” and “dependent”	Constructive
2 <i>With a partner, compare</i> definitions and check them with the dictionary	Interactive
3 <i>Generate</i> a final definition based on original, partner, and dictionary versions	Constructive
4 <i>Write</i> the correct definitions in a notebook	Active
5 <i>With a partner, generate</i> definitions of “independent clause” and “dependent clause”	Interactive
6 <i>Check</i> the definitions using the dictionary	Active
7 <i>With a partner, compare</i> definitions and check them with the dictionary	Interactive
8 <i>Glue</i> the notes into a notebook to use as examples	Active
9 <i>With a partner, identify</i> correctly punctuated complex sentences and <i>justify</i> the choice	Interactive
10 Teacher <i>tells</i> the students which answers are correct	Passive
11 <i>With a partner, create</i> a general rule for using commas in complex sentences	Interactive
12 <i>As a whole class, discuss</i> sentence rules and definitions	Interactive
13 <i>Glue</i> the notes into a notebook to use as examples	Active
14 <i>With a partner, practice combining</i> two short sentences into one complex sentence	Interactive

Results

Lesson Design Analysis: Complex Sentences

One teacher designed a lesson to teach 79 6th-grade students about independent clauses, dependent clauses, and sentence combining. The teacher developed three lesson versions to contrast active, constructive, and interactive instruction. For each lesson, the teacher's plans and materials were analyzed to determine the constituent tasks, which were then categorized according to the ICAP level (Table 1). In the Active version, students engaged in 8 tasks. Although 5 tasks (62.5%) were categorized as active, 3 tasks (37.5%) offered meaningful opportunities for constructive activity. Thus, the Active version was *recategorized* as “constructive.” In the Constructive

version, students completed 12 tasks: 7 were constructive (58.3%), 4 were active (33.3%), and 1 was passive (8.3%). Thus, the teacher's original "constructive" label for this lesson was retained. Finally, in the Interactive version, students completed 14 tasks: 7 were interactive (50.5%), 2 were constructive (14.3%), 4 were active (28.6%), and 1 was passive (7.1%). Thus, the teacher's original "interactive" label for this lesson was retained.

This evaluation demonstrated two key points regarding lesson design. First, teachers' intentions may be countered by incorporating tasks of a different level. The teacher's active lesson contained constructive tasks involving generation of examples, resulting in the lesson being relabeled as constructive. Importantly, we are not arguing that lessons should only contain tasks of one type; learning tasks at different levels each have their role to play. Although ICAP does not specify how activities should be sequenced, a student debate (interactive) may be more productive if debaters first articulate their arguments on their own (constructive), which may require reading and note-taking (active) to collect evidence. The core idea, however, is that if the teacher's goal is to support a particular level, then the types of activities included in the lesson must be carefully chosen.

Student Work Analysis

An analysis of all cases of student products was beyond the scope of this paper. To exemplify a student work analysis, we consider students' responses on one worksheet (Figure 1) in which they practiced combining sentences (i.e., in Table 1, see Active task #8, Constructive task #12, and Interactive task #14).

Figure 1 displays two example worksheets from the Complex Sentences lesson. Each worksheet is titled "Combining Two Ideas in a Single Sentence" and includes a cartoon character and a math problem (x 12/12). The instructions state: "For each pair of sentences, write one sentence that includes the ideas from both sentences. Turn one of the sentences into a phrase and place it at the beginning to tell about the subject of your sentence." The worksheets contain six pairs of sentences, each with a handwritten student response. The responses use various conjunctions like "When", "While", "After", and "As" to combine the two ideas from each pair of sentences.

Figure 1. Two example worksheets from the Complex Sentences lesson. The grades provided by the teacher (e.g., "+12/12") were not included in the ICAP evaluation process.

The worksheet contained six pairs of sentences. Students could choose the conjoining word, where to place the conjoining word (i.e., at the start of the first sentence or between the sentences), and whether to maintain or swap the ordering of the original sentences. To code for constructive activity, we examined the variability of students' strategies. Students earned up to 6 points based on whether they used a single *word insertion strategy* for all pairings (0 points) versus varying the strategies equally (6 points). Likewise, students earned up to 6 points based on whether all sentences used one *ordering strategy* (0 points) or varied the ordering strategies equally (6 points). Finally, students earned up to 6 points for *each unique conjoining word* (1 point per word). Higher total scores (i.e., > 12 points) showed more constructive activity by implementing more varied strategies whereas lower scores (i.e., < 6 points) indicated less constructive activity (Table 2).

Not every student completed the assigned tasks ($n = 69$). Across students with available data, average constructive activity for the worksheet was only 6.0 ($SD = 3.3$). Students showed the most constructive activity in terms of varying conjoining word choice rather than other strategies, but students exhibited minimal constructive activity, overall. According to the ICAP framework, students in the interactive version should have outperformed the others, which appeared to be untrue in this case. However, without data on partners'

contributions or dialogue, it was impossible to diagnose what occurred between partners (e.g., whether they were co-constructing knowledge or engaging in off-task conversation).

The results of this student work analysis demonstrated a potential disconnect between lesson intentions and actual implementation by students. Although the task offered several opportunities to constructively explore different sentence combining techniques, students used relatively few of them. Instead, students may have been using only one or two rules for combining sentences by rote instead of thoughtfully constructing new sentences.

Table 2: Mean (and standard deviation) scores for worksheet constructive activity.

Strategy	Active ($n = 19$)	Constructive ($n = 26$)	Interactive ($n = 24$)	F	p
Total Score	5.5 (2.5)	6.8 (3.6)	5.5 (3.6)	1.28	.284
Word Insertion	1.3 (1.5)	2.1 (1.9)	0.8 (1.5)	3.93	.024
Sentence Ordering	0.4 (1.1)	1.6 (2.2)	1.1 (1.9)	2.34	.104
Word Variability	3.8 (1.2)	3.1 (1.1)	3.7 (1.5)	1.92	.155

Conclusion

The methodology presented here suggests that ICAP is useful and viable for instructional design. In this paper, we considered how ICAP principles can be used to assess lesson design and implementation. Such evaluations can reveal how the individual tasks that comprise a lesson may support or undermine broader curriculum goals. For example, learning tasks might support a lower level of cognitive engagement than intended or students may not enact the task at the target level. Such failures of instructional design are not new but ICAP provides a framework for analyzing and specifying these issues in a fine-grained manner.

Importantly, ICAP also informs the analysis and development of new lessons. When combined with evaluation, ICAP can guide iterative lesson design and improvements. For example, in the Complex Sentences lesson, the teacher might revise the instructions for the sentence combining task (Figure 1) to explicitly require constructive use of diverse strategies. Similarly, the teacher might create more precise sequences of active, constructive, or interactive tasks that build from lower to higher levels of cognitive engagement. Although ICAP does not specify an ideal ordering of tasks or engagement levels, researchers could use this methodology to test hypotheses about instructional sequences. That is, one might contrast the efficacy of lessons that progress from low-to-high cognitive engagement, high-to-low cognitive engagement, or stagger the engagement levels.

Any evaluation method is limited by the available data. For instance, no trace data were collected regarding students' dialogue in the interactive version of the Complex Sentences lesson. Thus, it was impossible to evaluate whether students co-constructed ideas while combining sentences. For researchers and educators seeking to use ICAP for instructional design, care must be taken to create and collect diagnostic examples of student work. If students are supposed to be constructive, do their materials allow them to record their generated responses? If students are supposed to be interactive, do their materials allow them to record their individual or mutual contributions to the final products? Addressing such questions further supports iterative instructional goals (e.g., asking partners to record their contributions may encourage interaction) and evaluation goals.

References

- Allen, D., & Tanner, K. (2005). Infusing active learning into the large-enrollment biology class: Seven strategies, from the simple to complex. *Cell Biology Education*, 4, 262-268.
- Armbruster, P., Patel, M., Johnson, E., & Weiss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE-Life Sciences Education*, 8, 203-213.
- Chan, J. F. (2010). *Designing and developing training programs: Pfeiffer essential guides to training basics*. Wiley and Sons: San Francisco, CA.
- Chi, M. T. H. (2009). Active-Constructive-Interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1, 73-105.
- Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS Model approach*. New York: Springer.
- Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off of a learner's mind: Instructional design for complex learning. *Educational Psychologist*, 38, 5-13.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant 305A110090 to Arizona State University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. The authors would like to thank Wan-Ting Huang, Seokmin Kang, Matthew Lancaster, Derek Stark, and David Yaghmourian for their assistance.

Sequencing Sense-Making and Fluency-Building Support for Connection Making between Multiple Graphical Representations

Martina A. Rau, Department of Educational Psychology, University of Wisconsin-Madison, marau@wisc.edu
Vincent Aleven, Human-Computer Interaction Institute, Carnegie Mellon University, aleven@cs.cmu.edu
Nikol Rummel, Institute of Educational Research, Ruhr-Universität Bochum, nikol.rummel@rub.de

Abstract: Multiple graphical representations can significantly improve learning, provided that students make connections between them. In doing so, they need to engage in sense-making processes to build up conceptual understanding of the connections, and in fluency-building processes to fast and effortlessly use perceptual properties to make connections. We investigate how these learning processes interact, and consequently, which learning process should be supported first. We contrast two hypotheses: (1) conceptual understanding facilitates fluency-building processes, and (2) fluency enhances sense-making processes. We conducted an experiment to investigate whether students learn better if they receive sense-making before fluency-building support, or fluency-building before sense-making support. We assessed students' learning outcomes, problem-solving behaviors, conceptual reasoning, and visual attention. Our results show an advantage for supporting sense-making processes before fluency-building processes rather than vice versa. We conclude that instructional materials with multiple representations should first support sense-making processes and then support fluency-building processes in connection making.

Introduction

Instructional materials in science, technology, engineering, and mathematics (STEM) domains often employ multiple graphical representations (e.g., circles and number lines of fractions, ball-and-stick figures and skeleton drawings of molecules), which use visual and perceptual elements rather than symbols to communicate information. A vast literature documents that multiple representations can enhance students' learning, provided that students make connections between them (Ainsworth, 2006; Bodemer & Faust, 2006). However, students tend not to spontaneously make connections, but they need to be supported in doing so (Bodemer & Faust, 2006).

In any domain, learning is likely to include sense-making processes and fluency-building processes. Sense-making processes are learning processes that lead to conceptual understanding by explicit and verbal reasoning. Fluency-building processes lead to more automatic knowledge that can be used fast and effortlessly. Accordingly, with respect to connection making between multiple graphical representations, we distinguish between sense-making processes that lead to conceptual understanding of the connections between graphical representations, and fluency-building processes that lead to the ability to fast and effortlessly make connections between them based on their perceptual characteristics.

Most prior research exclusively focused on supporting *sense-making processes* involved in students' conceptual understanding of connections between multiple representations. This research shows that supporting students to make connections based on elements that – across representations – correspond to one another enhances their learning (e.g., Bodemer & Faust, 2006; Seufert & Brünken). Recent research on perceptual learning has investigated the effects of support for *fluency-building processes* in connection making on students' learning (Kellman, Massey, & Son, 2009). In these studies, students learned to find corresponding representations while focusing on perceptual characteristics of the representations, without conceptually reflecting on the connections. Students who received fluency training showed significantly higher learning gains compared to students who did not receive such training (Kellman et al., 2009). However, Kellman and colleagues did not investigate interactions between conceptual understanding of connections and perceptual fluency.

In a prior experiment, we found that both sense-making and fluency-building support is necessary for students' learning from multiple graphical representations of fractions (Rau, Scheines, Aleven, & Rummel, 2013): neither sense-making or fluency-building support alone were effective, but only the combination of sense-making and fluency-building support significantly enhanced students' learning of fractions (compared to a single-representation control group). To gain further insight into the mechanisms of the interaction between sense-making and fluency-building support, we conducted a mediation analysis based on errors students made during the learning phase (Rau et al., 2013). Our findings indicate that conceptual understanding of connections enhanced students' benefit from fluency-building support. We did not, however, find any evidence that fluency in connection making enhanced students' benefit from sense-making support. This prior research leads to the hypothesis that instruction is most effective if we provide students with sense-making before fluency-building support (*understanding-first hypothesis*). By contrast, if fluency enhances students' benefit from sense-making support, we should provide fluency-building before sense-making support (*fluency-first hypothesis*).

In support of the *fluency-first hypothesis*, one might argue that fluency in connection making between representations reduces cognitive load during learning activities with multiple representations (Koedinger et al., 2012). Indeed, Kellman and colleagues (2009) argue that fluency results from automating the perceptual task of connection making, thereby freeing cognitive resources for subsequent activities. Thus, supporting fluency-building processes first may free cognitive resources that students can then invest in making sense of connections between graphical representations. Thus, the fluency-first hypothesis predicts that instruction is most effective if it provides fluency-building support before sense-making support (fluency-sense condition).

By contrast, the *sense-first hypothesis* proposes that conceptual understanding is necessary for students to attend to relevant features of the graphical representations while they work on fluency-building support. Not having conceptual understanding might lead students to employ inefficient learning strategies (e.g., trial and error), which might impede their benefit from fluency-building support. Indeed, education practice guides seem to implicitly agree with this view. For example, according to the NCTM (2010), understanding of fractions representations is expected by grade 5, but the ability to efficiently work with fractions representations is expected later: by grade 8. In sum, the understanding-first hypothesis predicts that instruction is most effective if it provides sense-making support before fluency-building support (sense-fluency condition).

In our prior study (Rau et al., 2013), we consistently provided sense-making support before fluency-building support. Therefore, the understanding-first and fluency-first hypotheses remain untested. Yet, we expect that the temporal sequence of sense-making and fluency-building support should maximize students' benefit from activities designed to support connection making. This question of optimal sequence is of broad relevance because connection making between multiple graphical representations is critical to students' learning of robust domain knowledge in many STEM domains. Furthermore, in investigating this question, our research helps close the gap between studies that have exclusively focused on sense-making support (e.g., Bodemer & Faust, 2006) or exclusively on fluency-building support (e.g., Kellman et al., 2009). In this paper, we present a lab experiment that contrasts the fluency-first hypothesis and the understanding-first hypothesis.

Experimental Study

Experimental Design and Procedure

To investigate these hypotheses, we conducted a lab experiment that contrasted different sequences of sense-making and fluency-building support. We conducted the experiment within the context of the Fractions Tutor: an intelligent tutoring system for fractions (fractions.cs.cmu.edu). The Fractions Tutor supports learning through problem solving with a variety of interactive graphical representations (Figure 1).

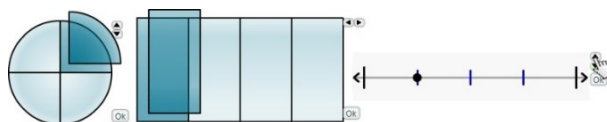


Figure 1. Interactive graphical representations used in the Fractions Tutor: circle, rectangle, and number line

Seventy-four students from grades 3-5 participated in the experiment. Table 1 details the sequence of activities for each condition. Students were randomly assigned to the sense-fluency and the fluency-sense conditions. Both conditions contained the same tutor problems, but they were provided in different orders. Students in the *sense-fluency condition* received sense-making support before fluency-building support. This procedure was implemented for each topic (i.e., equivalence and comparison; see Table 1). By contrast, students in the *fluency-sense condition* received fluency-building support before sense-making support, again for each topic.

The experiment was conducted in two phases. Due to delayed arrival of eye-tracking equipment, 38 students participated in phase 1 of the experiment, without eye tracking. The remaining 36 students worked with the SMI RED 250 remote eye-tracking system. The procedure for phases 1 and 2 was identical except for the collection of interview data (detailed below) and a calibration procedure (1-2 minutes prior to the pretest).

Table 1: Sequence of activities by experimental condition

Activity Type	Sense-fluency condition	Fluency-sense condition
1. Test	Pretest: near / far transfer	Pretest: near / far transfer
2. Tutor: equivalence	<i>Sense-making support</i> : 4 tutor problems	<i>Fluency-building support</i> : 4 tutor problems
	<i>Fluency-building support</i> : 4 tutor problems	<i>Sense-making support</i> : 4 tutor problems
3. Quiz: equivalence	Reproduction-understanding, reproduction-fluency	Reproduction-understanding, reproduction-fluency
4. Tutor: comparison	<i>Sense-making support</i> : 4 tutor problems	<i>Fluency-building support</i> : 4 tutor problems
	<i>Fluency-building support</i> : 4 tutor problems	<i>Sense-making support</i> : 4 tutor problems
5. Quiz: comparison	Reproduction-understanding, reproduction-fluency	Reproduction-understanding, reproduction-fluency
6. Test	Posttest: transfer of fractions knowledge	Posttest: transfer of fractions knowledge
7. Interview	Retrospective interview on tutor problems	Retrospective interview on tutor problems

Materials

Students worked with the Fractions Tutor's units on equivalent fractions and fraction comparison topics. Each tutor activity was designed to support either sense-making or fluency-building processes. *Sense-making support* first presented students with a worked example that used a familiar graphical representation (i.e., circle or rectangle) to illustrate how to solve a problem. Students completed the last step of the worked-example problem themselves. Next, with the worked example still on the screen, students solved an analogous problem with a less familiar representation (i.e., the number line). At the end of the problem, students were prompted to relate the two representations to each other. *Fluency-building support* was similar to Kellman and colleagues' (2009) fluency training for perceptual expertise in connection making. Students were presented with a variety of graphical representations and, using drag-and-drop, had to sort them into sets of equivalent fractions, or order them from smallest to largest. Students were encouraged to solve the problems visually rather than computationally.

Assessments

We assessed learning outcomes with respect to reproduction of conceptual understanding, reproduction of fluency, and transfer of fractions knowledge. *Understanding-reproduction* quiz items assessed students' conceptual understanding of connections between graphical representations. *Fluency-reproduction* quiz items assessed students' fluency in making connections. We assessed students' *transfer of fractions knowledge* based using pretests and posttests. The transfer test included test items without graphical representations.

In addition, we assessed *conceptual reasoning* using retrospective interviews. For all students, we randomly selected one problem of each type for retrospective interviews. In phase 1, the interviewer asked the student immediately after completing the problem how he/she solved each problem step. In phase 2, we used eye-gaze recordings as cues for the interviews. Protocols obtained from retrospective interviews were coded for conceptual and surface-level processing of connections between multiple graphical representations, and for conceptual reasoning about fractions (independent of graphical representations), based on a coding scheme used in our prior research (Rau, Rummel, Aleven, Pacilio, & Tunc-Pekkan, 2012).

Further, we assessed visual attention behaviors during the learning phase using eye tracking. To analyze the eye-tracking data, we created areas of interest (AOIs) for each representation presented in the Fractions Tutor problems. We considered *frequency of switching* between different AOIs, which has been used to indicate perceptual integration (e.g., Johnson & Mayer, 2012). We computed the frequency of switching between graphical representations as the number of times a fixation on one AOI was followed by one on a different AOI. Second, we considered the duration of fixation after the first inspection of an AOI. The first inspection of an AOI is considered to indicate initial processing of material (e.g., Mason et al., 2013). The duration of fixations after the first inspection is considered to reflect intentional processing to integrate the information with other information (e.g., Mason et al., 2013). We computed the *duration of second-inspection fixations* on each AOI as the sum of fixations that occurred after the first fixation on AOIs for area models and number lines.

Finally, we collected *errors rates* while students worked with the tutor based on the tutor logs. We considered a step to be correct if the student solved it without hints or errors. We computed error rates separately for equivalence-sense, equivalence-fluency, comparison-sense, and comparison-fluency problems.

Results

Table 2. Frequencies of utterances coded as representation connections and conceptual reasoning

	Sense-fluency condition	Fluency-sense condition
1.1. Representation-surface	1	0
1.2. Representation-concept-incorrect	6	2
1.3. Representation-concept-correct	17	14
2. Representation-fluency	108	94
3. Concept-correct	305	245

Five students were excluded because they did not complete the Fractions Tutor or were statistical outliers at the pretest. Thus, we analyzed data from $N = 69$ students ($n = 37$ in sense-fluency, $n = 32$ in fluency-sense).

First, we consider the effects on learning outcomes. The *understanding-first hypothesis* predicts that the sense-fluency condition will outperform the fluency-sense condition on measures of fluency in making connections. We conducted repeated measures ANCOVAs with transfer pretest score and time spent on the Fractions Tutor as covariates and performance on the *reproduction-fluency* quiz, averaged across equivalence and comparison topics, as dependent measure. There was a significant advantage of the sense-fluency condition over the fluency-sense condition, $F(1,65) = 4.52$, $p < .05$, $\eta^2 = .07$. The *fluency-first hypothesis* predicts that the fluency-sense condition will outperform the sense-fluency condition on measures of conceptual understanding of connections. Using performance on the *reproduction-understanding* quiz averaged across equivalence and comparison topics as dependent measure, we found no significant effect of condition ($F < 1$). Both hypotheses predict that the optimal sequence of sense-making and fluency-building support will enhance students' perfor-

mance on the *transfer test*. Using performance on the *transfer test* as dependent measure, we found a marginally significant interaction of test time with condition, $F(1,66) = 3.76, p < .10, \eta^2 = .05$, but no significant main effects of condition or test ($F_s < 1$). A posthoc comparison showed a marginal advantage of the sense-fluency condition at the posttest, $F(1,65) = 3.05, p < .10, \eta^2 = .05$.

Second, we consider the effects on conceptual reasoning, based on retrospective interviews. Table 2 depicts the number of utterances of representation connections and conceptual reasoning. Interrater reliability between two independent coders on 33% of the transcripts was substantial ($\kappa = .66$). To investigate the effects on the interview measures, we computed chi-square tests. The *understanding-first hypothesis* predicts that the sense-fluency condition will make more connections while reflecting on fluency-building problems than the fluency-sense condition. Using frequency of *representation-fluency* utterances as the dependent measure, we found no significant difference between conditions ($\chi^2 < 1$). The *fluency-first hypothesis* predicts that the fluency-sense condition will make more connections while reflecting on sense-making problems than the sense-fluency condition. Since there were almost no utterances coded as representation-concept-correct connections (see Table 2), a chi-square test on representation-sense utterances was not warranted. Both hypotheses predict that the optimal sequence of sense-making and fluency-building support will result in more *conceptual reasoning* about fractions. Using concept-correct utterances as dependent measures, we found a significant difference between conditions, $\chi^2(1, N = 550) = 6.55, p < .05$, with the sense-fluency condition exhibiting more conceptual reasoning than the fluency-sense condition. Finally, fluency-building problems elicited significantly more connection-making utterances than the sense-making problems did, $\chi^2(1, N = 241) = 110.25, p < .01$.

Third, we consider effects on visual attention behaviors. Four students were excluded due to a tracking ratio below 70%. Thus, we analyzed data from $N = 24$ students from phase 2 ($n = 12$ in each condition). We conducted repeated measures ANCOVAs with transfer pretest and time spent on the tutor as covariates and visual attention measures on equivalence and comparison problems as dependent variables. The *fluency-first hypothesis* predicts that the fluency-sense condition will exhibit more integrative eye-gaze behaviors while working on sense-making problems. Using frequency of switching as dependent measures, we found no main effect of condition, $F(1, 21) = 1.11, p > .10$, but a significant main effect of topic, $F(1, 21) = 11.19, p < .01, \eta^2 = .35$, and a marginal interaction of topic with condition, $F(1, 21) = 4.09, p < .10, \eta^2 = .16$. Post-hoc comparisons showed that the fluency-sense condition switches marginally more frequently between representations on equivalence-sense problems, $F(1, 21) = 3.53, p < .10, \eta^2 = .14$, but not on comparison-sense problems ($F < 1$). Using duration of second-inspection fixations as dependent measures, we found a significant main effect of condition, $F(1, 21) = 4.43, p < .05, \eta^2 = .17$, and a significant interaction of topic with condition, $F(1, 21) = 7.09, p < .05$, but no significant main effect of topic ($F < 1$). Post-hoc comparisons showed that the sense-fluency condition exhibits significantly longer second-inspection fixations on area models on comparison-sense problems, $F(1, 21) = 5.95, p < .01, \eta^2 = .22$, but not on equivalence-sense problems, $F(1, 21) = 1.43, p > .10$. There were no significant effects of condition on duration of second-inspection fixations on number lines ($ps > .10$). The *understanding-first hypothesis* predicts that the sense-fluency condition will exhibit more integrative eye-gaze behaviors while working on the fluency-building problems. Using eye-gaze measures collected while students worked on fluency-building problems, we found no significant differences between conditions ($F_s < 1$).

Finally, to better understand the effects on visual attention behaviors, we computed correlations with the error rates on equivalence-sense and comparison-sense problems. Frequency of switching on equivalence-sense problems correlated positively with students' error rates ($r = .358, p < .10$), indicating that a higher frequency of switching is associated with lower problem-solving performance. Duration of second-inspection fixations on area models on comparison-sense problems correlated negatively with students' error rates ($r = -.357, p < .10$), indicating that shorter fixations are associated with higher problem-solving performance.

Discussion and Conclusion

We conducted an experiment to contrast the *understanding-first* and the *fluency-first hypotheses*. Results from the *learning outcomes* provide some support for the understanding-first hypothesis. The sense-fluency condition significantly outperformed the fluency-sense condition on the reproduction-fluency quiz and marginally on the transfer posttest. By contrast, our results do not support the fluency-first hypothesis.

The analysis of *process-level measures* provides insights into the mechanisms that may underlie the advantage of the sense-fluency condition. First, the analysis of the *retrospective interviews* shows that the sense-fluency condition engages in more conceptual reasoning than the fluency-sense condition. We also found that fluency-building support elicits significantly more connection-making utterances than sense-making support. (These types of activities involve a large number of representations, and may simply afford greater opportunity for connection making.) A reasonable interpretation of these findings may be that the combination of sense-making and fluency-building support is effective because fluency-building support aids students in making a large number of explicit connections between graphical representations. However, only when students receive sense-making support *before* fluency-building support might they be able to take advantage fluency-building support because they can integrate the perceptual knowledge with conceptual understanding.

The analysis of the *eye-tracking data* provides somewhat less conclusive insights. We found differences only on sense-making problems, not on fluency-building problems. Seemingly consistent with the fluency-first hypothesis, students in the fluency-sense condition (compared to the sense-fluency condition) switched more frequently between representations on equivalence-sense problems. Although frequency of switching is often considered to indicate integrative processing (e.g., Johnson & Mayer, 2012), we found that frequency of switching is (marginally) associated with lower performance on these problems. Rather than indicating integration across representations, frequency of switching may indicate confusion. Inconsistent with the fluency-first hypothesis, the fluency-sense condition shows shorter second-inspection fixations on area models than students in the sense-fluency condition on comparison-sense problems. Shorter second-inspection fixations on area models were associated with lower performance on these problems. This finding might indicate that receiving fluency-building support before sense-making support inhibits students' integration of new information with the area models. These interpretations of the eye-tracking data remain speculative, but they illustrate that we do not yet fully understand which measures of visual attention to consider, or whether they indicate productive or unproductive learning processes. It is crucial, therefore, that we explore the relationship between measures of visual attention and measures of other learning processes, such as the error rates we collected in our study.

In summary, our findings provide some support for the understanding-first hypothesis. Although our results were not uniformly strong, a reasonable interpretation may be that understanding of connections between representations *is necessary* for students' benefit from fluency-building support because it enables students to relate connections between representations to conceptual knowledge about fractions. Thus, we cautiously recommend that instructional designers of multi-representational learning materials provide sense-making *before* fluency-building support to enhance students' acquisition of fluency in connection making and of robust domain knowledge that transfers to novel task types. Whether or not these conclusions hold for other domains than fractions learning remains to be empirically tested. We consider our study to be a first step towards closing the gap between research that has exclusively focused on sense-making processes in connection making (e.g., Bodemer & Faust, 2006) and studies that have mainly focused on perceptual fluency in connection making (e.g., Kellman et al., 2009). Since many STEM domains employ multiple graphical representations and emphasize the importance of both conceptual understanding of the connections between these representations as well as the ability to fluently make connections between them, we anticipate that future research that investigates the interaction between sense-making and fluency-building processes will have broad impact to many domains.

References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183-198.
- Bodemer, D., & Faust, U. (2006). External and mental referencing of multiple representations. *Computers in Human Behavior*, 22(1), 27-42.
- Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, 18(2), 178-191.
- Kellman, P. J., Massey, C. M., & Son, J. Y. (2009). Perceptual learning modules in mathematics: enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science*, 2, 285-305.
- Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, 36(5), 757-798.
- Mason, L., Pluchino, P., & Tornatora, M. C. (2013). Effects of picture labeling on science text processing and learning: evidence from eye movements. *Reading Research Quarterly*, 48(2), 199-214.
- NCTM. (2006). Curriculum focal points for prekindergarten through grade 8 math: A quest for coherence. Reston, VA.
- Rau, M., Rummel, N., Aleven, V., Pacilio, L., & Tunc-Pekkan, Z. (2012). How to schedule multiple graphical representations? A classroom experiment with an intelligent tutoring system for fractions. In J. van Aalst, K. Thompson, M. J. Jacobson & P. Reimann (Eds.), *Proceedings of the 10th ICLS – Vol. 1* (pp. 64-71). Sydney, Australia: ISLS.
- Rau, M. A., Scheines, R., Aleven, V., & Rummel, N. (2013). Does representational understanding enhance fluency or vice versa? Searching for mediation models. In S. K. D'Mello, R. A. Calvo & A. Olney (Eds.), *Proceedings of the 6th International Conference on EDM* (pp. 161-169).
- Seufert, T., & Brünken, R. (2006). Cognitive load and the format of instructional aids for coherence formation. *Applied cognitive psychology*, 20, 321-331.

Acknowledgments

This work was supported by the National Science Foundation, REESE-21851-1-1121307 and by IES, NCER-CASL Award No. R305A120734. We thank Ken Koedinger, Brian Junker, Jay Raspat, Michael Ringenberg, Angela McCarthy, Siyan Zhao, Lavender Yi, Jessica Han, Lisa Kwon, the Datashop and CTAT teams.

Taking DALITE to the Next Level: What Have We Learned from a Web-Based Peer Instruction Application?

Elizabeth S. Charles and Chris Whittaker, Dawson College,
echarles@dawsoncollege.qc.ca, cwhittaker@dawsoncollege.qc.ca

Nathaniel Lasry and Michael Dugdale, John Abbot College,
lasry@johnabbott.qc.ca, michael.dugdale@johnabbott.qc.ca

Kevin Lenton, Vanier College, lentonk@vaniercollege.qc.ca
Sameer Bhatnagar, Dawson College, sbhatnagar@dawsoncollege.qc.ca
Jonathan Guillemette, McGill University, jonathan.guillemette@mail.mcgill.ca

Abstract: The Distributed Active Learning Interactive Technology Environment (DALITE) is a web-based tool designed on the principles of Peer Instruction. DALITE promotes student's self explanation and asynchronous explanation to others. This design-based research project involved practitioners and researchers in the co-design process. In this paper we describe the main features of DALITE, its extended system and its implementation. We report on its effectiveness as a tool for teaching and learning physics.

Introduction

A fundamental concern in science education, and physics in particular, continues to be the difficulty students experience building robust understandings of core principles and concepts along with their ability to use them in different settings – i.e., conceptual change and transfer of learning. Recent studies and practice-based efforts to address these problems tell us that science learning and teaching can benefit from pluralistic approaches (e.g., Treagust & Duit, 2008). These include changes to the ways students engage with the content and each other, the ways teachers orchestrate and use new pedagogical approaches, and the ways we design tools that support these various socio-cognitive and socio-cultural processes.

The Distributed Active Learning Interactive Technology Environment (DALITE) is a web-based tool that aims to promote conceptual learning while working within an asynchronous mode of student engagement. It involves the learners in a variety of tasks including writing explanations for conceptual questions, reflecting on and comparing these explanations to those of peers and experts, and taking part in the social construction of the database repository by voting on the most convincing explanations. It is part of a larger study and system of social constructivist pedagogical practices conceived to promote learning in physics at the postsecondary level. Its design draws on social conceptions of conceptual change, recognition of the role of context (e.g., Engle, 2006) and the success of the practical approach to conceptual learning called Peer Instruction (Mazur, 1997).

The study and evolution of DALITE is a design-based research (DBR) experiment (Hoadley, 2002). Its development team is made up of researchers and practitioners (college-level physics instructors) working together in using a co-design approach. DALITE is currently in its third iteration. In this paper we provide a brief background of the theoretical foundations of its design and features, including the extended system it has been embedded within (tagging and concept mapping activities). Additionally, we provide an evaluation of its effectiveness, to date, and speculate on its potential as a tool to support students' conceptual change as well as teachers' efforts to enact an active learning curriculum.

Background

Investigating how students learn physics has been a perennial concern not only of physics education research (PER) but also of the learning sciences (e.g., diSessa & Sherin, 1998). The body of research generated by both communities confirms that conceptions of the physical world, such as force, motion, and acceleration, are difficult to change with traditional instruction (e.g., Hestenes, 1992). However, studies of social constructivist instruction, popularly referred to as *active learning* pedagogy, report findings of statistically significant gains in students' conceptual understanding in physics and other science disciplines (Meltzer & Thornton, 2012). Of particular interest are implementations that focus on promoting conceptual change by placing an emphasis on intentional reflection (Sinatra & Pintrich 2003). Adding to this are questions about the processes involved in self-explanation (Chi, de Leeuw, Chiu & LaVancher, 1994) versus forms of peer explanations such as reciprocal teaching (Palincsar & Brown, 1984), and other collaborative and discursive practices (Stahl, 2006). In fact, it might be argued that there is value in examining the processes of what might be referred to as "interactive explanation" (Ploetzner, Dillenbourg, Preier & Traum, 1999). This interception between explaining to others, as well as reflecting on one's own explanation provides a power nexus for investigation. We propose that such a nexus is found in the variation on Peer Instruction that is at the heart of our designed intervention, DALITE.

Peer Instruction Approach to Learning

Peer Instruction (PI) is an example of an evidence-based pedagogical innovation popularized by Eric Mazur (Mazur, 1997). Its method of engaging students in scientific discourse focuses on acts of explanation, comparison, and reflection that lead to conceptual change. Meltzer (2013) states that, at the postsecondary level, PI is one of the most widely used active learning approaches in North America. No doubt in large part because of the growing body of research supporting claims of its efficacy in producing statistically significant conceptual gains (e.g., Crouch & Mazur, 2001).

In PI implementations, instructors present students with multiple-choice conceptual questions that the students answer using wireless handheld devices, colloquially referred to as *clickers*. These initial polling activities provide instructors with real-time feedback on the status of students' understanding. Answering these questions allows instructors to know whether or not concepts are known, somewhat known or unknown to students. If conceptual understanding falls within the "known" range (correctly answered by more than 70% of students), the teacher can move forward to another concepts and questions. If it is "unknown" (correctly answered by less than 30% of students), the teacher is advised to revisit the ideas. The real peer-to-peer interactions only come into play with the "somewhat known" concepts (30-70% correctly answered). When responses fall within this range, students are asked to turn to their neighbor and discuss their answers and reasoning. It is arguable that these discursive practices allow students to engage in sense making and intentional reflection on these specific concepts.

Some of the most successful implementations of PI have been those found in large lecture halls with hundreds of students. In such settings, rich discussions can arise because of the larger probabilities of having greater diversity among students, which undoubtedly acts to amplify the cognitive dissonance. However, in smaller classrooms there is often less diversity between students' answers and their understandings leading to a paucity of conceptual discussions. In such cases, PI has not always worked well. Adding to this, there is the question of what happens if we were to take PI online. How would it work if peers cannot interact in real-time? With growing interest in active learning pedagogies, which benefit from having students prepared ahead of class work – i.e., the flipped classrooms – there is added pressure on getting design elements for digital and online learning right.

Digital and Online Instruction of Physics

Computer supported learning environments to promote learning in physics is not new. A major initiative in this area is the Andes project, an intelligent tutoring system for a first year college-level physics course (Gertner, Conati & VanLehn 2000). It coaches students through the problem solving process step by step and provides hints should the student get stuck. Andes and similar tutoring systems are very successful and produce significant learning gains compared to traditional instruction. However, some have criticized Andes, and other similar tutoring systems, for failing to get at deep learning. In particular, three weaknesses have been identified as the failure to get students: (1) to use the language of the discipline (i.e., talk science); (2) to reflect more deeply on the learning; and, (3) to work on developing their conceptual knowledge (Graesser, VanLehn, Rosé, Jordan & Harter, 2001). Viewing these as challenges to be overcome when designing an online learning environment, we consider these as the foundation of our design features.

Building Collective Artifacts

Recent studies conducted by Slotta and colleagues show an increased sense of group regulation (aka agency) and collective responsibility when students contribute to commonly shared "knowledge base" resources (Slotta & Najafi, 2010). In addition it has long been shown that providing students with a sense of contribution has implications for promoting epistemic agency (Scardamalia & Bereiter, 2006).

Our Four Design Principles

In the process we can examine the four design principles we extracted from the literature: (1) promoting the use the language of the discipline (i.e., physics talk); (2) promoting reflections on explanations, one's own and that of others – i.e., the interactive explanation; (3) promoting deeper understandings of conceptual structure; and (4) promoting students' agency and responsibility for examining their peers' arguments and assessing their correctness and quality.

Methods

This project uses a design-based research (DBR) methodology. DBR allows for the design of tools or conceptual models that help us better understand the conditions under which the context and/or the intervention can promote better learning outcomes, and in turn, to adapt these to support better learning (Anderson & Shattuck, 2012). Data collected were for the purpose of informing us on the design principles related to promotion of conceptual understanding and conceptual change. Mixed-methods were used that include: standardized pre-post questionnaires (i.e., the Force Concept Inventory (FCI), Hestenes et al., 1992); and course grades. Qualitative

data on student's conceptual understanding is documented in their DALITE rationales. Student interviews were conducted and include video recordings of think-aloud protocols that help to reveal how DALITE was used and how it was perceived as a tool to promote conceptual learning. Classroom observations were also collected to document the ways in which DALITE was used as part of the instructor's system of active learning practices.

Context and Participants

The study is situated within physics classrooms in three English-speaking colleges in Quebec. Four instructors participated and each of them was also a member of the research team (see Table 1). Student participants (N=168) were first year science majors, ages 17-19, enrolled in one of five sections of a 15-week introductory physics course – approximately equivalent to first-year university. DALITE was assigned weekly as homework via the web. Approximately four to five questions were assigned weekly along with other readings and problem-solving activity. DALITE was brought into the classroom setting regularly, which included having the instructor follow up with the correct answers and elaboration on questions that were identified as challenging. Additionally, it was made part of an extended activity that involved concept mapping and tagging activities. These will be discussed only briefly.

Table 1: Number of instructors and students participating in the study.

	College X	College Y	College Z
# of Instructors	n=1; 2 sections	n=1; 1 section	n=2; 1 section each
# students	n=30 and n=31	n=30	n=36; n=41
Classroom design	Active Learning classroom	Active Learning classroom	Hybrid classrooms

Designing DALITE

DALITE was conceived as a way to harness the benefits of PI and address some missed opportunities. In traditional enactments of PI, student conversations disappear into the ether. Though instructors sometimes overhear conversations, there is no trace left behind. More importantly, instructors seldom know what types of arguments and reasoning are convincing to students. What discursive elements help students change their answers – whether they change towards or away from the correct answer. In addition, there is little documentation of the accumulated database of conceptual questions, and what can be learned from how they are responded to. In short, how the question frames the context. Might there be different effects regarding the timing of assigning questions, might there be an issue of the context of delivery or the wording. Can we promote better forms of *intercontextualization* with the sequencing and design of questions? DALITE as a solution provides students with a diversity of explanations for all the possible answer choices. It allows students to interact with rationales from students at different institutions or even “peers” who took the class previously.

The DALITE infrastructure is made up of the following components: (1) a student registration and software application management; (2) a framework for data mining and tracking of student interactions in real time including the instructional scripts; (3) a central database or repository; and (4) data displays for instructors. The platform uses “Agile” development practices with the aim of ensuring future availability, scalability, and performance. The database repository is composed of two parts: (1) the curriculum content – conceptual multiple-choice questions (sometimes referred to as *concept test* questions); and, (2) the student-generated answers and rationales for these answers.

To date, the curriculum database contains over 120 questions spread across the three main topics generally covered in an introductory physics course – i.e., kinematics, dynamics, and energy and momentum. These questions are designed to be roughly at the first-year university level. Influenced by the Ohio State concept test questions (Lee, Ding, Reay, & Bao, 2011), many questions are organized into sets of three to four questions on a single concept that progressively increase in difficulty. These sets of increasingly difficult multiple-choice problems are built on similar deep structures with different surface features, or similar surface features with different deep structures. Instructors have control over the selection and assignment of questions using a specially designed teacher portal. Each problem set can be customized to meet the perceived knowledge level of the students.

The second database repository, the student-generated rationales, has been developed through a “seeding” process. That is, the database asks about 20 students to answer the questions and write rationales, without working through the full DALITE script. This process enables the first participants in the system to see other students' rationales. However, it places constraints on the development of new questions entering the system. In addition, because rationales are student-generated we believe it necessary to develop a mechanism of cleaning up and categorizing the database. This has led to the implementation of a voting system – students have the option of giving a “thumbs up” to the rationale that convinced them. In doing so, these ratings are a design element. In the future a heuristic will be designed to highlight these popular rationales. Lastly, nonsense rationales will be eliminated (e.g., unreadable text, meaningless strings of symbols).

Lastly, the data display for instructors brings DALITE into the classroom. The display provides an interface to allow the instructor to review students' progress in real-time as well as provide a tool for in-class review. It has proven to be more important than we had thought. We discuss this in the upcoming section.

DALITE Scripts: How Does It Work?

The script for DALITE mirrors much of what we imagine students do when they engage in the discursive practices of PI. As such, the DALITE script consist of the following six steps: (S)elect multiple choice answer; (W)rite rationale; (R)econsider based on alternatives; (R)evote; (V)ote on most convincing; and (R)evuew expert rationale. In step 1, students are presented with a multiple-choice conceptual question. They are asked to select an answer from the multiple choices. In step 2, they then write an explanation for their choice, what we call rationales. In step 3, they are asked to reconsider their answer based on another possible alternative, the aim is to replicate the experience of the “turn to your neighbor” phase in PI. If their answer is incorrect, they are presented with student rationales for the correct answer as well as rationales from other students on the same incorrect answer they chose. The aim of this comparison is to present the contrast and cognitive dissonance of traditional PI. If their answer is correct, they are presented with rationales from other students on the same correct answer as well as student rationales for the most popular wrong answer; the aim of this comparison is to test for fragile understanding or lucky guessing. In step 4, the student is asked to consider whether one of the rationales was particularly convincing, if yes they are asked to vote it “thumbs up”. In step 5, students are asked to re-choose an answer for the original question: either their original answer, or the other answer that was just presented to them, based on the reading of these rationales. Lastly, step 6, they are presented with a normative rationale of an expert, but are not given “the” answer; the aim of this decision being to delay feedback and increase self-regulation of criteria and standards.

Other Features Designed to Extend the DALITE System

In addition to the online components, we consider DALITE to be embedded into an extended system that includes a Tagging and a Concept Mapping tool. The tagging tool is digital and paperbased. It is designed to prompt students' thinking about the deep structure of the content contained in the DALITE questions. As such, this tagging tool takes students through a series of cascading concepts – from general to specific. It starts with a DALITE question, then asks the students to reflect on and identify/tag key concepts, first individually, then collaboratively in small groups.

The concept mapping tool is computer-based, and presently uses C-Map (citation). It takes the opposite approach to the tagging tool. It starts by asking students to work collaboratively to identify connections and state relationships between a restricted set of concepts – in the process creating a concept map. It then asks students to add in the DALITE questions to the appropriate area of the map. At the end of this process, students are asked to work on the maps individually, as a reflection exercise.

These tools were implemented separately. Two section worked with the tagging tool (College X). Meanwhile another two sections worked with the concept mapping tool (one section at each College Y & Z). In all instances students were asked to write rationales for the DALITE questions after the activity.

Building on DALITE's Implementation

DALITE implementation over the five sections of 168 students produced approximately 7182 student-generated rationales. The actual distribution by college is described below (see Table 2). These variations in the number of DALITE questions assigned, with the respective variations in the number of rationales written, allowed us to examine the impact of these different modes of implementation. Results also show a statistical difference between DALITE students (average mean Hake gain = 0.49) compared to non-DALITE comparison group (mean = 0.31). Interestingly, the FCI gains for students in these five sections are near identical (0.59). These gains are calculated as the number of transitions of wrong to right answers divided by those that were initially wrong. These results suggest that using DALITE for conceptual gains may not be dependent on the quantity of questions but more likely the choice of questions – i.e., the difficulty and timing (pre-post instruction).

Table 2: Descriptive statistics on the rationales generated by students in the five sections, across the 3 colleges.

# Rationales written/ student	College X ₁	College X ₂	College Y	College Z ₁	College Z ₂
Mean	51	56	36	40	34
Mode	58	66	48	50	48
Median	56	63	39	45	36

Discussion

In regards to the four design features described earlier: (1) The DALITE rationales show that students have been moving towards identifying what we consider the “trigger feature” when explaining their answers. The

trajectory of their rationales also shows more complete explanations over time. In doing so, these data suggest that the design of repeatedly asking for rationales can promote improved use of physics talk. (2) Our think-aloud protocols and interviews with students show that they have a high level of metacognitive activity when using and discussing their use of DALITE. In fact, one surprising finding is that the young women (?) spent considerably more time reflecting on their explanations. One student, whose first language was not English, stated that the reading of rationales has taught her how to better read and understand physics explanations on the internet, which are frequently in English. 3) Observations of students' tagging activity as well as their interviews suggest that the combined use of DALITE and tagging have promoted their understanding of the deeper, structural similarities between questions, regardless of the surface features. Subsequent assessment activities, referred to as "sorting tasks" show that these students are better able to identify similarities between questions compared counterparts who have not used the DALITE system.

The instructor display of student results has proven to be a very important feature of DALITE. Arguably, it the most practical tool for instructors. It provides immediate and detailed feedback to instructors and greatly supports the flipped classroom method or active learning approach in general. Students are better prepared for class and teachers can identify conceptual issues before class and tailor their class to focus on these specific issues. Also it allows students to focus on both the collective and the individual – what does the class know, what do specific individuals need to know (where are they falling short), a feature we only identified in student interviews.

References

- Anderson, T., & Shattuck, J. (2012). Design-Based Research A Decade of Progress in Education Research?. *Educational Researcher*, 41(1), 16-25.
- Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, 18(3), 439-477.
- Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69, 970.
- diSessa, A. A. & Sherin, B.L. (1998). What changes in conceptual change? *International Journal of Science Education*, 20 (10), 1155-1191
- Engle, R.A. (2006). Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. *Journal of the Learning Sciences*, 15(4), 451-498.
- Gertner, A. S., & VanLehn, K. (2000). Andes: A coached problem solving environment for physics. In *Intelligent Tutoring Systems* (pp. 133-142). Springer Berlin Heidelberg.
- Graesser, A. C., VanLehn, K., Rosé, C. P., Jordan, P. W., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI magazine*, 22(4), 39.
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The physics teacher*, 30, 141.
- Hoadley, C. (2002). Creating context: Design-based research in creating and understanding CSCL. In G. Stahl (Ed.), *Computer Support for Collaborative Learning*, (pp. 453-462). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lee, A., Ding, L., Reay, N. W., & Bao, L. (2011). Single-concept clicker question sequences. *The Physics Teacher*, 49(6), 385-389.
- Mazur, E. (1997). *Peer Instruction : A User's Manual*. Upper Saddle River, NJ: Prentice Hall
- Meltzer, D. E., & Thornton, R. K. (2012). Resource letter ALIP-1: active-learning instruction in physics. *American journal of physics*, 80(6), 478-496.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and instruction*, 1(2), 117-175.
- Pfundt, H., & Duit, R. (2009). Bibliography: Students' alternative frameworks and science education. Kiel, FGR: Institute for Science Education.
- Ploetzner, R., Dillenbourg, P., Preier, M., & Traum, D. (1999). Learning by explaining to oneself and to others. *Collaborative learning: Cognitive and computational approaches*, 103-121.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. *The Cambridge handbook of the learning sciences*, 97-115.
- Sinatra, G. M., & Pintrich, P. R. (2003). The role of intentions in conceptual change learning. *Intentional conceptual change*, 1-18.
- Slotta, J. D., & Najafi, H. (2010). Knowledge communities in the classroom. *International encyclopedia of education*, 8, 189-196.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.
- Treagust, D. F., & Duit, R. (2008). Conceptual change: A discussion of theoretical, methodological and practical challenges for science education. *Cultural Studies of Science Education*. Published online 29 March 2008: doi:10.1007/s11422-008-9090-4.

Representational Competence and Spatial Thinking in STEM

Mike Stieff, Matthew Lira and Dane DeSutter, University of Illinois-Chicago
mstieff@uic.edu, mlira2@uic.edu, ddesut2@uic.edu

Abstract: Spatial ability predicts success in STEM fields, particularly chemistry. As such, new educational models have called for learning environments that improve spatial ability. These environments neglect the role of representational competence in supporting spatial thinking in individual STEM fields. This short paper reports a preliminary investigation concerning the unique contribution of representational competence to spatial thinking in the discipline of organic chemistry. Using authentic disciplinary tasks we show that student achievement and response time depends more upon their developing representational competence in chemistry than mental rotation ability and that the format of a disciplinary representation can significantly mediate spatial thinking. Given these findings we argue that new learning environments that target representational competence may be more effective at supporting spatial thinking than those that attempt to train generic spatial ability.

Spatial Thinking in STEM Disciplines

Spatial thinking is a central component of problem solving at all levels of STEM (Science, Technology, Engineering, and Mathematics) instruction (National Research Council, 2006). In both beginning and advanced classrooms, students are tasked with identifying important spatial relationships relevant to scientific concepts and predicting how transformations of those relationships affect physical and biological systems. Spatial thinking ranges in complexity from reasoning about simple geometric relationships, such as the distance between two points on a geologic map, to complex spatiotemporal dynamics, such as reasoning about chemical reaction processes in three-dimensional space. Spatial thinking is challenging and has been cited as a primary barrier to success in STEM classrooms and to entry in STEM careers (Humphreys, Lubinski, & Yao, 1993; Lubinski, 2010; Wai, Lubinski, & Benbow, 2009). Students who struggle with spatial thinking in entry-level STEM courses are less likely to enjoy STEM instruction and pursue STEM professions less frequently (Shea, Lubinski, & Benbow, 2001).

The emphasis on spatial thinking in STEM disciplines suggests that individual differences in spatial ability are related to STEM achievement. Indeed, correlation studies indicate that achievement in several STEM fields is, to some degree, dependent on spatial ability, such as mental rotation and perspective taking (Kozhevnikov, Hegarty, & Mayer, 2002; Pribyl & Bodner, 1987; Sorby, 2009; Turner & Lindsay, 2003). Many STEM achievement assessments require students to visualize the spatial transformation of complex structures, mentally rotate imagined objects to compare spatial features, and assume imagined perspectives to draw two-dimensional diagrams from different orientations. Thus, students who score poorly on several spatial ability measures (e.g., mental rotation, perspective taking, or spatial visualization) are seen to perform poorly in STEM fields that emphasize spatial reasoning (Shea et al., 2001).

The primacy of spatial thinking in the STEM classroom has given rise to a body of learning environments that attempt to support and train spatial thinking in different disciplinary contexts. These studies are quite diverse in their goals and design principles. Designs range from those that attempt to train basic visuo-spatial ability skills through sustained practice (Miller & Halpern, in press; Sorby, 2001) to those that scaffold spatial thinking with concrete and virtual models (Stull, Barrett, & Hegarty, 2013; Stull, Hegarty, Dixon, & Stieff, 2012). These efforts have yet to result in large improvements in STEM achievement or degree attainment, although several individual designs show promise for improving spatial thinking in general and on specific tasks (Uttal et al., 2013). As such, curricular and instructional designs that aim to improve spatial thinking in STEM disciplines remain important targets of current reform efforts.

Despite the correlation between spatial ability and STEM achievement, it is not clear if spatial ability is the primary factor that determines success in STEM achievement. Recent studies suggest that student success in STEM fields is partially dependent upon representational competence (Kozma & Russell, 1997; Stieff, 2011; Stull et al., 2012) and discipline-specific problem solving strategies (Schwartz & Black, 1996; Stieff, 2007; Stieff, Hegarty, & Dixon, 2010). Representational competence comprises a set of skills that include the ability to analyze features of a representation, transform one representation into another, generate novel representations, explain the utility of a given representation, and explain the unique affordances of different representations (diSessa & Sherin, 2000; Russell et al., 1997). These skills are necessary for students to interpret and relate the wide variety of external representations across STEM domains that highlight or mask spatial information to varying degrees. For example, Figure 1 depicts an example assessment item from organic chemistry that requires students to translate from a *dash-wedge perspective formula* to a *Newman projection* to identify high and low energy spatial conformations of an organic molecule. Students' with limited representational

competence who are unable to perform the initial translation are unlikely to identify how the different spatial conformations correlate with energy (Raje & Stieff, 2009, April). Innovative learning environments that aim to improve representational competence may constitute more productive avenues to support student success than those that emphasize generic spatial thinking. The skills composing representational competence reflect authentic disciplinary practices and ways of knowing not addressed by an environment that emphasizes abstract spatial thinking.

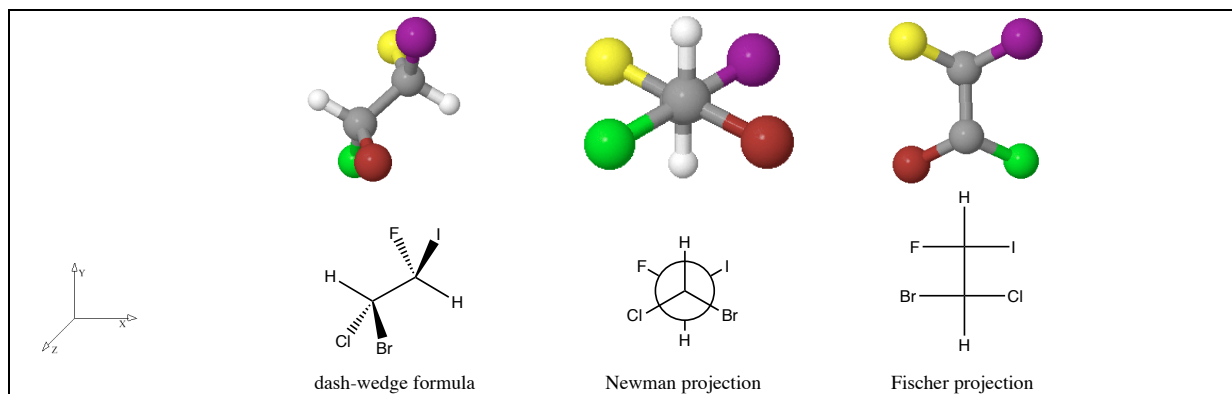


Figure 1. The dash-wedge, Newman, and Fischer representations each use different formalisms to represent the same three-dimensional information in two-dimensional diagrams. In the figure, a concrete molecular model of (1*S*, 2*R*)-1-bromo-1-chloro-2-fluoro-2-iodoethane has been spatially transformed to align with each respective diagram. The model depicted here uses colors to represent individual elements, which are represented symbolically in the diagrams. Grey = carbon (C), white = hydrogen (H), green = chlorine (Cl), brown = bromine (Br), yellow = fluorine (F) purple = iodine (I).

Present Study

In this investigation we explore the relationship between spatial ability and representational competence in an experimental study of spatial thinking in the STEM discipline of organic chemistry. Arguably, organic chemistry emphasizes spatial thinking more so than any other post-secondary STEM course, as the discipline's overarching learning objectives include identifying spatial relationships in hydrocarbons and explaining the chemical and physical properties of a compound resulting from molecular structure. To investigate the relationship between spatial ability and representational competence we compared student achievement on three assessments that varied by representational format. Formats included (1) only three-dimensional models that made spatial relationships salient using shading and perspectival cues (3D-Model), (2) only two dimensional diagrams that made spatial relationships implicit using disciplinary formalisms (2D-Diagram), and (3) mixed formats that include multiple representations of a molecular structure (2D-3D Mixed). Each assessment was constructed of items that required students to make similarity judgments about pairs of molecular representations to determine whether each pair represented the same molecule or a mirror image (i.e., enantiomeric relationship). Such items are authentic assessment items employed in organic chemistry classrooms to evaluate student knowledge of stereochemistry (Stieff, 2007). Using accuracy and response time data we predicted that student achievement would be higher on the assessments that included a single representation whether a 3D-Model or a 2D-Diagram than on the mixed representation assessment. Conversely, we predicted that response time would be higher on the mixed representation assessment. We base these predictions on the assumption that representational competence, as opposed to spatial ability, is the primary barrier to spatial thinking in the STEM disciplines. In other words, student failure will result from challenges relating multiple representations of spatial information independent of their mental rotation ability.

Method

Participants

Twenty-eight (16 female) undergraduate chemistry students participated on a volunteer basis. Each participant was recruited from the population of students who had completed at least five weeks of instruction in organic chemistry at a major research university. All students had received instruction in content assessed in the study.

Instruments

Three tests of mental rotation that varied by the included representation (3D-Model, 2D-Diagram, 2D-3D Mixed) were constructed and employed (see Figure 2 for examples). Each test included stimuli that consisted of pairs of molecular representations that were either identical or mirror image reflections. Each test contained

six asymmetrical stimuli rotated in the picture plane. Identical object pairs were presented once at five unique angles ranging from 0 to 180 degrees in 45-degree increments. Six mirror-reflected object pairs were presented at three randomly selected angles ranging from 0 degrees to 180 degrees in 45-degree increments. Thus, each test included 48 trials. The design was not balanced with regard to mirror image pairs, which were not analyzed: Such pairs were included to prevent participants from assuming that all pairs were identical. Mental rotation was evaluated by the Vandenberg Test of Mental Rotation (Vandenberg & Kuse, 1978).

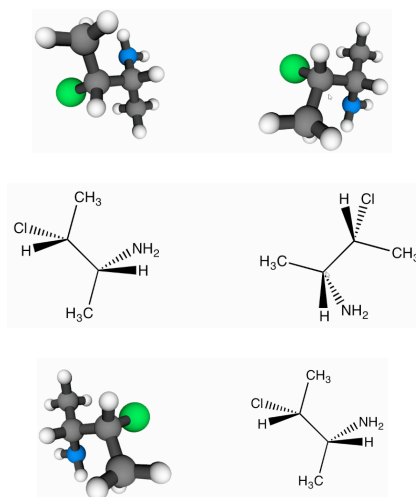


Figure 2. Example stimuli: From top to bottom: 3D-Model, 2D-Diagram, and 2D-3D Mixed item.

Procedure

The study utilized a repeated-measures design with representational format (2D vs. 3D vs. 2D-3D) as a within-subjects variable. All participants completed all blocks with each block randomly presented first. The experiment was presented with ePrime v.2.0. First, participants viewed a screen instructing them to compare stimuli pairs to determine whether each pair contained identical structures. Participants cued the presentation of each stimulus by pressing “SPACE” and responded that pairs contained identical objects by pressing “I” or pressing “E” if they were mirror images (or enantiomers). Participants completed 6 practice trials followed by 144 experimental trials. All participants completed each test in approximately 30-40 minutes during which the keyed response time from stimulus onset was recorded for each trial. The mental rotation test was then administered. Participants received \$20 USD.

Results

First, accuracy was calculated as the average number of correct similarity judgments on identical pairs and analyzed via repeated-measures ANOVA with representational format as the within-subjects variable controlling for mental rotation ability. A main-effect of representational format ($F(2,26) = 28.9, p < .001, \eta_p^2 = .53$) and spatial ability ($F(2,26) = 8.70, p = .007, \eta_p^2 = .25$) was observed. As illustrated in Figure 3, planned contrasts revealed that students performed worst when evaluating 2D-3D Mixed Representations ($M = .77, SD = .15$) than they did evaluating 3D-Models ($M = .89, SD = .07, F(1, 26) = 8.7, p = .007, \eta_p^2 = .25$) or 2D-Diagrams ($M = .83, SD = .15, F(1,26) = 8.51, p = .007$). There was no difference in accuracy between 2D and 3D conditions, $F(1,26) = 1.96, p = .17$. Second, response time was calculated as the average time to respond to a stimulus pair and compared across assessments as above. A significant main effect of representational format was observed ($F(2,27) = 136.68, p < .001, \eta_p^2 = .73$) but no relationship between spatial ability and response time was evident ($F(2,26) = .51, p = .48$). As illustrated in Figure 4, planned contrasts revealed that students responded slower to 2D-3D Mixed Representations ($M = 17853\text{ms}, SD = 5068\text{ms}$) than to either 3D-Models ($M = 8764\text{ms}, SD = 3395\text{ms}, F(1,26) = 64.2, p < .001, \eta_p^2 = .71$) or 2D-Diagrams ($M = 10759\text{ms}, SD = 4018\text{ms}, F(1,26) = 136.7, p < .001, \eta_p^2 = .835$).

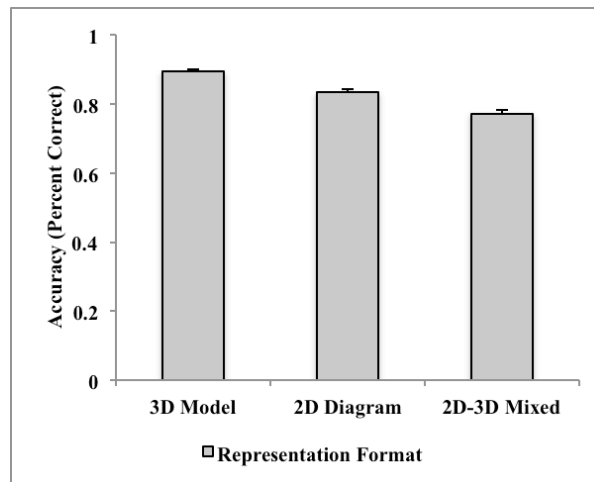


Figure 3. Participant accuracy across conditions. Student achievement was significantly higher in the 3D-Model condition than each of the other two conditions. Performance was lowest in the 2D-3D Mixed condition.

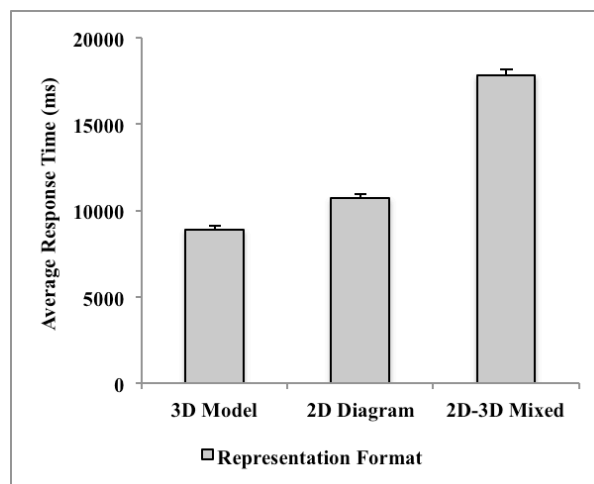


Figure 4. Participant response time across conditions. Response time was significantly higher in the 2D-3D Mixed condition than each of the other two conditions. Performance was lowest in the 3D-Model condition.

Conclusions

These results are consistent with our initial predictions that students would perform better and faster when comparing representations with similar formats (2D, 3D) than those with mixed formats (2D-3D). As above, all students were highly accurate in their evaluation of identical pairs; a meaningful decrement in performance was observed only for assessment items that contained mixed representational formats. Ostensibly, this increase in response time reflects the additional processing demands required to translate between representations to make an identity judgment. Finally, although mental rotation ability correlated moderately with achievement, the results suggest that this correlation depends upon how salient spatial information appears in disciplinary representations. Indeed, our analysis revealed that the format of the disciplinary representations explained more than twice the variance in achievement than spatial ability. Thus, these results suggest that mental rotation *and* representational competence are necessary to solve mixed format items. This result is consistent with prior research that has demonstrated the task-specificity of strategy use (Stieff, 2007) and suggests that future learning environments that target representational competence may be more effective at supporting spatial thinking in STEM disciplines than those that attempt to train generic spatial ability in isolation.

References

- diSessa, A. A., & Sherin, B. L. (2000). Meta-representation: An introduction. *Journal of Mathematical Behavior*, 19(4), 385-398.
- Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology*, 78(2), 250-261.

- Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Spatial abilities in problem solving in kinematics. In M. Anderson, B. Meyer & P. Olivier (Eds.), *Diagrammatic representation and reasoning* (pp. 155-171). Berlin: Springer-Verlag.
- Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, 34(9), 949-968.
- Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. Gilbert (Ed.), *Visualization in science education* (pp. 121-146). London: Kluwer.
- Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences*, 49, 344-351.
- Miller, D. I., & Halpern, D. F. (in press). Can spatial training improve long-term outcomes for gifted STEM undergraduates? *Learning and Individual Differences*.
- National Research Council. (2006). *Learning to think spatially*. Washington, D.C.: National Academies Press.
- Pribyl, J. R., & Bodner, G. M. (1987). Spatial ability and its role in organic chemistry: A study of four organic courses. *Journal of Research in Science Teaching*, 24(3), 229-240.
- Raje, S., & Stieff, M. (2009, April). *An examination of the cognitive mechanisms underlying chemistry misconceptions*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Russell, J., Kozma, R., Jones, T., Wykof, J., Marx, N., & Davis, J. (1997). Use of simultaneous-synchronized macroscopic, microscopic, and symbolic representations to enhance the teaching and learning of chemical concepts. *Journal of Chemical Education*, 74(3), 330-334.
- Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science*, 20, 457-497.
- Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, 93(3), 604-614.
- Sorby, S. A. (2001). A course in spatial visualization and its impact on the retention of female engineering students. *Journal of Women and Minorities in Science and Engineering*, 7(2), 153-172.
- Sorby, S. A. (2009). Education research in developing 3-D spatial skills for engineering students. *International Journal of Science Education*, 31(3), 459-480.
- Stieff, M. (2007). Mental rotation and diagrammatic reasoning in science. *Learning and Instruction*, 17(2), 219-234.
- Stieff, M. (2011). Fostering representational competence through argumentation with multi-representational displays *Proceedings of the 9th International Conference on Computer-Supported Collaborative Learning* (Vol. 1, pp. 288-295): Erlbaum.
- Stieff, M., Hegarty, M., & Dixon, B. L. (2010). Alternative strategies for spatial reasoning with diagrams. In A. K. Goel, M. Jamnik & N. H. Narayanan (Eds.), *Diagrammatic Representation and Inference* (pp. 115-127). Berlin: Springer.
- Stull, A. T., Barrett, T., & Hegarty, M. (2013). Usability of concrete and virtual models in chemistry instruction. *Computers in Human Behavior*, 29, 2546-2556.
- Stull, A. T., Hegarty, M., Dixon, B. L., & Stieff, M. (2012). Use it or lose it: Representational translation with concrete models. *Cognition & Instruction*, 30(4), 404-434.
- Tien, L. T., Roth, V., & Kampmeier, J. A. (2002). Implementation of a peer-led team learning instructional approach in an undergraduate organic chemistry course. *Journal of Research in Science Teaching*, 39(7), 606-632.
- Turner, R. C., & Lindsay, H. A. (2003). Gender differences in cognitive and noncognitive factors related to achievement in organic chemistry. *Journal of Chemical Education*, 80(5), 563-568.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A., Warren, C., & Newcombe, N. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2), 352-402.
- Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations: A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599-604.
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over fifty years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101, 817-835.

Acknowledgements

The research reported here was supported by a grant from the National Science Foundation (REESE 1102349). The opinions expressed are those of the authors and do not represent the views of these agencies. The authors thank Mary O'Donnell for her assistance with data collection and entry.

Expansive Framing and Preparation for Future Learning in Middle-School Computer Science

Shuchi Grover and Roy Pea, Graduate School of Education, Stanford University, Stanford, CA
Stephen Cooper, Computer Science Department, Stanford University, Stanford, CA
Email: shuchig@stanford.edu, roypea@stanford.edu, coopers@stanford.edu

Abstract: Educators aspire to transfer of learning as a goal of their teaching. Expansive Framing and Preparation for Future Learning (PFL) are new perspectives on how to foster and assess transfer. As computing education makes its way into K-12 schools, efforts are underway to introduce children to programming in block-based environments like Scratch and Alice. This paper reports on a design-based research in progress that employs ideas of Expansive Framing and PFL to pedagogy and assessments in a middle school introductory CS curriculum that uses Scratch, and includes designed measures for evaluating how well it prepares students for success in future computing experiences with text-based programming.

Introduction

Computational Thinking (Wing, 2006), now widely recognized as a necessary skill for today's generation of learners, is increasingly being introduced in middle and early high school via programming in block-based introductory environments such as Scratch, Alice, and MIT App Inventor, among others (Grover & Pea, 2013). Ideally, educators would like these first experiences to be framed in such a way that learners can transfer their learning successfully to future computational experiences, which are likely to be in the context of higher-order, text-based programming languages. However, mediating transfer of learning from one context to another or even from one grade to the next is known to be difficult in STEM domains including programming (Kurland, Pea, Clement, & Mawby, 1986; Pea, 1987).

In designing and evaluating a 6-week introductory CS curriculum for middle school, we were guided by both work on “expansive framing” as a pedagogy to promote transfer and the *Preparation for Future Learning* (PFL) approach to assessing transfer. Expansive framing and PFL have yet to be investigated in the context of curricular interventions in CS Education. This paper describes our investigations around introductory CS instruction designed to nurture deep learning and expansive framing of CS constructs rather than focusing solely on the shallow surface features of the block-based programming environment, with the rationale that such an approach will prepare students for better success in future computational experiences, especially with text-based programming languages. The research effort includes the design and use of “dynamic” PFL assessments to assess how well students can apply their understanding of computing constructs learned in the context of Scratch to algorithmic solutions expressed in a more advanced text-based programming language.

Research Framework: Mediating Transfer and Preparation for Future Learning (PFL)

Transfer and PFL are embodied in notions of *deeper learning*, a topic at the center of a recent National Academy of Sciences synthesis report on developing transferable knowledge and skills for 21st century life and work (Pellegrino & Hilton, 2012). The need for classroom learning to be transferrable so it can be applied in practice and contexts outside of school is in keeping with the idea of ‘learning and becoming in practice’, the theme for the 11th International Conference of the Learning Sciences, 2014. The seminal literature on how people learn points to several critical features of teaching and learning that affect people’s ability to transfer and suggests ways to facilitate transfer (Bransford, Brown & Cocking, 2000). While there is no single prescribed strategy for fostering learning for transfer, suggested instructional strategies aim to help students assemble new mental platforms for subsequent learning. As Engle et al. (2012) note, *expansive framing* fosters an expectation that students will continue to use what they learn later, create links between learning and transfer contexts so that prior learning is viewed as relevant during potential transfer contexts; and encourage learners to draw on their prior knowledge during learning, which may involve them transferring in additional examples and making generalizations. Educational psychologists also argue, “learners who compare cases will develop a more general problem-solving schema that primarily captures the common structure of the cases rather than the surface elements” (Gentner, Loewenstein, & Thompson, 2003). Consequently, in contrast to cases studied individually, analogous representations compared as part of a more expansive framing should be more easily retrieved when the learner encounters a new case with a similar structure.

The preponderance of studies in education literature however suggests that appropriate transfer comes with difficulty (Pea, 1987). Bransford and Schwartz (1999) also critique traditional tests of transfer for predominantly testing direct application of one’s previous learning to a new setting or problem with no opportunities for learners to demonstrate their abilities *to learn* to solve new problems. To this end, they call for broadening previous conceptions of transfer by including an emphasis on people’s “preparation for future

learning" (PFL) where the focus shifts to *assessments of people's abilities to learn from new resources*. The PFL perspective suggests that assessments of people's competencies can be improved by involving assessments that provide opportunities for new learning. Such "dynamic assessments" (Campione & Brown, 1990; Schwartz & Martin, 2004; Schwartz, Bransford & Sears, 2005) measure how well students "transfer in" skills to apply to their new learning rather than simply testing how well they "transfer out" of situations to solve problems.

Few studies in the realm of computing education have attended to transfer of learning from visual block-based environments to text-based ones. Dann et al. (2012) mediated transfer of learning from a special version of the Alice visual environment to the text-based Java environment. They contend that by using the exact same example in Alice and Java, their students succeeded with better learning results. More recently, Touretzky et al. (2013) used the idea of presenting contrasting cases and analogous representations of the use of the same computing constructs in a one-week summer camp that had 11 to 17-year-old children transition from Kodu to Alice to Robotics NXT-G in a structured way. By scaffolding instruction to help children see analogies between formalisms in each language, they sought to foster deeper conceptual understanding. For example, their strategies attempted to help children appreciate that "WHEN/DO in Kodu, If/Then in Alice, and SWITCH blocks in NXT-G all function as conditional expressions, even though they look different".

Building on this earlier research, we argue that in computing education, successfully mediating transfer will depend on expansively framed development of deeper conceptual understanding of computational thinking elements that children experience in their introductory computational learning. These include the ability to decompose problems and compose solutions, to understand fundamental notions of flow of algorithmic control that are broadly applicable, and to build practices with an academic vocabulary of the domain that will help students not only communicate computational ideas effectively but also aid in future programming experiences. Additionally, we contend that a concern for transfer of computational thinking experiences will be advanced using PFL assessments that measure readiness to work with more advanced programming environments. Successful PFL would require that students develop not only strong algorithmic thinking skills but also an *understanding of the underlying structures of programs beyond the syntax and surface features of the environment in which children are initially learning programming to more expansive frames in which similarities in deep structures across programming environments are anticipated, recognized and productively used*. Unlike earlier research that has attempted this by employing different programming languages to help students abstract deeper features of constructs, *our work is distinct in that we apply these ideas while using a single programming environment (Scratch), by employing the strategies described below*.

The remainder of the paper describes the features of an introductory CS "mini-course" inspired by the transfer of learning rationale above, and the empirical investigations for studying students' PFL as a result of this curricular intervention in a public middle school classroom.

Methods

This section describes the design-based research around a six-week middle school curriculum titled *'Foundations for Advancing Computational Thinking' (FACT)* designed to promote deeper engagement with foundational CT concepts and assess students through tests of direct application of skills as well as dynamic PFL assessments. The research question guiding this effort is: Does the FACT curriculum promote an understanding of algorithmic concepts that goes deeper than tool-related syntax details, as measured by Preparation for Future Learning (PFL) assessments?

Curriculum Design

Our curriculum focused on core CS concepts that would universally be identified as foundational to any computing experience for middle school. These include structured problem decomposition, and algorithmic notions of flow of control including conditional selection and repetition. The programming unit of the Exploring CS curriculum for high school (<http://www.exploringcs.org/>) inspired the curriculum and use of Scratch.

Instructional Approaches for Promoting Transfer and Deeper Learning

Our approach to teaching for transfer relies on using expansive framing and analogous representations of algorithmic solutions to help learners perceive these in forms more expansive than the constraints of a specific syntactical structure. We predict that guiding students to draw analogies between different formalisms will foster deep and abstract understanding of fundamental concepts of computational thinking.

To this end, the curriculum introduced new computational concepts through a "problem" example that required the use of the concept, demonstrations in Scratch, and an explanation of the concept using new computing vocabulary. Additional examples *using English & pseudo code were used to describe algorithms* involving the use of the concept so that students could see the concept being employed in algorithmic solutions that were represented in ways distinct from Scratch, thus setting them up for a more expansive framing of their learning than learning to program in Scratch alone. Short formative assessments included exercises and quizzes

involving pseudo code and short programming exercises in Scratch, and an end-of-unit activity involving use of the concept in the context of a more substantial programming task (in Scratch).

We used pseudo-code not only to describe and deliberately lay out the sequence involved in organizing the algorithmic steps to accomplish a goal (which has its own benefits), but also to introduce students to *analogical terms and representations of algorithmic solutions distinct from the Scratch environment*. Our reasoning was that this design would bolster familiarity with textual representations of programs, and analogous terms and description of loops and conditional structures that were different from Scratch. For instance, Scratch has only “REPEAT” and “REPEAT UNTIL” blocks for bounded and unbounded iteration. However, using terms like “WHILE” or “FOR” in pseudo-code aim to help students recognize that different terms can be used to describe the idea of repetition of steps (even though there are subtle differences in the ways in which these constructs operate in different programming languages). This approach was taken throughout the course accompanied by suggesting relevance of these terms and ideas in programming experiences in text-based languages such as Java and Python, e.g. *“Even though a loop in other languages like Java or Python will be expressed with terms like While or For, they help to accomplish the same things in an algorithmic process like the Repeat Until loop does in your Scratch program that finds the average test score for a class.”* At various points in each unit, students were also given an opportunity to examine the same algorithm put together in Java or some other language, so even though they were not being taught Java, the analogical instances helped students see the deeper structure of the program, for example, the rather simple “SAY” command in Scratch accomplishes the same goal as the more convoluted “system.out.print”. Space constraints preclude inclusion of figures to demonstrate these analogous snippets of code shown to the students.

Dynamic Assessments for Assessing PFL

The curriculum design included the design of dynamic assessments for assessing PFL that attempt to measure how well students “transferred in” their conceptual understanding of computing constructs to learn from a new resource and apply it to understand code presented in a text-based language. The questions are described along with their goals and student results in the Results & Discussion section below.

The problems were thus preceded by “new learning” in the form of descriptions of how the syntax for the fictitious (Pascal/Java-like) text-based language worked. Two different types of syntax were explained, followed by questions each involving programs coded using the new syntax. For example, the following explanation formed part of the new (Pascal-like) syntax description that preceded Questions 1 and 2.

‘<--’ (left arrow) is used to assign values to variables. *For example:* `n <-- 5` assigns the value 5 to the variable n
 If there are blocks of compound statements (or steps), then the **BEGIN**..**END** construct is used to delimit (or hold together) those statement blocks (like the yellow blocks for REPEAT and IF blocks in Scratch).
FOR and **WHILE** are loop constructs like REPEAT & REPEAT UNTIL

WHILE (*some condition is true*)
BEGIN
 ... (Execute some commands)
END

Figure 1. Sample new syntax specification preceding questions in “dynamic” PFL assessment

In order for students to tackle Questions #3, #4 and #5, they had to use a Java-like syntax preceding the questions that were explained in a similar fashion to the prior Questions 1 and 2.

Study and Data Measures

The curriculum was taught for six weeks in April-May, 2013 in a public school classroom of 25 children from 7th and 8th grade (20 boys and 5 girls, mean age: ~13 years) enrolled in a semester-long *Computers* elective. As an elective class, students were self-selected. This accounted for the low numbers of female students. The class met for 55 minutes each day four times per week. The lead researcher on this effort was also the curriculum developer and teacher for the pilot 6-week FACT course. The regular *Computers* teacher was present in the class at all times.

Data Measures

Beyond a survey to assess prior programming experience, the following data were gathered for assessing students’ learning through the FACT curriculum:

- *Measures of computational learning:* Students were given pre-post tests that measured their understanding of computational concepts especially in the context of Scratch. Questions were borrowed from prior work involving middle school kids and Scratch (Ericson & McKlin, 2012; Zur Bargury, Pârv & Lanzberg, 2013).
- *Preparation for Future Learning Test:* The post-test included a section pertaining to testing for PFL using the five questions as described in the table below. A researcher unrelated to the project graded the test.

Results and Discussion

The results on two out of the five PFL questions are described in the table below that describes the grading procedure as well. In Question #1, 8 out of the 25 respondents misunderstood the question and explained the code instead. Question #2 has been excluded from the table. It used a FOR loop, which has been found to be problematic for novice programmers (Robins, Rountree & Rountree, 2003). Not surprisingly, less than 50% of the students could tackle that question correctly. Question #3 met with the most success. 84% of the respondents got it completely correct or close to correct. Questions #4 & #5 were based on the same snippet of code that was very similar to #3 but with the added complexity of checking for divisibility by 2 and then incrementing one or the other counter variable. The answers were coded with one of 4 scores (3, 2, 1 or 0) as in Question #3. The wording of Question #5 also seemed to have caused some confusion, with some students giving the number of variables that the code used rather than how many numbers are processed by the loop. In both Questions #4 and #5, roughly 65% of the students got the answer correct or mostly correct.

Table 1: Sample PFL questions, their goals, and results describing student responses

PFL Assessment Question	Goal of Question & Results	
#1: When the code below is executed, what is displayed on the computer screen? <pre>PRINT("before loop starts"); num <-- 0; WHILE (num < 6) DO BEGIN num <-- num + 1; PRINT("Loop counter number", num); END PRINT("after loop ends");</pre>	Goal: To test whether students were able to transfer in ideas of (a) sequence and what comes before and after the loop in addition to the things that happen within the loop, (b) looping (using a WHILE loop here), (c) how variable values change with each iteration of the loop, and (d) understanding the loop terminating condition. Results: 32% of respondents misunderstood the question (and explained the code instead). 71% of those who followed the required format of the response (i.e., 68% of the total) got it correct while 19% were off by 1 (<i>num</i> went up to 5 rather than 6). 71% of the total number of students paid attention to the Print commands before and after the loop.	
#3: Describe in plain English what this piece of code is doing. What are FirstCounter & SecondCounter keeping track of? <pre>int TotalCount = 0; int FirstCounter = 0; int SecondCounter = 0; while (TotalCount < 100) { int num; num = InputFromUser(); if (num == 0) { FirstCounter++; } else { SecondCounter++; } TotalCount++; }</pre>	Goal: To test whether students could make sense of (a) a FOR loop and (b) how the variables <i>num</i> and <i>i</i> changed with each loop iteration. The results are shown below (the column to the right indicates number of students)	
	Correct (3 points) <i>Example:</i> "It is asking for 100 user inputs. If the input is 0, then it changes FirstCounter by 1. If it is something else, then it changes SecondCounter by 1."	19
	Mostly Correct (2 points) <i>"it keeps track of 0s for 100 responses."</i>	2
	Mostly Wrong (1 point) <i>Example:</i> variable total set to 0 / variable First set to 0 / variable second set to 0 / repeat total till total <100 / / input variable num / if num is equal to 0 / set first counter else second"	1
	Wrong (0) <i>Example:</i> "The code is asking for numbers that are less than 100. The two numbers."	3

We found the results to be largely encouraging. It is worth noting that: (1) In most cases, students were able to correctly get a sense for the program flow and at a fundamental level understood the concept of looping or conditional execution in the code, even though their responses were not always completely accurate. (2) The students whose responses were consistently right or consistently (completely) wrong mapped closely to the best and worst performers on the Scratch test. This is consistent with earlier literature that contends that skills mastery in the original context is essential for transfer (Kurland, Pea, Clement, & Mawby, 1986). (3) The nature of some of the errors committed in the PFL test were similar to those committed on the Scratch test, suggesting weak initial learning of some concepts. (4) Problems such as the "off-by-1" looping error (Question #5) or issues with the FOR construct (Question #2) are common among older novice programmers at the undergraduate level as well. (5) Most of the PFL question involved loops with variable manipulation, a topic that students found challenging and performed poorly on in the Scratch test as well.

Conclusion and Future Work

As a second iteration of this design-based research, the same curriculum was taught in the same classroom with a new cohort of students comprising 20 boys and 8 girls (mean age ~12.35 years). The assessments were largely unchanged from this study. The wording of Questions #1 and #5 was rectified to improve clarity of what was

being asked. We also added additional PFL assessment questions, and survey questions that elicit student beliefs of future applicability of their learning from this course. Data from this study are currently being analyzed.

Overall, the results of the PFL test were promising, and in answer to the research questions that guided this study, we found that students were able to transfer many algorithmic ideas from the 6-week Scratch curriculum to their new learning, and broadly interpret programs in a text-based language although the mechanics of some constructs were difficult to grasp. Most students evidenced an inherent understanding of the algorithmic flow of control even in a completely new programming context.

This paper suggests a successful approach to using the powerful ideas of Expansive Framing and Preparation for Future Learning (PFL) in introductory CS teaching and learning at the middle school level. Efforts such as the one described in this paper also provide middle school teachers with curricular and pedagogical ideas that promote a deeper engagement with computational thinking concepts, even as they use friendly block-based environments like Scratch. This study also makes an important contribution to the idea of textual but language-agnostic assessments that can be used as PFL assessments after teaching middle school students introductory CS using environments like Scratch and Alice. To conclusively establish the merits of this curricular approach, comparative investigations would be needed with children who are learning programming in Scratch/Alice using other types of curricula.

References

- 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 (pp. 61–101). Washington, DC: American Educational Research Association.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Mind, brain, experience and school (expanded edition)*. Washington, DC: National Academy Press.
- Campione, J.C., & Brown, A.L. (1990). Guided learning and transfer: Implications for approaches to assessment. In N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (pp. 141–172). Hillsdale, NJ: Erlbaum
- Dann, W. P., et al. (2012). Mediated transfer: Alice 3 to Java. In *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*, 141–146.
- Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist*, 47(3), 215–231.
- Ericson, B., & McKlin, T. (2012). Effective and sustainable computing summer camps. *Proceedings of the 43rd ACM technical symposium on CS Education*, 289–294.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95(2), 393–408.
- Grover, S., & Pea, R. (2013). Computational thinking in K–12: a review of the state of the field. *Educational Researcher*, 42(1), 38–43.
- Kurland, D. M., Pea, R. D., Clement, C., & Mawby, R. (1986). A study of the development of programming ability and thinking skills in high school students. *Journal of Educational Computing Research*, 2(4), 429–458.
- Pea, R. D. (1987). Socializing the knowledge transfer problem. *Int'l Journal of Ed. Research*, 11(6), 639–663.
- Pellegrino, J. W. & Hilton, M. L. (Eds.). (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: National Academies Press.
- Robins, A., Rountree, J., & Rountree, N. (2003). Learning and teaching programming: a review and discussion. *Computer Science Education*, 13(2), 137–172.
- Schwartz, D. L., Bransford, J. D. & Sears, D. (2005). Efficiency and innovation in transfer. In J. Mestre. (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1–51). Greenwich, CT, Information Age Publishing.
- Schwartz, D. L. & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, 22(2), 129–184.
- Touretzky, D. S., Marghitu, D., Ludi, S., Bernstein, D., & Ni, L. (2013). Accelerating K-12 computational thinking using scaffolding, staging, and abstraction. In *Proceeding of the 44th ACM technical symposium on computer science education* (pp. 609–614). New York: ACM.
- Wing, J. 2006. Computational Thinking. *Communications of the ACM*. 49(3), 33–36.
- Zur Bargury, I., Pârv, B. & Lanzberg, D. (2013). A nationwide exam as a tool for improving a new curriculum. In *Proceedings of ITiCSE'13*, 267–272. ACM.

Conceptualizing Teacher's Practices in Supporting Students' Mathematical Learning in Computer-Directed Learning Environments

Aaron M. Kessler, University of Pittsburgh, 5500 Wesley W. Posvar Hall, aaronmkessler@gmail.com
Melissa D. Boston, Duquesne University, 600 Forbes Ave, bostonm@duq.edu
Mary Kay Stein, University of Pittsburgh, 828 LRDC, mkstein@pitt.edu

Abstract: In this paper, we suggest a different conceptualization of teachers' practices in supporting students' mathematical learning in a cognitive tutor or computer-directed learning environment. The CT program examined in this study aimed to enhance students' learning of proportional reasoning through an engaging engineering/robotics context. Students of teachers who assumed an active role in engaging students in mathematical conversations away from the cognitive tutor exhibited significant learning gains, whereas students in settings using the cognitive tutor as designed (with limited interaction from the teacher) did not exhibit learning gains. This research has implications for reconsidering the teacher's role in cognitive tutor or computer-directed learning environments.

Purpose and Framework

Cognitive tutors are a specific type of intelligent tutoring system (ITS) that use technologies associated with curriculum sequencing, intelligent solutions analysis and/or problem solving support to assist students' learning (Brusilovsky & Peylo, 2003). Research on the impact of cognitive tutors (CT) on students' learning of mathematics often assumes a teacher-less environment, with the CT directing the teaching and learning of specific production rules associated with students' ability to correctly complete mathematics problems (Koedinger & Corbett, 2006; Anderson, et.al, 1995). Because teaching and learning are designed to occur between the student and CT, little (if any) role is ascribed to the teacher.

Our research provides empirical evidence for an innovation in practice; specifically, a conceptualization of how teachers might implement CT and computer-delivered mathematics instruction to optimize students' learning of mathematics. This work stems from a design-based project, in which our research group's specific role within the larger project was to examine and enhance the mathematics teaching and learning components of a robotics-themed cognitive tutor (CT) program. The program aims to introduce robotics and proportional reasoning to students in 4th-8th grade, as programming the robots involves proportional relationships (e.g., linear distance and turn angle are linear functions of wheel size and motor rotations). Proportional reasoning is a key mathematical idea in middle school (National Council of Teachers of Mathematics, 2006; Lobato, et.al, 2010); hence this program aimed to enhance students' learning of important mathematics through an engaging engineering/robotics context.

As a design study, our work proceeded in three phases. In Phase 1, we reviewed the tutor materials and determined that the vast majority of tasks (97% of all tutor tasks) engaged students in only lower levels of mathematical thinking; specifically, procedures without connections to meanings, concepts, or understanding (Stein, Grover, & Henningsen, 1996). In mathematics education research, procedural knowledge of this type is considered "low cognitive demand" and has been empirically associated with low student achievement in mathematics (e.g., Stein & Lane, 1996; Hiebert & Stigler, 2004). Conversely, declarative knowledge is consistent with high cognitive demand, and research suggests that cognitively challenging tasks, with high-level demands maintained during implementation, are associated with positive student learning (Boaler & Staples, 2008; Stein & Lane, 1996; Tarr et al., 2008).

The CT tasks initially observed by our research team in Phase 1 required students to perform procedures (e.g., multiply or divide to complete a data table) without invoking a conceptual understanding of proportional relationships. In fact, only 2 of 62 original CT tasks (3%) provided students opportunities for higher-level thinking and reasoning. Furthermore, at early implementation sites (prior to our involvement), no learning gains appeared on a test of proportional reasoning (Weaver & Junker, 2004). Driven by our hypothesis that higher-cognitive-demand tasks would increase students' learning, the remainder of Phase 1 consisted of the development and integration of a set of cognitively demanding tasks into the CT system, tasks that engaged students in thinking about the role of unit rates and scale factors in proportional relationships. CT tasks were revised to have high-level demands, with new materials containing 16 high-demand tasks (16/48; 33%) that could engage students in developing a conceptual understanding of proportional relationships and strategies (e.g., unit rates and scale factors). We analyzed CT mathematics instructional materials using existing frameworks developed to examine the cognitive demand of mathematical tasks (i.e., the Task Analysis Guide (Stein, Smith, Henningsen, & Silver, 2009) and shifts in their implementation through lesson phases (i.e., the

Mathematics Task Framework [Stein, Grover, & Henningsen, 1996]. We also drew from research delineating students' development of proportional reasoning (Lamon, 2007).

Phase 2 primarily consisted of observations of the implementation of the new, higher-cognitive demand materials using an observation instrument (the Interaction Tracker) that was based on a conceptualization of teaching and learning known as the "instructional triangle" (Cohen, Raudenbush, & Ball, 2003; Stein & Kim, 2012). At one implementation site, researchers noted that instructors engaged students in discussions *away from* the CT screen to help students understand the underlying mathematical ideas (e.g., unit rate and scale factor). These interactions between teachers and students around the mathematics seemed particularly fruitful for impacting students' learning, and distinctly different from interactions noted in other settings. Seeing these types of interactions occurring in only one setting caused the team to consider the possible role of the instructors in the CT program. We hypothesized that different teacher-student-computer interaction styles would generate differences in students' learning, and that different interaction styles may be associated with teachers' pedagogical content knowledge (PCK) of proportional reasoning.

In Phase 3, based on this second hypothesis, we developed materials that introduced teacher-mediation into the system. Specifically, we inserted materials that were designed to guide teacher-student interactions at points where students had just finished learning or dealing with difficult concepts. Specifically, facilitation materials were created that encouraged teachers to draw students' attention back to the declarative knowledge pieces of the program. These materials were meant to foster interactions between students and teacher away from the CT that were still grounded in the unique robot context and that provided opportunities to talk more deeply around the mathematics. In this paper, we present research testing our hypotheses at phases 2 and 3:

1. Did implementation sites using the high-cognitive-demand CT tasks exhibit increases in students' learning?
2. Are different types of teacher-student-computer interactions associated with teachers' level of PCK?
3. Did implementation sites using the high cognitive demand CT tasks and embedded opportunities for interaction (about the mathematics) exhibit increases in students' learning?

Method

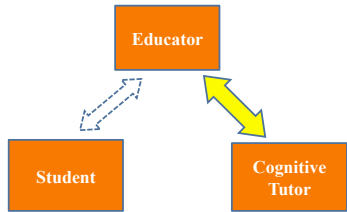
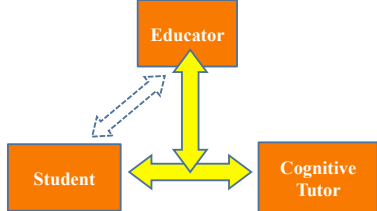
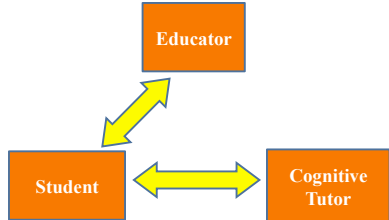
In Phase 2, the project worked with 3 implementation sites representing 111 students, observing the impact of the revised CT materials from phase 1 on learning outcomes and students' interactions with the system. During these implementations, the research team made regular observations and assessed student knowledge by administering a pre-post proportional reasoning written assessment using an equivalent forms design (e.g., Weaver & Junker, 2004). These student pre- and post-test scores were compared using paired-values t-tests.

Teachers in the three Phase 2 implementation sites were observed as they implemented the CT program, and an "Interaction Tracker" (informed by Stein and Kim, 2012) that was developed to represent the observed interactions. The Interaction Tracker was then tested in three additional settings. We compared interaction patterns of teachers with different levels of PCK in proportional reasoning, based on a project-developed assessment of teachers' use of proportional reasoning strategies to solve challenging problems and ways in which they indicated they would support students' use of proportional reasoning strategies. We designed, piloted, and checked the reliability of this assessment. Specifically, we double-coded 7 teacher-PCK tests (3 pilot teachers and 4 project teachers) with 89% exact-point reliability. Teacher-PCK data exists for 8 of 10 facilitators, representing 5 of 6 implementation sites (one facilitator was an expert from the project team). We also examined whether different interaction patterns supported students' engagement in high-cognitive-demand tasks.

During these Phase 2 observations we identified four distinct types of interactions between the teacher, students, and CT (see Table 1). Educators ($n = 10$) at each of the six implementation sites, through phase 2 and phase 3, were observed by at least one member of the team, during at least two sessions.

Table 1. Interaction patterns

Interaction Pattern	Description	Instructional Triangle Diagram
Cognitive Tutor (CT) Environment as Designed	Students interact with the CT with minimal or no interaction between the teacher and student or teacher and CT. The CT provides all learning opportunities; all math activity occurs between the CT and the student (indicated by the double-headed yellow arrow)	

Interaction Pattern	Description	Instructional Triangle Diagram
Educator Taking Over the CT Environment	The educator (instead of the student) interacts directly with the CT, often physically taking over the computer (e.g., reads CT screen to the student, rephrases questions on the CT, enters responses into the computer). Most or all activity with the CT is filtered through the teacher. Teacher-student interactions are teacher-directed, reducing students' engagement with the mathematics to brief responses.	
Educator Facilitating the CT Environment	The educator addresses procedural or technical aspects of the students' interaction with the CT around mathematics (e.g., calling attention to the need to round to 2 decimal places, suggesting use of the hint button). The mathematical activity remains between the student and the CT, with the educator facilitating this activity. The educator does not address aspects of the student's specific ways of thinking or components of the underlying mathematics not present on the CT screen.	
Educator Facilitating the Mathematics	The educator interacts with the student about the underlying mathematics or about the students' specific mathematical thinking. The interaction is no longer directly about what is happening on the screen. Meaningful mathematical activity occurs between the teacher and student, and then the student uses the mathematics to engage with the CT.	

To examine the impact of an increased role for teachers in the CT environment, we re-designed the materials to include embedded opportunities for teacher-student interactions around the mathematics in Phase 3 of the program. In the original materials, students progressed through a series of computer-delivered, cognitive tutor modules at their own pace. The new structure contained three whole-group discussion points (D1, D2, and D3 in Figure 1) and facilitation materials designed to increase students' engagement with the robot engineering context (D1), the mathematical ideas (D2) (e.g., proportional reasoning, scale factor, and unit rate), and to use the mathematics to accomplish the goals of the robot-engineering context (D3).

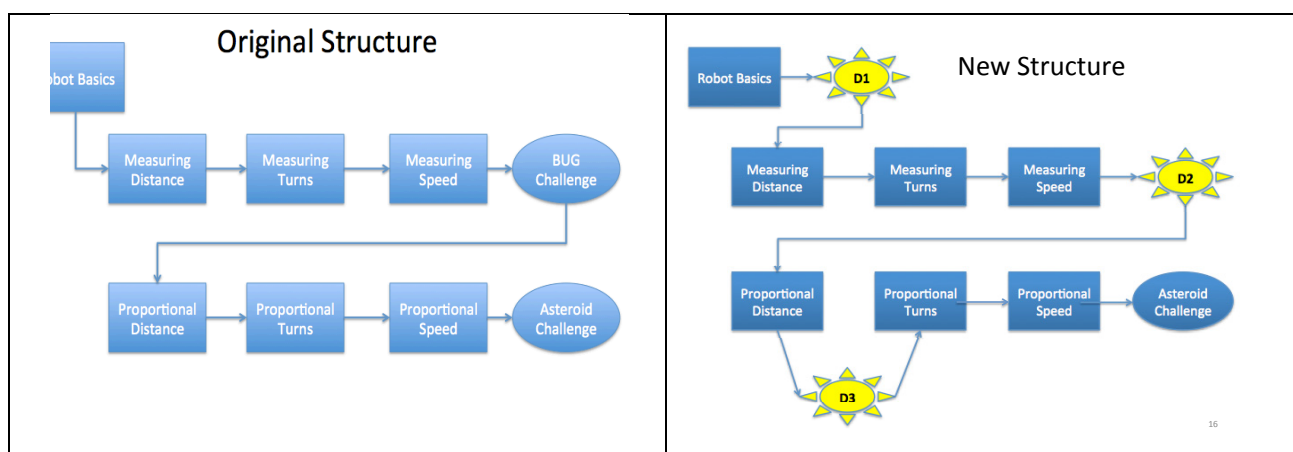


Figure 1. Original and revised structure

In phase 3, a member of the design team piloted the new teacher-mediated materials in an afterschool program ($n_4 = 12$ students) in grades 5-7. Following this pilot, two teachers from different districts also implemented the new materials ($n_5 = 12$ students; $n_6 = 16$ students). Students' work on the project lasted approximately 3 weeks at each site and included students in grades 7-8. Along with the pre/post assessments of students' knowledge, observations were carried out at both sites using the Interaction Tracker.

Findings

Research Question 1. In examining the three implementation sites in Phase 2 that used the high-cognitive-demand CT materials ($n = 111$ students), students showed no learning gains on the proportional reasoning written assessment ($p = .48$). Hence, simply incorporating the high-cognitive-demand tasks did not generate the anticipated increases in students' learning.

Research Question 2. Teachers' interaction patterns were examined, and for each teacher, researchers identified a clear main interaction pattern and secondary interaction pattern(s). The 10 teachers were then grouped by their scores on the PCK tests, and a relationship between PCK and interaction pattern appeared to exist (see Table 2).

Research Question 3. In piloting the Teacher-Mediated materials (Phase 3), the instructor in one implementation site was a member of the research team (Site 4; $n_4 = 12$). In the two school-based implementation sites ($n_5 = 12$, $n_6 = 16$), students' pre/post assessments increased (Site 5, 56% to 57%; Site 6, 59% to 64%). When comparing the pre/post scores across all three sites that used the new teacher-mediated materials, including the pilot, we see a significant increase from $M = .596$ to $M = .633$, ($n = 41$, $p = .04$).

Looking across Research Questions 2 and 3 to further explicate the connection between the observed interaction patterns and students' learning, the research team is more deeply examining Group 1 and Group 2 teachers and implementation sites (see Table 2). Preliminary results support the connection between the "facilitating the mathematics" pattern and increased student-learning.

Table 2. Teachers' interaction patterns and pedagogical content knowledge scores

Teacher	Certification	PCK Score	Interaction Patterns
Group 1: <i>Facilitating the Mathematics</i> as the Primary Instruction Style			
Site 4	Expert	High PCK	<i>Facilitating the Mathematics</i> Facilitating CT Environment Designed CT environment
Site 2 –Teacher	Math	83%	<i>Facilitating the Mathematics</i> Designed CT environment Facilitating CT Environment
Site 2 – Student Teacher	Math Pre-Service Teacher	76%	<i>Facilitating the Mathematics</i>
Group 2: <i>Facilitating the Mathematics</i> as a Secondary Instruction Style			
Site 3 – T2	Math	58%	Facilitating CT Environment <i>Facilitating the Mathematics</i> Designed CT environment
Site 3 – T1	Science	56%	Facilitating CT Environment <i>Facilitating the Mathematics</i>
Site 6	Technology/Industrial Arts	56%	Facilitating CT Environment <i>Facilitating the Mathematics</i>
Group 3: <i>Facilitating the Mathematics</i> Not Present			
Site 1 – T2	Elementary	53%	Facilitating CT Environment Taking over the Environment
Site 1 – T1	Elementary	50%	Facilitating CT Environment
Site 5	Technology/Industrial Arts	22%	Designed CT environment
Site 2 - LST	Learning Support	--	Taking over the Environment

Implications

Our work echoes a consistent finding in mathematics education research: cognitively challenging tasks are a necessary but not sufficient condition for promoting students' learning of mathematics, and implementation appears to be key. In our work, revising the CT tasks did not enhance students' learning, but coupling such tasks with implementation that maintains high cognitive demands (e.g., as enacted in the "Facilitating the Mathematics" interaction pattern and as designed into the teacher-mediated materials) was essential for increasing students' proportional reasoning strategies and understanding of proportional relationships. This represents a shift in how the practice of using CT's in the classroom has been looked at in the past.

The role of teachers' pedagogical content knowledge suggests the need for professional development around the teaching of important mathematical ideas. Crafting the types of interactions that facilitate the mathematics in CT and computer-delivered learning environments requires a deep understanding of the underlying mathematical ideas that is often not included as part of current CT resources. Building this capacity in teachers who will be implementing this CT program shows promise in laying a foundation for quality mathematics-based interactions with students.

Finally, across sites and contextual factors, the addition of teacher mediation to the system had a positive impact on students' ability to learn proportional reasoning. This finding has general implications for reconsidering the "designed" role of the teacher in CT or computer-delivered instructional settings. Teachers acting to "facilitate the mathematics" may promote greater students' learning than if the CT or computer is assumed to direct the teaching and learning in a "teacher-less" environment.

References

- 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.
- Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *The Teachers College Record*, 110(3), 608-645.
- Brusilovsky, P., & Peylo, C. (2003). Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education*, 13(2), 159-172.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119-142.
- Hiebert, J., & Stigler, J. W. (2004). Improving Mathematics Teaching. *Educational leadership*, 61(5), 12.
- Henningsen, M. & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28, 524-549.
- Koedinger, K. R., & Corbett, A. T. (2006). Cognitive tutors: Technology bringing learning science to the classroom. *The Cambridge handbook of the learning sciences*, 61-78.
- Lamon, S. J. (2007). Rational numbers and proportional reasoning: Toward a theoretical framework for research. *Second handbook of research on mathematics teaching and learning*, 1, 629-667.
- Lobato, J. & Ellis, A. (2010). *Developing Essential Understanding of Ratios, Proportions & Proportional Reasoning: Grades 6-8*. Reston, VA: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics (NCTM). (2006). Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence. Reston, VA: NCTM.
- Stein, M. K., Grover, B. W., & 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(2), 455-488.
- Stein, M. K., & Kim, G. (2012). The role of mathematics curriculum materials in large-scale urban reform: An analysis of demands and opportunities for teacher learning. In Remillard, J. T., Herbel-Eisenmann B. T., & Lloyd, G. M. (Eds), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction*, (pp. 37-55). New York, NY: Routledge Taylor & Francis Group.
- Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, 2(1), 50-80.
- Tarr, J. E., Reys, R. E., Reys, B. J., Chávez, Ó., Shih, J., & Osterlind, S. J. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education*, 247-280.
- Weaver, R., & Junker, B. W. (2004). Model specification for cognitive assessment of proportional reasoning.

Teaching Struggling Middle School Readers to Comprehend Informational Text

Donna Caccamise, Angela Friend, Christine Groneman, Megan Littrell-Baez, and Eileen Kintsch
University of Colorado, Boulder, Institute of Cognitive Science, 1777 Exposition Dr., Boulder, Colorado 80301
Email: donna.caccamise@colorado.edu, angela.friend@colorado.edu, christine.groneman@colorado.edu, megan.littrellbaez@colorado.edu, eileen.kintsch@colorado.edu

Abstract: This project involves developing a reading comprehension curriculum based on the Construction-Integration model (Kintsch, 1998) and implementing a pilot efficacy study in middle-school classrooms. The curriculum (BRAVO) explicitly defines the cognitive processes involved in skilled reading, teaching students how to establish local and global text coherence and use background knowledge to create a mental model – the prerequisite to deep, meaningful learning. A crucial and unique component is the use of sequenced texts related to a single over-arching topic, enabling students to use their expanding prior knowledge while reading complex expository texts. BRAVO, consistent with Common Core State Standards, bridges the divide between grade-school narratives and secondary-level textbooks.

Introduction

This report describes an ongoing project to develop a reading comprehension curriculum targeted to middle school children and aligned with Common Core State Standards (CCSS) for comprehending informational texts. Drawing on current cognitive theory and research on comprehension and the limited results of classroom- and lab-based studies, the **Boulder Reading Intervention (BRAVO)** offers a detailed, content-based literacy curriculum. BRAVO is being employed in classrooms as part of the language arts curriculum in participating schools. Most of the instructional components are familiar to educators; what is new here is that the components are organized and motivated by an explicit theory, the Construction-Integration model of comprehension (Kintsch, 1988; 1998). What the theory offers, thus, is a narrative that makes sense of literacy instruction. Our goal here is to figure out just how the various components of the theory need to be integrated and presented so that future teachers can use them creatively and effectively in their work, and not just vainly attempt to follow a script.

Significance of the Project

The BRAVO curriculum addresses an important and widely recognized problem in education, namely, that: many children in America read well below national standards. Based on the CI Model, BRAVO helps students understand difficult content-rich, informational text, teaching them to detect and remedy comprehension failures, overcome misconceptions, understand the text meaning and build knowledge. Using middle school science content, students learn how to form coherent and cohesive understandings of the ideas in a text, while integrating the new content with prior knowledge to build rich and flexible knowledge of the topic. Strategy instruction in BRAVO takes place as students work through a graduated series of texts of increasing complexity all related to a single, important theme. They learn appropriate mental processes by means of concrete, targeted activities to help them unlock the meaning of new text while drawing on their growing repository of knowledge about the topic. They work with the new content in various ways so that it becomes integrated in their existing store. This principled approach with its joint focus on comprehension and knowledge building, is what differentiates BRAVO from other reading programs.

Background

The importance of improving reading achievement is not in dispute (e.g., Biancarosa & Snow, 2004). It is critical to address this problem if we are to prevent students from dropping out of high school because they are unable to read well enough to learn from the complex, content-rich materials encountered in the secondary curriculum (Caccamise & Snyder, 2009; Carnegie Council on Advancing Adolescent Literacy (CCAAL), 2010). When students begin upper elementary school, they are expected to read content-area texts (often in the form of a textbook); however, they are rarely provided with instruction on *how* to read informational texts. The result is the well-known fourth-grade slump. Beyond Grade 4, students with inadequate reading skills have difficulty catching up (e.g., Francis & Stuebing, 1996). Even many college-bound students lack the more advanced literacy skills needed to understand content area textbooks and to learn from what they read (Allen & Sconing, 2005).

The problem that many poor readers share is not their inability to read the words on a page but rather their inability to comprehend what they read (Biancarosa & Snow, 2004). Poor readers struggle with issues related to inadequate knowledge: knowledge of content of course, but also knowledge of syntax, of syntactic

markers for logical relations and coherence, and knowledge of the diverse text structures used in expository writing (Hirsch, 2006; Torgesen et al., 2007). Remediation of adolescent literacy difficulties, therefore, is not simply a matter of revisiting elementary school instruction or using easier materials. Instead, it is a matter of providing these readers with explicit, concrete tools for getting the meaning of what they are reading.

Faulty literacy skills are only part of the problem. Building the requisite domain knowledge is also essential (Caccamise & Snyder, 2005; 2009; Kintsch, 1998; Kintsch & Kintsch, 2005). As Torgesen et al. (2007, p. 63) point out, there is “compelling evidence that as students improve their knowledge in any specific area, their ability to comprehend text in that area improves. Consequently, any effort to improve adolescent literacy in the long term must focus on the efficacy of teaching essential content within and across grade levels as an important reform goal”. BRAVO, addresses both the literacy and the knowledge problem together by embedding the instruction of explicit reading comprehension skills within content-area learning.

Attributes and Shortcomings of Existing Practice

A widely used approach to comprehension instruction consists in teaching middle school students comprehension strategies like the ones expert readers use, to be applied in situations where comprehension breaks down. However, this approach also presumes that all students come adequately prepared to form their own interpretations of informational text, that they are able to detect a failure to comprehend, and that they know which strategy or strategies to apply to achieve a more complete and accurate understanding. Expository texts are especially problematic, not only because of the unfamiliar subject matter, but due to the higher density of ideas, lack of coherence and more complicated references employed in such texts. Hence, getting the meaning from assigned readings in content area classes often poses an insurmountable obstacle for these readers. Students who struggle at the text level have no basis from which to determine which strategy or strategies might help them improve or expand their comprehension.

Therefore, we argue that strategy instruction alone cannot meet the needs of readers with comprehension problems, who need to learn how to connect the words, phrases and sentences and larger text segments into coherent, comprehensible ideas — a perspective that is lacking in most comprehension strategy interventions. Moreover, curriculum materials that typically emphasize broad topic coverage are not conducive to building coherent subject matter knowledge that is well elaborated and integrated with personal knowledge. This kind of deep understanding, according to Kintsch (e.g., 1998; Kintsch & Kintsch, 2005), is crucial for creating lasting, usable knowledge.

A Better Pedagogical Model

According to Biancarosa and Snow (2004; also Caccamise & Snyder, 2009; CCAAL, 2010), direct, explicit comprehension instruction, guided by a principled understanding of the learning process and that is *focused on content* is necessary to improve middle and high school literacy. The BRAVO curriculum provides explicit comprehension instruction within a content-area context. The intervention is derived from Kintsch's (1988, 1998) Construction-Integration (CI) Model of Text Comprehension, which explicitly defines the cognitive processes involved in successful comprehension and learning. Thus, the CI model provides the blueprint for designing instruction, materials, and assessments in BRAVO.

Theoretical and Empirical Rationale: The CI Model of Text Comprehension

The CI model (Kintsch, 1988; 1998), as well as other recent cognitive models (e.g., Graesser, Singer, & Trabasso, 1994; Just & Carpenter, 1987; van den Broek, Young, Tzeng, & Linderholm, 1999) mark a change in comprehension research from an earlier focus on memory representation (i.e., what readers remember *after* they read) to a specification of the mental processes involved in reading (i.e., what readers do *while* they read). This theoretical direction since the 1990s has sought to understand how readers make sense of text word-by-word, line-by-line, and moment-by-moment, developing an understanding of the content as they go along, an understanding that is guided by top-level goals, beliefs, personal skills, and knowledge. Readers' understanding changes and develops as they read, based on the words on the page and on the meaning they make in their minds from the text and from their own knowledge.

The CI Model describes the cognitive processes of comprehension as the interaction of the *textbase* (*ideas in the text*) and the *situation model* (*integration with own knowledge*). The reader develops a textbase representation from the propositions, or idea units, explicitly stated in the text, inferring information as needed to form a coherent understanding of the text content. The reader builds a situation model representation by integrating the textbase with his or her own background knowledge that is relevant to the text or that is required for comprehension (Kintsch, 1988; 1998). Developing a textbase and building a situation model are dynamic interactive processes that happen simultaneously. Successful comprehension at both levels of processing depend on the reader's background knowledge: knowledge about the topic, about language use, about expository text structures, as well as knowledge about how strategic processes interact to support the construction of a coherent, accurate and personalized memory representation.

Readers who attend only to the words, phrases and sentences will develop a shallow understanding that does not last. Building a situation model is essential for deep comprehension that is sustained over time. A situation model enables the reader to find a place for new knowledge in the brain by linking the new information to existing pieces of knowledge. These links in the memory network may later serve as cues for retrieving the information in future applications (Ericsson & Kintsch, 1995; Kintsch & Kintsch, 2005).

Many factors prevent inexperienced readers from comprehending text deeply, i.e., from developing a coherent and well-integrated memory representation; some are inherent in the reader, others are features of the texts themselves. For example, readers may lack sufficient topic knowledge to realize what particular non-text knowledge or facts may be relevant to understanding the text at hand (e.g., McNamara & Kintsch, 1996; Voss & Silfies, 1996). In addition to inadequate conceptual background and vocabulary, readers may have difficulty parsing the more complex and less familiar syntactic forms and organizational structures of expository prose (e.g., Meyer, 1975). Difficult, inhospitable text characterizes many content area readings and likewise impedes understanding. Such text factors include the following: ambiguous or indirect references; absence of information that establishes an appropriate context for the content and that provides a link to existing background knowledge; lack of clearly signaled connections between ideas or events; irrelevant details; and sentences with a high density of concepts and propositions (Best, Rowe, Ozuru, & McNamara, 2005; Britton & Gülgöz, 1991; McKeown, Beck, Sinatra, & Loxterman, 1992; McNamara, Kintsch, Songer, & Kintsch, 1996). Textbook writing often avoids linking ideas with connectives and subordinate relationships in the interest of making shorter sentences that conform to readability formulas (Kintsch & Kintsch, 2005). Science textbooks especially feature brief paragraphs of short, choppy sentences, listing facts and definitions with little elaboration or explanation of how they are related. Lacking knowledge of the basic coherence-making processes, struggling readers are unable to make sense of such materials. Short sentences may lure readers into a false sense of understanding, such that they often do not even realize when their understanding is incomplete or how to bootstrap their understanding with appropriate problem-solving strategies (e.g., McNamara, 2007). However, even readers with adequate skills and knowledge, often fail to engage in the effortful, inferential processes needed for deep understanding (Kintsch & Kintsch, 2005; McNamara & Kintsch, 1996). Thus passive, unengaged reading is also a problem for many students.

The BRAVO Intervention

Our project explores the notion that direct, explicit teaching and practicing of the cognitive processes of comprehension *within the context of content learning* can, in fact, enhance cognitive engagement during reading and improve comprehension of text. In accordance with the CI model's emphasis on knowledge building as fundamental to successful comprehension, we believe that an instructional model that teaches comprehension processes simultaneously with content knowledge will be more effective for readers with poor comprehension. McKeown, Beck, and Blake (2009) offer supporting evidence for this argument in their two-year study of 5th graders comparing strategy instruction with and without content focused instruction. Consistent, though small effects favoring the content group were found on recall and on one transfer task. McKeown et al. (2009) argue that although it is important to teach students about reading strategies, instruction should focus primarily on the meaning of the text: how to construct a coherent understanding of the content.

This is precisely the goal of BRAVO. In BRAVO knowledge building at both the textbase and situation model levels is supported by working through a series of carefully sequenced texts and learning activities pertaining to a single overarching topic. Rather than end-of-chapter questions and strategy instruction applied to different text genres and topics, students learn how to deeply comprehend and how to monitor and remediate comprehension problems in the context of learning thematically related content. During the 8-week course of instruction students expand their knowledge of the topic, using texts that vary conceptually, structurally, and linguistically. They become familiar with different mid-level organizing structures used in informational text, such as sequence, cause-effect, problem-solution, and compare and contrast.

The curriculum is divided into four modules. In expert reading, comprehension occurs at different levels of processing more or less simultaneously and in an integrative fashion. However, for ease of presentation and learning the first three modules in this curriculum deal with comprehension problems that occur at each level in turn, organized from easier to more difficult. The goal throughout is coherence building: within and between sentences in Module 1, between paragraphs and larger segments of text in Module 2, finally dealing with inferential complexity in Module 3. In Module 4 students review these comprehension skills, while integrating the content across all the chapters to form a situation model of the overall theme of ecology.

The texts that students read were manipulated so that students learn to recognize and repair particular problems at each processing level. Thus the three texts for each module are sequenced to first illustrate the particular comprehension problem, followed by "problem" texts that feature coherence gaps and diminishing textual support. In this way, students learn to treat the comprehension of complex text as a problem-solving activity for which they learn the appropriate remedial strategies. The text manipulations reflect real "problems" that students encounter in "authentic" texts such as the lack of explicit signal words to indicate topic transition,

or the use of pronouns and other co-referents to refer back to the subject from an earlier line/section in the text. The texts in Module 3 were manipulated to elicit deeper inferencing on the part of the reader with each subsequent text.

An overview of the instructional curriculum is as follows:

In Module 1 students learn to identify various signals to local, inter-sentential connections, for example, by identifying the referents of pronouns and synonymous terms, identifying and filling in missing connections (e.g., *then*, *and so*, *for example*, *because*, *also*, etc.), and clarifying unfamiliar vocabulary. In so doing they become familiar with more complex sentence structures and learn to form gist statements at the paragraph level as they re-read the text chapter.

Module 2 focuses on identifying textual signals to the underlying structure, (e.g., headings, subheadings, boldface terms, and global level connectives (e.g., *in sum*, *an example of X*, *as a result*, *in contrast*, etc.) that signal different kinds of mid-level and overall structures (e.g., compare-contrast, steps in a process, definition-example, problem-solution); they organize the content into graphic organizers that make the structure explicit; they write summaries that subsume several paragraphs and the entire text chapter. As they proceed through successive chapters they also select key concepts to include in a concept map of the overall ecology theme.

Module 3 addresses the more effortful, inferential kinds of processes that are essential to constructing deep and lasting, personalized memory for the material. Students learn about inferring beyond the text meaning, how to recognize when inferences need to be made and which inferences are appropriate. Appropriate inferences are ones that fill out and elaborate the author's meaning, for example, relevant implications, predictions, self questioning and explaining, critical evaluations, stating how current text relates to previous ones, and the like.

Each sequenced text in Modules 1-3 introduces science concepts while also referring back to concepts presented in previous chapters, helping students reinforce concepts and relations among concepts both within and across texts. The single overview text in Module 4 sets the stage for integration and review, in which students focus on semantically linking ecology concepts from the previous texts, while engaging in all of the coherence building and inferential tools used by expert readers. An important activity is labeling the relations between concepts in their concept maps, thereby creating an overt situation model of their understanding. Finally, students consolidate their knowledge by writing an opinion essay addressing a broad problem in the field of ecology.

Interim Outcomes

In our most recently completed Colorado implementation (AY2012-2013) to test the feasibility of BRAVO as a model for how to improve comprehension instruction for middle school struggling readers and learners, middle school teachers were provided with professional development and ongoing mentoring as they implemented the 8-week program. Before starting BRAVO and right after completing the program, students read one of two counterbalanced, unfamiliar expository texts followed by a written recall task. A paired-samples t-test was conducted to compare recall performance collected before and after implementing BRAVO in the classroom. Results show a significant improvement in the scores for the posttest ($M = 7.90$, $SD = 3.22$) compared to the pretest ($M = 5.73$, $SD = 2.54$), $t(44) = 4.418$, $p < .001$. The effect size for this difference is .37. When compared against the annual reading gain effect size of .24 for an average middle school student across seven nationally normed tests, it suggests that BRAVO improved comprehension skills. Furthermore, as the students in this sample were in remedial reading support classes, the improvement in this population is quite promising. A larger, controlled efficacy study is underway this year in both general classrooms and reading resource classrooms.

Relevance to the Conference Theme

Comprehension instruction in BRAVO focuses directly and explicitly on the mental processes that serve the cumulative building of content knowledge during reading. The conference theme, "Learning and Becoming in Practice" is realized here by developing students' content knowledge as they learn about and practice expert comprehension skills. Thus, at the conclusion of a lesson sequence we anticipate that students who have exhibited poor understanding of content area reading assignments will be empowered to read complex, grade-level informative texts, texts that they were previously unable to fathom. In the pilot study currently underway we expect to demonstrate (a) the feasibility of comprehension instruction based on theoretically specified cognitive processes within a content oriented framework; and (b) the fidelity of implementing the intervention in a typical educational setting.

References

- Allen, J., & Sconing, J. (2005). Using ACT Assessment Scores to Set Benchmarks for College Readiness. ACT Research Report Series, 2005-3. *American College Testing (ACT), Inc.* Retrieved from http://209.235.214.158/research/researchers/reports/pdf/ACT_RR2005-3.pdf.
- Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorder, 25*, 65-83.
- Biancarosa, G., & Snow, C. (2004). *Reading next: A vision for action and research in middle and high school literacy: A report from Carnegie Corporation of New York*. New York, NY: Alliance for Excellent Education.
- Britton B. K., & Gülgöz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology, 21*, 333-356.
- Caccamise, D., & Snyder, L. (2009). Comprehension instruction in the 21st century. *Perspectives on Language and Literacy, 35*, 23-26.
- Caccamise, D., & Snyder, L. (2005). Constructive models of reading comprehension: The situation model and related constructs. *Topics in Language Disorders, 25*, 5-20.
- Carnegie Council on Advancing Adolescent Literacy (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success*. New York, NY: Carnegie Corporation of New York.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211-245.
- Francis, D.J., & Stuebing, K.K. (1996). Development lag versus deficit models of reading. *Journal of Educational Psychology, 88*, 3-17.
- Graesser, A. C., Singer, M., & Trabasso T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371-395.
- Hirsch, E. D., Jr. (2006). *The knowledge deficit*. New York, NY: Houghton Mifflin, Co.
- Just, M.A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston, MA: Allyn & Bacon.
- Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*(2), 163.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge; New York, NY, USA: Cambridge University Press.
- Kintsch, W., & Kintsch, E. (2005). Comprehension. In S. G. Paris & D. S. Stahl (Eds.), *Children's reading comprehension and assessment*. (pp. 71-92). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- McKeown, M. G., Beck I. L., & Blake, R. G. K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly, 44*, 218-253.
- McKeown, M. G., Beck, I. L., Sinatra, G. M., & Loxterman, J. A. (1992). The contribution of prior knowledge and coherent text to comprehension. *Reading Research Quarterly, 27*, 79-93.
- McNamara, D. S. (Ed.) (2007). *Reading comprehension strategies: Theories, interventions, and technologies*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes, 22*, 247-287.
- McNamara, D. S., Kintsch, E., Songer, N., & 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), 1-43.
- Meyer, B. J. F. (1975). Identification of the structure of prose and its implications for the study of reading and memory. *Journal of Reading Behavior, 7*, 7-47.
- Torgesen, J. K., Houston, D. D., Rissman, L. M., Decker, S. M., Roberts, G., Vaughn, S., Wexler, J., Francis, D. J., Rivera, M. O., & Lesaux, N. (2007). *Academic literacy instruction for adolescents: A guidance document from the Center on Instruction*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71-98). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Voss, J. F., & Silfies L. N. (1996). Learning from history texts: The interaction of knowledge and comprehension skills with text structure. *Cognition and Instruction, 14*, 45-68.

Acknowledgements

This project has been funded by the Department of Education, Institute of Education Sciences (IES) grant #: R305A110467, Donna Caccamise, Principal Investigator. Fostering Comprehension & Knowledge-Building in Middle School Struggling Readers

Becoming a Computer Scientist: Early Results of a Near-Peer and Social Justice Program with Latino/a Children

Jill Denner and Jacob Martinez, Education, Training, Research, jilld@etr.org, jacobm@etr.org
Heather Thiry, Golden Evaluation and Policy Research, heather.thiry@goldeneval.com

Abstract: This report describes the results of a pilot study of CSteach, an afterschool program designed to get Latino/a students on the pathway to becoming a computer scientist. The hypothesis was that a social justice perspective and the use of near-peer teachers are motivating for elementary school children with little prior exposure to computer science. Using a mixed-methods approach, the study employed pre-post surveys, focus group interviews, observations, and an assessment rubric for student projects. The data show measurable increases in students' computer science identities, capabilities, and networks of support, and provide information that was used to strengthen the next phase of the program. High school-aged near-peer teachers also benefited from the program. They described growth in their leadership and communication skills and recognized their position as role models for the younger students. The paper has implications for designing culturally relevant learning environments.

Deep learning of the kind valued in the learning sciences requires students to be motivated to participate and persist in the face of challenges or uncertainty. This is particularly important in the field of computer science (CS) for two reasons. First, CS is not required or available in most US schools, and stereotypes based on gender and race/ethnicity prevent many students from taking or finishing these classes (Margolis, 2008). Second, learning to think computationally requires cycles of design, test, and revise (Lee et al., 2011), and beliefs about a lack of innate ability prevent many students from persisting when their initial program does not work. This study focuses on Latinos, who are the fastest growing ethnic minority group in the US but are underrepresented in CS. They are 16% of AP test takers, but only 1% of the AP Computer Science test takers; those that took it scored far below their peers (College Board, 2011). In order to increase the numbers of Latinos in CS, we need to create learning environments that excite them about becoming a computer scientist starting in elementary school and motivate them to persist in deep CS learning.

In this paper, we describe research on CSteach, an afterschool program that aims to move Latino/a students down the path to becoming a computer scientist by fostering capacities, identities, and relationships starting in 5th grade. Our Design Experiment involves an iterative cycle of implementation, data collection, and revision that is well-suited to developing programs that aim to avoid a deficit perspective when promoting learning experiences for marginalized populations (Collins, Joseph, & Bielaczyc, 2004). The design of the CSteach program was based on the understanding that CS, like all knowledge, is situated within a cultural context (Nasir & Hand, 2006) and that educational approaches should affirm and draw on intellectual resources in a given culture or community (Bang & Medin, 2010). CSteach draws on research in mathematics that shows that creating learning experiences that are culturally relevant and connect with social justice issues can engage students in learning in a way that is deep, meaningful, and that contributes to the development of a positive identity (Leonard et al., 2010). CSteach is the first known effort to integrate CS and social justice in elementary schools, and it builds off examples of “computing for the social good” in college (Goldweber et al., 2011).

Our hypothesis is that by integrating social justice into a CS curriculum, it would make it more relevant and increase students' motivation to pursue and persist, particularly when faced with challenging CS concepts in an after school setting. Key determinants of motivation and cognitive engagement are the value students place on the subject matter, their feelings of competence, their opportunities for autonomy or agency, and the extent to which their needs for relatedness are met (Blumenfeld et al., 2006). The incorporation of social justice is designed to increase the value of CS, and hands-on, scaffolded computer experiences and projects are designed to increase feelings of competence and opportunities for autonomy and agency. To meet relatedness needs, the classes are led by near-peers—high school students from the same community who teach, offer stories, and build connections that help 5th graders navigate competing expectations across their worlds of home, school, and peers (Cooper, Domínguez, & Rosas, 2005). While a near-peer teaching strategy has been used in educational settings for decades, there is very little research on effective implementation strategies, the benefits for the teachers and learners, and no examples of its use to engage underrepresented students in CS. The relational aspects of learning are well documented, and are part of what Pea (2004) describes as the process of “becoming” as children engage in a mutually influential interaction with the world. CSteach builds on these perspectives in that the teaching approach involves modeling and supporting experiential learning, with limited direct instruction.

CSteach is being pilot tested with Latino students, due to their underrepresentation in computing fields. While there is great variation in the group of students called “Latino/a,” studies across regions and samples have led to some consistent themes. The focus of this study is on students of Mexican origin, who make up 63% of the US Latino population and accounted for three quarters of the growth in the US Latino population in the last decade (Ennis, Ríos-Vargas, & Albert, 2011). We use the term “Latino,” because it is commonly used in our community.

CSteach Strategies

CSteach employs exemplary practices for engaging Latino/a students in CS (Computer Science Collaborative Project, 2012). This includes *culturally-targeted programming*, which makes connections to students’ values (Sheridan, Clark, & Williams, 2013) and engages Latino/a youth by building connections to their identity and culture and addressing the needs of the community, not just the individual (Cooper et al., 2005; Solórzano, Villalpando, & Oseguera, 2005). For example, students use technology to research social justice issues, and to identify strategies to address them in their local community. Another strategy is to *challenge the image of CS* including commonly held beliefs that computer scientists are white or Asian, male, geeky, and unconcerned with the social good. To this end, CSteach builds a network of support for 5th grade students through collaboration with their peers, and instruction by high students that understand local challenges (e.g., financial, stereotypes, etc.) and role model an interest and curiosity about CS. Activities are also designed to address CS Standards (Computer Science Teachers Association, 2011). For example, activities both on and off the computer *introduce CS concepts* like algorithms and networks; these are applied in students’ final project where they use a child-friendly programming environment to create an animation that informs people about their social justice topic.

The CSteach curriculum was developed by a team of computer science educators and social justice experts to ensure that CS concepts are introduced in an exciting and developmentally appropriate way that is relevant to students’ experiences. CS and social justice are integrated in several ways. For example, students identify their own personal network and generate ideas about how that network can help them make positive change; then they learn about computer networks, and how information is transferred within the network in ways that are similar to their own personal network. In another activity, students learn to efficiently and safely do online research about a social justice topic, and then apply algorithmic thinking (a logical sequence of steps) to make a plan to address that problem. In a subsequent activity, they apply algorithmic thinking to create and debug a computer program. In this paper, we examine the effectiveness of these strategies by addressing the following questions:

- Does the pilot data suggest that the CSteach program can increase the computer science identities, capacities, and networks of support of Latino/a students?
- How can we maximize the impact of a near-peer approach on learning and identity development in K-12 computer science?
- How can we integrate social justice with CS principles, and does this motivate Latino/a students to pursue and persist in CS?

Methods

Participants

A pilot study in Spring 2013 included 37 5th grade students (mean age=10) and five high school students (mean age=16) in a class that met 2 hours/week for 12 weeks as part of an afterschool program at two schools. Over half of the 5th graders (68%) were female, and 86% self-identified as Latino/a. More than half (54%) of students had a mother or maternal figure who had not completed high school, and 69% of students’ fathers or paternal figures had not completed high school. Most (72%) of the students reported that a language other than English was spoken in their household at least half the time. The five high school near-peer teachers were all female; three seniors, one junior and one freshman; four of them identified as Latina and one identified as African-American.

Research Design and Data Collection Methods

The evaluation of the pilot CSteach program used a mixed-method approach to gather formative feedback on how the program was working, as well as a preliminary assessment of whether the activities were influencing the development of students’ computer science identities, capacities, and networks of support to study computing. Findings from surveys, focus groups, student projects, and observations will be presented.

All 5th grade students and high school student near-peer teachers (Tech Teachers) completed a *pre-post survey* that built on and adapted existing instruments to apply to the CSteach focus and population. The scales measured change in students’ expectations for success in college and career, perceived support to study

computers, CS identity, beliefs about the value of computing, confidence and ability to use technology to address community needs, expectations for success in computing, and ability to evaluate web-based information. A *focus group* was conducted with the five high school Tech Teachers at the end of the semester and lasted 60 minutes. The questions focused on challenges and how they handled them, how the 5th grade students benefited from the program, how they benefited from the program, and suggestions for improving their training and preparation.

Student *projects* were scored using a rubric that was based on the Computer Science Teachers Association (CSTA) national standards, which include computational thinking, as well as awareness about the community, global and environmental impacts of technology. The projects included a set of Power Point slides and an animation created in the Scratch programming environment. The slides were scored on 4 characteristics (e.g., contains a graph; poses questions about a social justice topic). The Scratch animation projects were scored on 6 characteristics (e.g., keyboard-driven events; moving characters). The raters used a 3-point scale (0=did not meet standard, 1=partially met standard, 2=demonstrated proficiency); two raters scored the projects and discussed discrepancies in scores until agreement was reached.

Data Analysis

The surveys were analyzed by conducting basic, descriptive statistics, such as frequencies, means, and crosstabs to explore patterns in the data. Additionally, paired samples t-tests, chi-square, and one-way analysis of variance (ANOVA) were conducted to compare students' pre-post responses on survey scales and to measure differences across key demographics like gender, parental education level, and prior computer use. Simple descriptive analyses were used to summarize the results of scoring student projects. The focus groups were transcribed verbatim and analyzed using domain analysis; codes were generated deductively, based on evaluation questions and theoretical constructs from the CSteach program design, and inductively, based on important themes that emerged from analyses. Codes were organized into taxonomies and patterns were compared across interviews. Sample analytic categories identified in the analysis include: Motivation for participating in the CSteach program, Teaching strategies, Challenges faced and how they dealt with them, Beliefs about computing and the social good, Tech Teacher outcomes, and Elementary student outcomes.

Results

The results provide a picture of the implementation and impact of the first (pilot) semester of CSteach, as well as information to guide the next phase of the program. The findings are organized by Research Question

- Does the Pilot Data Suggest that the CSteach Program Can Increase the Computer Science Identities, Capacities, and Networks of Support of Latino/a Students?

Although the sample size is small, we detected modest gains in most of the domains measured by the survey among the elementary students, and received positive evaluations of the impact of the program on both elementary and high school students' identities, capacities and networks of support. The results also provide information about how the program can be strengthened. Based on their surveys, the strongest gains were in 5th grade students' confidence to use a computer to address needs in their community (statistically significant change). There were also increases in self-efficacy to use and study computers, and in their CS identity, specifically the belief that computer science is a creative endeavor. In the focus groups, high school students talked about how the 5th graders changed. Clara [names were changed] stated "They're definitely more confident. They call me over sometimes and say, "teacher look at what I know how to do," and I'm like, "that's awesome." Natalia commented "I think they are more confident. At the beginning of the year, there were some girls that were very unsure of the computers, but like now, I see them helping each other. And they really know how to work the computer and all the programs."

Results from surveys, student projects, and interviews suggests that there were some areas in which CSteach engaged students in understanding and applying CS and social justice concepts. In their surveys, students reported significant increases in the frequency with which they evaluate the credibility of information, such as judging whether internet sites are trustworthy, and in the ratings of their ability to organize and manipulate information. The results of scoring student projects suggest that students developed CS capacities, including proficiency in manipulating data and evaluating information, and in the ability to gather and manipulate data using digital tools. The projects also show that most were able to integrate a social justice topic into a programming activity (83%), and social justice questions into a Power Point presentation (78%). Natalia, a Tech Teacher, commented on students' growing capacity to use technology for academic purposes, rather than just for fun: "Now they know how to make a chart off of data they collected. They know how to do a presentation, they know how to do Power Point, they know how to do a lot of things for middle school, and that will prepare them."

The data also suggest some impact on students' perceived networks of support to pursue computing. For example, the Tech Teachers described how students began to see their peers as tech savvy. Sofia, another Tech Teacher, said that students began to see one another as knowledgeable and capable in computing: "When they use Scratch, one thing that's good is that sometimes I'll admit that I don't know how to do it, and there's always someone who gets it and I'll send them to go help another student. So then they're like teaching themselves."

Similar to the 5th grade students, the Tech Teachers' largest gains were in their ability to use computers to address community needs, and in their identity as computer scientists. Several important themes emerged in interviews, including the development of leadership skills, influences on career and educational paths, and career preparation and personal growth. Alisa said "The CSteach program benefited me by outfitting me with knowledge and skills involving computers, leadership and overall abilities to interact with others, that I didn't know I had in the first place." The Tech Teachers also developed an identity as role models to the 5th graders, which increased their confidence to be teachers and leaders, as well as their motivation to do so. Ana said: "The fact that you're older, it gives you a sense of leadership. Since they're younger, they're like, 'Oh, she's older, so I'm just gonna look at her as a role model or a leader.'"

- How do we maximize the impact of a near-peer approach on learning and identity development in K-12 computer science?

In the pilot study, the near-peer approach involved five female high school students that had participated in other technology-based after school programs; they were divided across the two pilot school sites to balance their experience and teaching confidence. Observations and focus group interviews were used to evaluate the effectiveness of the near-peer approach and identify ways to strengthen it. For example, we tested a model where the Tech Teachers led 30% of the activities at the beginning of the semester, and then slowly transitioned to leading 70% by the end of the semester. Although this helped to ease the Tech Teachers into their role, it also set up an expectation by both the 5th grade and high school students that the adult teacher was the primary authority, and reduced the message that the near-peers were tech savvy role models. We now have the Tech Teachers leading most of the class activities starting at the beginning of the semester; the adult teacher introduces each activity to set the tone and expectations, and the Tech Teachers take the 5th graders through each step and provide support. This helps to clarify their role vis-à-vis the adult teacher.

In addition to increasing the leadership of the high school students, the pilot study suggested other ways to strengthen the near-peer component are to create a targeted recruitment strategy and provide more training. For example, some of the high school students were motivated to participate because they liked computers and wanted to get younger children excited about them, but several did it just for the money. The ideal near-peers would have a stronger orientation to computers, and be excited about sharing their passion. The high school students also needed additional training (e.g., in Scratch) to have the confidence to help the younger children. As Sofia said, "They're more tech-savvy. They were introduced to these things sooner than we were." Others gained a greater appreciation for what it means to teach: "After this, I feel totally sympathetic to all my teachers. This stuff is hard!" The data were used to inform what was covered in the next round of high school student recruitment and training, the adult teacher training, and in the curriculum.

- How can we integrate social justice with CS principles, and does this motivate Latino/a students to pursue and persist in CS?

CSteach aims to increase the relevance of CS by integrating it with social justice issues in two ways. First, we aim to motivate students to deeply engage in CS concepts and use them to create a project on a topic that is meaningful to them and their community. Topics so far have included bullying, drugs and alcohol, animal cruelty, poverty, among other issues. The Tech Teachers thought that the social justice aspect really was motivating for the children. Adriana said: "I know a lot of people feel really passionate about their thing...about their social justice thing." But they also thought it would be more motivating if the students presented their projects to others: "I think that would be pretty cool if their parents could come, and they could tell them, 'Hey, this is what I think is wrong in our community, and this is what I think we should do.'"

The analysis of their projects also suggested that the topic was motivating: over half used event-driven dialogue (one of the more complex features) to either raise awareness or propose ways to address their social justice topic.

The second way that CS was integrated with social justice was to use information technology to showcasing the issue and the steps to address it, including the strengths of their community. Most students could apply algorithmic thinking to both address the social justice issue and to create a programming project about that topic. However, the Tech Teachers suggested adding more reflection about what they were learning, as well as more explanation about why they were learning it: "Show them stuff that they can do with this knowledge, that's what I wanna do. Like you can do this or you can build this, but you need to learn the basics first." To this

end, we rewrote the curriculum to clearly describe all the CS principals that were addressed in each activity, as well as how to reinforce them in a language that the 5th grade students could understand. This meant letting go of some CS topics, like vector graphics, because they were more difficult to explain by linking to a social justice principle.

Conclusion

The results suggest that the CSteach program has the potential to support Latino/a youth on the path to becoming computer scientists. The pilot study suggests that combining near-peer teaching and computing for the social good was motivating, and has the potential to increase students' CS identities, capacities, and networks of support; a necessary step toward deep learning. However, additional work needs to be done to increase students' engagement with CS (e.g., include fewer CS concepts and address them with greater depth), strengthen the social justice component (e.g., scaffold final projects to more strongly reflect how technology can be used to address community needs), and maximize the impact of the near-peers (e.g., more targeted recruitment and training). These results are being used to inform the next stages of implementation, and the curriculum continues to be revised based on input from the near-peers, the adult teachers, and the partnering afterschool program. In the 2013-14 school year, CSteach is being implemented in 8 schools with 160 elementary and 17 high school students.

References

- Bang, M., & Medin, D. (2010). Cultural processes in science education: Supporting the navigation of multiple epistemologies. *Science Education*, 94(6), 1008-1026.
- Campbell, M. (2011). Hispanic students: 2010 statistical survey. *Hispanic Outlook*, 46-48. Retrieved on October 29, 2013 from <https://www.wdhstore.com/hispanic/data/pdf/jan3-hispanicstudents.pdf>
- Blumenfeld, P.C., Kempner, T.M., & Krajcik, J.S. (2006). Motivation and cognitive engagement in learning environments. In R.K. Sawyer (Ed.), *Handbook of Learning Sciences*, 475-488. New York: Cambridge.
- College Board (2011). *AP CS: Principles Curriculum Framework* retrieved from <http://www.collegeboard.com/html/computerscience/index.html>
- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, 13(1), 15-42.
- Computer Science Collaborative Project (2012). Exemplary practices for engaging Hispanic/Latino(a) youth in computer science. (www.cscproject.org/index.php?q=node/162).
- Cooper, C.R., Domínguez, E., & Rosas, S (2005). Soledad's dream: How immigrant children bridge their multiple worlds and build pathways to college. In C. R. Cooper et al. (Eds.) *Developmental pathways through middle childhood: Rethinking context and diversity as resources* (pp. 236-259). NJ: Erlbaum.
- Ennis, S.R., Ríos-Vargas, M. & Albert, N.G. (2011). The Hispanic population: 2010. 2010 Census Briefs. Retrieved on January 7, 2012 from <http://www.census.gov/prod/cen2010/briefs/c2010br-04.pdf>
- Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., Malyn-Smith, J., & Werner, L. (2011, March). Computational Thinking for Youth in Practice. *ACM Inroads*, volume 2, no 1, 32-37.
- Leonard, J., Brooks, W., Barnes-Johnson, J., & Berry, R.Q. (2010). The nuances and complexities of teaching mathematics for cultural relevance and social justice. *Journal of Teacher Education*, 61(3), 261-270.
- Margolis, J. (2008). *Stuck in the shallow end: Education, race, and computing*. Cambridge, MA: MIT Press.
- Nasir, N.S. & Hand, V.M. (2006). Exploring sociocultural perspectives on race, culture, and learning. *Review of Educational Research*, 76(4), 449-475.
- Pea, R. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences*, 13(3), 423-451.
- Sheridan, K.M., Clark, K., & Williams, A. (2013). Designing games, designing roles: A study of youth agency in an urban informal education program. *Urban Education*, 48(5), 734-758.
- Solorzano, D., Villalpando, O., & Oseguera (2005). Educational inequities and Latina/o undergraduates in the US: A critical race analysis of their educational progress. *Journal of Hispanic Higher Education*, 4, 272-294.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. CNS-1240756.

Science, Technology, Body and Personhood: The Concept of Health Emerging in High-Tech Modern Medicine Practice

Federica Raia^{1,2}, Martin Cadeiras², Ali Nsair², Daniel Cruz², Grecia Ramos¹, Claire Alvarenga¹, Valeria M. Rivera¹, Kristina Barrientos¹ and Mario C. Deng²

University of California, Los Angeles

¹Graduate School of Education and Information Studies, Department Education

²David Geffen School of Medicine, Department of Medicine

Email: raia@gseis.ucla.edu, MCadeiras@mednet.ucla.edu, ANsair@mednet.ucla.edu, DCruz@mednet.ucla.edu, greciaramos48@yahoo.com, clairealvarenga@yahoo.com, vmrivera01@ucla.edu, kbarrientos90@gmail.com, MDeng@mednet.ucla.edu

Abstract: In this participatory action research project, we studied the process of practice in high-tech modern medicine as a particularly powerful domain in which integration of science, technology, body and personhood needs to take place. We posit that this integration demands a conceptual evolution of a static concept of health into a concept of health that is based on the ability to adapt and self-manage. There are different level of learning about living with high technology modern medicine, including from the patient perspective learning about own body, learning to live with the machine, learning to live with uncertainties, and, from the physician perspective, to understand the relationship between supporting the patient in developing a new sense of self, for technology to be completely integrated in the patient's life, past, present, future and self and the ownership over the specific critical health decisions with its physiological manifestation and course.

Introduction

In this paper we explore the process of practicing high-tech modern medicine as it is the chronic and acute advanced heart failure practice, often in the setting in critical care, as a particularly powerful domain in which integration of science, technology, body and personhood takes place. Specifically, we investigate how stakeholders (patient, family, healthcare professionals) in Advanced Heart Failure (AdHF) medical encounter negotiate the tensions between formal definitions of health and choice in the actual practice. This study is of a particular interest for the field of the learning sciences because in agreement with Rogers Hall's call (Hall, 2005), it opens up new perspectives between the larger scales of classification systems at the *society/policy level* such as the ones outlined in the Patient Protection and Affordable Care Act 2010 (ACA) (Fineberg, 2012) and the lower scale, such as the level of an individual making informed decisions and of person's choice – *Individual level* (Washington & Lipstein, 2011). This research focuses on the middle level – *practice level* – where people negotiate these tensions in actual practice, the person-in interaction level (Hall, 2005).

Our focus on the practice level where participants are in interactions has important implications in the context of the rich tradition of research on illness in the sociology and anthropology of medicine and medicine education. In the medicine education field, the attention is focused on medical school students and most of the literature centers on General Practice and the application of models of care such as patient centered care (Stewart et al., 2003). In sociology and anthropology of medicine, there is extensive research on chronic illness either at the policy societal level or on the individual level. Patient-centered perspective provides a rich literature on the experience of chronic illness (Pierret, 2003) as biographical disruption (Bury, 1982) and processes of adaptation to a new life (Charmaz, 1995; Pierret, 2003). Yet, this research has been mostly removed from the rich context of the specific disease as an ontological reality of the body with few important exceptions (e.g. Mol, 2002) and from high-tech modern medicine of chronic and acute practice, often in the setting in critical care whose situated specificity of disease it is not reducible into a melting pot of chronic illness (Timmermans & Haas, 2008). The big impact that the field of science studies has had and continues to have in the learning sciences is based on the fundamental focus on practice, showing that in the scientific practice of each discipline, criteria, argumentation, instrumental expertise and methodological sensitivities develop with their own specificity (Knorr-Cetina, 1981; Latour & Woolgar, 1979). We claim that in the context of medical practice, as seen in the study on science in action (Latour, 1987), discipline-dependent knowledges of practice, concepts of care, disease, decision making and self do emerge with their own specificity. The experience of the body and of the self, interacting with physiological constraints and manifestations and the interactions with high technology make participation in the practices a learning experience. This presupposes both: what participants become and how they act as knowers (Roth & Lee, 2007).

We take the situated specificity of disease in the medical practice as the starting and returning point of all our discussion. Specifically, this study is situated in the rich context of AdHF where high technological

advances play a fundamental role in people's lives and practices offering a window on what high-tech modern medicine entails for individuals in their interactions at the practice levels. Specifically, the study is situated in the AdHF context of practice as a preeminent example of high-tech modern medicine with its first implantation of a total artificial heart as a permanent device in 1984 and the first use of a wearable left ventricular assist device in 1994 (Goldstein et al., 1998). This means that 'biographical disruptions' (Bury, 1982) entail integration of technological advances such as being kept alive mechanically, while continuing participation in family, work and social life; unknown 30 years ago. These experiences are novel to patients and their family and unknown to physicians.

Society Level / Individual-Level: Health, Choice and Decision-Making

Society/Policy Level: In the USA, ACA has authorized the creation of the Patient-Centered Outcomes Research Institute (PCORI). This non-profit organization is entrusted by the Congress to identify national priorities for research and to fund research guided by patient-centered questions that can provide information about the best available evidence to help patients and their health care providers make more informed decisions, better understand the prevention, treatment and care options available to them, and the science that supports those options (Washington & Lipstein, 2011). At the *Individual Level*, providing information to best support understanding and make informed decisions has on one hand contributed to moving the process beyond informed consent, rooting it in the multifaceted and more complex (Dy & Purnell, 2012) frame where all stakeholders involved – patient, caregivers and health professionals – work together (Charles et al., 1997), each bringing a complementary and necessary perspective to decision support (Elwyn et al., 2010). At the same time, it is based on normatives of autonomy, individuality and effectiveness which entail responsibility, and accountability for the consequences of people decision (O'Neill, 2002).

Practice Level: How are these concepts of choice, individuality, responsibility and autonomy negotiated in the context of high-tech modern medicine (Timmermans & Haas, 2008) in which AdHF practice unfolds? We address this question in our study

Methods

This study is part of an ethnographic and participatory research project (2011-2013) conducted in a large University Medical Center in California (Raia and Deng, in press).

Data Collected and Analysis

The data have been collected in three stages:

Stage.1:*Narratives.* We conducted exploratory interviews of stakeholders and 1 hour to 2 hour-long open-ended interviews of participant patients and their family to collect their narratives on the experiences in advanced heart failure.

Stage 2. *Medical Encounter Recordings.* Three recruited AdHF cardiologists were audio- and video-recorded in their AdHF medical encounters with their patients (N=25). Of the 25 patients in the study, 18 patients and their family were recorded in their medical encounters with the participating physician for a period of 1-2 years AdHF Medical encounters were audiotaped in *in-patient* medical rounds and audio and videotaped in *out-patient* clinic medical encounters:i) In *in-patient* medical rounds medical encounters AdHF cardiologists are audio-recorded while on rounds in the hospital on AdHF patients who are hospitalized (AdHF in-patients). These patients are often admitted or transferred from a regional hospital into the larger hospital in an advanced decompensated and often life-threatening condition. It is during these hospital rounds that AdHF cardiologists sometimes meet the patients and their caregiver(s) for the first time. Often, it is during the stay in the hospital that patient and caregivers are faced with fundamental decision making of heart transplantation or mechanical circulatory support implantation. Each patient is in a single-bed room. In the Cardiothoracic Intensive Care Unit (CTICU) patients are recovering from cardiothoracic surgery for either an assist mechanical circulatory support device implantation or heart transplantation; in Coronary Care Unit (CCU) and Cardiac Observation Unit (COU) patients are under observation and often listed for heart transplantation. In-patient medical rounds are structured so that the attending AdHF cardiologist, the physician on call responsible for the hospitalized patients, walks from unit to unit visiting each patient in her/his hospital room. ii) Out-patient clinic medical encounters- In these encounters, AdHF patient, not more hospitalized or not yet hospitalized, come in the hospital for a medical appointment and meet with the AdHF cardiologist in examination rooms.

Stage 3. *Co-Generative dialoguing (CoGen)* (Elder and Levin, 1991). AdHF cardiologists, whose interactions have been recorded in Stage 2, participate in weekly audio/video-recording viewing sessions as part of the research team. These research encounters (Stage 2) are videotaped (. The diverse perspectives of participating in the study and the researchers are shared, reflected upon and elaborated allowing all to develop and understanding of the issue at stake. In these weekly sessions, four AdHF participating cardiologists, together with student researchers and researcher, reviewed the taped medical encounters in which they participated and discuss the elements that emerged to be important to them. Reviewing together the data in

CoGen session allowed to unfold a richer perspective on the *practice* of high-tech modern and amongst other things to address questions most relevant to practice, depending on the clinical circumstances and on the physician's stance and knowledge at specific times in the medical encounters (Wirtz et al., 2006). It also allowed checking for validity of the emerging patterns identified in this study. As in the tradition of participatory research we all co-author in this paper. Data from Stage 2 medical encounters data was transcribed by student researchers co-authors and the first author of this paper and checked for agreement.

Major Findings, Conclusions, and Implications

When a patient and family member are confronted by choices of heart transplantation or mechanical circulatory support device implantation, to make a decision, they must accept that a substitution of a malfunctioning organ (in the first option) or a substitution of the pumping function (in the second option) is possible. This requires one's body or loved one's body to be seen as *conspicuous* (Heidegger, 1962) in its malfunctioning. This gaze on the properties and natural function of the body parts, not as a whole 'me,' moves to consider it as an instrument that has failed to do its job, with malfunctioning or broken-down part(s). In this, according to Heidegger, *familiarity* with what was known has broken down (Heidegger, 1962). But what happened when a patient wakes up from the surgery and finds him/herself with a heart transplantation or mechanical circulatory support device implantation?

The following excerpts two encounters between one of the participant-researcher doctors and Mr. J, a 33 year old patient, constitutes a representative example.

Context: Mr. J is in his thirties and always been physically very active. His condition of AdHF developed about ten years ago but at the moment has dramatically deteriorated. He is emergently hospitalized and resuscitated after having suffered a sudden death due to lethal arrhythmia. Mr. J accepts to be a candidate for heart transplant with the only alternative choice to die possibly very soon from another episode of fast chaotic heart rate for which he has been hospitalized. He is listed on the highest emergency category in the USA national list for heart transplant but, based on the unpredictable course of a reoccurring fast chaotic heart rate the AdHF health professional team had suggested an implantation of BiVAD (a mechanical circulatory support device) as a bridge to transplantation, if another of these episodes with loss of conscience were to occur. Mr. J remained hospitalized in intensive care in Cardiac Observation Unit under close monitoring and the lethal arrhythmia followed much sooner than expected. Mr. J is just recovering from the BiVAD implantation surgery in the CTICU, when Dr. D enters. The plan for Mr. J is to completely recover from the BiVAD surgery and be discharged from the hospital to qualify again to be a candidate for heart transplant in accordance with the USA national directives for heart transplantation listing. Because of the BiVAD implantation, Mr. J also needs to have two trained caregivers who can assist him living with the assist devise. Dr. D and Mr. J meet for the second time now.

Encounter 9

- 1 Dr. D: Hello young man!
- 2 How is life?
- 3 Mr. J: I'm fine
- 4 (3 sec pause)

Three second is a long silence as as long silences are reported to cluster around the one-second interval (0.9–1.2 s) (Jefferson, 1989). Here the doctor is waiting for the Mr. J to continue talking –as he usually does, but Mr. J does not. Silence. If we compare it with another day, we see that Mr. Jay respond differently to the doctor salute:

Encounter 10

- 1 Dr. D: Hi young man
- 2 Mr. J: Hi!
- 3 Dr. D: How is life?
- 4 Mr. J: Getting better. Had a nice breakfast.
- 5 Dr. D: I saw you (3 s)
- 6 Mr. J: Got the eggs
- 7 Got the turkey sandwich
- 8 Got a nice little fruit cup
- 9 My feet are starting to feel a little bit better

As Dr. D. describes in CoGen, entering the room on Day 9, he expected to see Mr. J to be “sitting up in the middle of bed and despite adversity: “a handsome young man with his long hair looking like a Jesus full of

inner life, vision energy and dignity holding on each side his two pumps,” Dr. D gesturing the expected Mr. J’s position gives an impression he is expecting to see somebody in charge: “a victorious young men” “Yes! Having survived a major heart attack, being resuscitated, and undergone a major heart surgery, he should be in the same framework as I am” and in this framework “*just* fine is not a good start after a major surgery” because “*just* I’m fine” does not grasp the immensity of what just happened, what Mr. J went through

Day 9 continues

- 5 Dr. D: good
- 6 So, on Friday a week ago when it came up
- 7 Friday morning was when we said let’s do
- 8 the high emergency status heart transplantation listing
- 9 and maybe we have an organ coming up during the weekend
- 10 That was also when we first met
- 11 Mr. J: yeah
- 12 Dr. D: And then we say ok and Dr. C was already little skeptical
- 13 and said maybe if anything happens we have to go ahead
- 14 and this was in the night
- 15 And so then it was clear
- 16 we have to recommend the Bi-ventricular Assist Device
- 17 Dr. D: How did it feel when the question came up now the assist device.
- 18 You were prepared on Saturday morning
- 19 you know on the 12th?
- 20 Mr. J: it was it was a little weird, you know,
- 21 kind of I mean (3.0 sec pause)
- 22 It’s just kind of
- 23 uh, uh a little trippy uhmm
- 24 hearing something like you know
- 25 like that
- 26 It weirded me out with gangliectomy you know
- 27 you don’t wanna do it (2.0 sec pause)
- 28 but if it’s really for the best you kind wanna do it,
- 29 but it’s just weird to think about it you know (1.48 sec pause)
- 30 and uhhh I am glad it’s happened you know (3 sec)
- 31 Dr. D: yeah, yeah
- 32 It’s a big (1.5 sec pause) thing
- 33 Mr. J: Yes! (2 sec pause)
- 34 Very big thing

At line 20, Mr. J starts sharing his experience for the first time. It is an experience of estrangement waking up with the machine as part of one’s own body. This experience is shared by other patients as Mr. J, a few months later, during an interview, recalls feeling “like a scary character in the movies, half man, half machine.”

There are the long silences (1.5 to 3 seconds) in Mr. J’s talk elaborating his experience. These silences are not interrupted by the physician. Dr. D notices also that Mr. James does not name the BiVAD, instead saying it was ‘trippy ... hearing something, you know, like that’ (line 20-25) and that he uses the word ‘weird’ three times (line 20, 26 and 29). It is on this word, ‘weird,’ that Dr. D operates the first transformation, from ‘weird’ into ‘big’ (line 32). Note that the pause of 1.5 seconds (line 32) after Dr. D uses the word ‘big,’ gives space to the word and greater emphasis on the transformation. ‘Yes! It is a very big thing.’

The patient starts talking and from here the doctor slowly moves into a discussion of what is coming in the future to maintain a space where the patient will in time elaborate the experience.

Our data show that accepting a life with heart transplantation or mechanical circulatory support device implantation is not done upon making the decision. It is a complex dynamic oscillatory process of adaptive synchronization. This is based on periods of patient’s anguish and despair, feeling lost in an *unfamiliar* life, and countervailing fluctuation of mood between disappointed expectation and support offered by doctors.

The expectation of the doctor entering Mr. J’s room is to find a person who won a battle against death in pursuing an agreed upon course of action (decision making on mechanical circulatory support device) that allows the patient to continue living his life. When the doctor meets the patient on Day 10 after surgery, there is no correspondence with his expectations. The patient is lost. In the decision-making framework, with normatives of autonomy, individuality and effectiveness, the patient’s reaction is not readable as being lost, but

is either ignored or, worse, is understood as “patient is not grateful for the gift of life and for the healthcare and family efforts in supporting the patient in the difficult decision.”

In high-tech modern medicine making a decision is understood as a process of oscillation between asynchronic state of being “not on the same page” and a return to a synchronic state. The interactional pattern sees the doctor respond ‘on not being on the same page’ with the goal of formulating for the patient possibilities for the future. When a synchronic state is reached in one encounter, the expectation is not that it is maintained, as doctors are observed to respond on this day ‘not being on the same page,’ understating that is a process of integration of science, technology, into a person’s life: a recursive transformation.

The understudying of decision-making in the practice of high-tech modern medicine where at the person level, there is learning to live with the machine and learning to live with uncertainties is an oscillatory process of synchronization fundamentally different from a point of arrival (decision) entailing responsibility, and accountability for the consequences of people decision (O’Neill, 2002). It is not regulated by normative of autonomy and individuality but by the imperative to care.

References

- Bury, M. (1982). Chronic illness as biographical disruption. *Sociology of Health and Illness*, 4, (2): 167-182.
- Charles, C., Gafni, A., & Whelan, T. (1997). Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). *Social Science & Medicine*, 44, 681-692.
- Charmaz, K. (1995). The body, identity, and self: Adapting to impairment. *Sociological Quarterly*, 36(4), 657-680.
- Dy, S.M., & Purnell, T.S. (2012). Key concepts relevant to quality of complex and shared decision-making in health care: A literature review. *Social Science & Medicine*, 74, 582-587.
- Elden, M., & Levin, M. (1991). Cogenerative learning: Bringing participation into action research. . In W. F. Whyte (Ed.), *Participatory action research* (pp. 127-142).
- Elwyn, G., Frosch, D., Volandes, A.E., Edwards, A., & Montori, V.M. (2010). Investing in deliberation: a definition and classification of decision support interventions for people facing difficult health decisions. *Medical decision making : an international journal of the Society for Medical Decision Making*, 30, 701-711.
- Fineberg, H.V. (2012). Shattuck Lecture. A successful and sustainable health system--how to get there from here. *The New England journal of medicine*, 366, 1020-1027.
- Goldstein, N.E., Lampert, R., Bradley, E., Lynn, J., & Krumholz, H.M. (2004). Management of implantable cardioverter defibrillators in end-of-life care. *Annals of internal medicine*, 141, 835-838.
- Hall, R. (2005). Reconstructing the learning sciences. *Journal of the Learning Sciences*, 14(1), 139-155.
- Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). New York: Harper.
- Knorr-Cetina, K. (1981). *The manufacture of knowledge : an essay on the constructivist and contextual nature of science*. Oxford ; New York: Pergamon Press.
- Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*: Harvard university press.
- Latour, B., & Woolgar, S. (1979). *Laboratory life : the social construction of scientific facts*. Beverly Hills: Sage Publications.
- Mol, A. (2002). *The body multiple : ontology in medical practice*. Durham: Duke University Press.
- Mol, A. (2008). *The logic of care : health and the problem of patient choice*. London ; New York: Routledge.
- O’Neill, O. (2002). *Autonomy and trust in bioethics*. Cambridge ; New York: Cambridge University Press.
- Pierret, J.(2003). The illness experience: state of knowledge and perspectives for research. *Sociology of Health& Illness*, 25, 4-22.
- Raia, F., & Deng, M. C., (in press) Relational Medicine –Personalizing Modern Healthcare: The Practice of High-Tech Medicine As A RelationalAct. Imperial College Press, London
- Roth, W.M., & Lee, Y.J. (2007). "Vygotsky's neglected legacy": Cultural-historical activity theory. *Review of Educational Research*, 77, 186-232.
- Stewart, M., Brown, J.B., Weston, W.W., McWhinney, I.R., McWilliam, C., & Freeman, T.R. (2003). *Patient-Centered Medicine: Transforming the Clinical Method*. Oxon: UK: Radcliff Medical Press Ltd.
- Timmermans, S., & Haas, S. (2008). Towards a sociology of disease. *Sociology of Health & Illness*, 30, 659-676.
- Washington, A.E., & Lipstein, S.H. (2011). The Patient-Centered Outcomes Research Institute—promoting better information, decisions, and health. *The New England journal of medicine*, 365, e31.
- Wirtz, V., Cribb, A., & Barber, N. (2006). Patient-doctor decision-making about treatment within the consultation - A critical analysis of models. *Social Science & Medicine*, 62, 116-124.
- Zimmermann, C. (2007). Death denial: obstacle or instrument for palliative care? An analysis of clinical literature. *Sociology of Health & Illness*, 29, 297-314.

“With-Me-Ness”: A Gaze-Measure for Students’ Attention in MOOCs

Kshitij Sharma¹, Patrick Jermann², Pierre Dillenbourg¹

1. Computer Human Interaction in Learning and Instruction (CHILI)

2. Center for Digital Education (CEDE)

École Polytechnique Fédérale de Lausanne, Lausanne

Email: {kshitij.sharma, patrick.jermann, pierre.dillenbourg}@epfl.ch

Abstract: We propose a gaze-based indicator of students’ attention in a MOOC video lecture. We report the results from an eye-tracking study during a MOOC lecture. We define the gaze-based indicator of students’ attention as “with-me-ness”. This answers a question from teachers’ perspective “*how much are the students with me?*” With-me-ness is defined at two levels: perceptual, following teacher’s deictic acts- and conceptual – following teacher discourse. We conducted an experiment with 40 participants and observed a significant and positive correlation between the two levels of with-me-ness and the posttest scores.

Introduction

We use eye-tracking methods to measure the students’ attention during MOOC lectures. In the background, the general question we address is “how can we help students to watch videos more efficiently?” In this study, we relate attention measures to how much students learn from a simple video.

There are many factors affecting classroom performance of the students: previous grades (Astin, 1971), students’ efforts and motivation (Grabe and Latta, 1981), socioeconomic differences (Kaplan, 1982), quality of schooling (Wiley, 1976), attention (Good and Beckerman, 1978;) and participation (Finn, 1989). We do not claim we address all of them. We focus on attention as one dimension of students’ performance that may be especially relevant when watching video lectures. We tackle this question from a teacher’s perspective: “*How much the student is with me?*” Accordingly, we call this gaze-based measure “With-me-ness”: is the student really “following” the lecturer, i.e. paying attention to the elements of the display that correspond to the instant behavior of the teacher? We selected two aspects of teacher’s behavior that will influence the student’s attention, what the teacher says and the teacher’s deictic references: in several MOOCs, learners may see the hand or the pen of the teacher in overlay to the slides.

Related Work

Attention as Factor Affecting Students’ Performance

Attention had been proved to be associated with performance in many studies related to visual tasks (Yantis and Jonides, 1984; Prinzmetal et. al., 1986; Juola et. al., 1991). In a visual comparison of two line segments Prinzmetal et. al. (1986) and Juola et. al., (1991) showed that the more attentive participants were more correct in selecting the longer line segment. Yantis and Jonides (1984) found similar results in visual perception tasks.

In a classroom attention is “*listening, sitting and working on assigned tasks*” -Homes et. al. (2006). In the context of academic performance previous research has shown strong association between students’ attention and academic performance (Finn, 1989).

We do not claim that visual attention is a deep indicator of learning activity: looking at a piece of information is often a condition to interact with it (read it, select it, move it) but it does not indicate how deeply the learner processes this information (does he understand or give meaning to it). However, in the context of video watching, it’s one of the rare behavioral information that can be collected. Moreover, as we will explain, “with-me-ness” is not simply measuring attention but co-attention, i.e. whether the learner is paying attention to the elements that the teacher is referring to, verbally or through deictic.

Eye-Tracking and Performance/Expertise

Previous research provides insights about the relationship between the gaze patterns and the behavioral and performance indicators in diverse scenarios. Existing results show a clear relation between gaze patterns and expertise. In a collaborative Tetris game, Jermann et al. (2010) showed that experts pay more attention on the stack of Tetronimoes where pieces land while novices allocate more attention to new pieces falling from the top.

Existing results also show a clear relation between gaze patterns and task based performance. In a pair-programming task, Jermann et al. (2012) showed that the pairs with high quality collaboration have synchronized more their gaze on different parts of a program than the pairs that do not collaborate well. In a similar task, Sharma et al. (2011) showed that the good performing pairs pay more attention to the data-flow of the program than the poor performing pairs. Moreover, Sharma et al. (2013) showed that while describing the functionality of a program the good performing teams had more gaze on the expressions in the program while

poor performing teams have equal distribution of gaze on different parts of the program during similar phase of the task.

The Present Study and Research Question

We present the results of an eye-tracking study contextualized within a MOOC class. We choose MOOC videos as a stimulus for the eye tracking because the effectiveness of video as a medium for delivery of educational content is a controversial issue in the current rise of online education. This issue has already been studied and established in literature (Paivio (1991), Mayer (2003), Schwartz (2007)). Through this contribution we propose a gaze-measure to capture students' attention (introduced in the section "With-me-ness") in the context of MOOC lectures. The present study addresses following methodological question:

1. How can we define attention through a gaze-measure? At what levels can we define the attention or from a teacher's perspective the measure of "With-me-ness"?

We describe with-me-ness at two different levels: perceptual and conceptual. Apart from the methodological question, through this contribution we address following research question:

2. How is perceptual and conceptual levels of with-me-ness is related to performance?

With-Me-Ness

With-me-ness is defined at two levels: perceptual and conceptual. There are two ways a teacher may refer to an object: with deictic gestures, sometimes accompanied by words ("here", "this variable") or only by verbal references ("the counter", "the sum") Deictic references are implemented by using two cameras during MOOC recording, one that captures the teacher's face and one, above the writing surface, that captures the hand movements. In some MOOCs, the hand is not visible but teacher uses a digital pen whose traces on the display (underlining a word, circling an object, adding an arrow) act as a deictic gestures. Perceptual with-me-ness measures if the student looks at the items referred to by the teacher through deictic acts. Conceptual with-me-ness is defined by the discourse of the teacher: do students look at the object that the teacher is verbally referring to, i.e., .e. that the teacher is referring to a set of objects that are logically or semantically related to the idea he is referring to. Figure 1 shows the relative temporal granularities of the two levels of with-me-ness and different levels of perceptual with-me-ness.

The notion of with-me-ness is also comparable with measures of gaze coupling that were developed in studies involving dual eye-tracking. Cross-recurrence (Richardson et. al., 2007) reflects how much the gazes of two people follow each other during interaction. Recurrence is highest during references and recurrence level is related to the quality of interaction (Jermann & Nüssli, 2012).

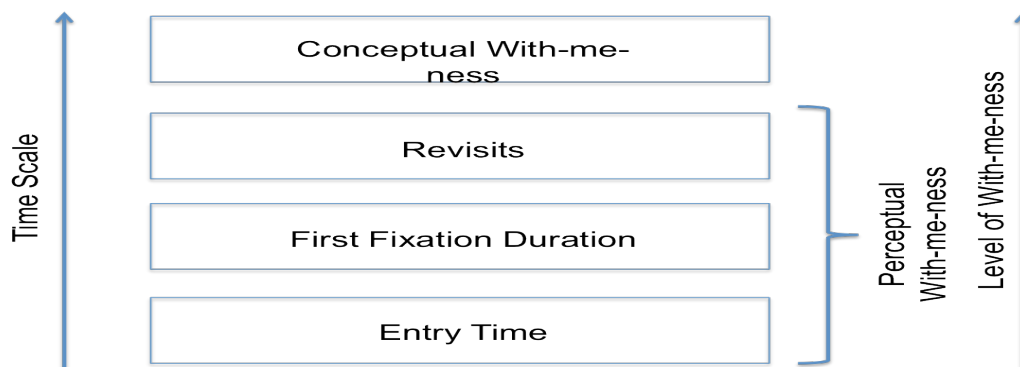


Figure 1. Temporal description of the two levels of with-me-ness and the sub-levels of perceptual with-me-ness.

Perceptual With-me-ness: The perceptual "With-me-ness" has 3 main components: entry time, first fixation duration and the number of revisits. Entry time is the temporal lag between the times a referring pointer appears on the screen and stops at the referred site (x,y) and the time student first time the student gaze stops at (x,y). First fixation duration is how long the student gaze stops at the referred site for the first time. Revisits are the number of times the student gaze comes back to the referred site.

Conceptual With-me-ness: The teacher may also refer. We measure how often a student looks at the objects verbally referred to by the teacher during the whole course of time (the complete video). In order to have a consistent measure of conceptual "With-me-ness" we normalize the time a student looks at the overlapping content by slide duration.

Experiment

In the experiment, the participants watched two MOOC videos from the course "Functional Programming Principles in Scala" and answer programming questions after each video (particular to videos).

The video was not made for this experiment but select from an existing popular MOOC. Participants' gaze was recorded, using SMI RED 250 eye-trackers, while they were watching the videos. Participants were not given controls for the playback for two reasons. First, the eye-tracking stimulus for every participant was the same, which facilitates the comparison participants. Second, the "time on task" remains the same for each participant.

40 university students from École Polytechnique Fédérale de Lausanne, Switzerland participated in the experiment. The only criterion of selecting the participant was the fact that each participant took the Java course in the previous semester, since this was a pre-requisite to understand the selected video. Upon their arrival in the experiment site the participants signed a consent form, then they answered three self-report questionnaires for a 20-point study processes (Biggs et. al., 2001), 10-point openness scale and 10-point conscientiousness scale. Then they took a programming pretest in Java. Then, they watched two videos from the MOOC course and after each video they answered programming questions based on what they were taught in the videos. In the following subsections, we describe different variables related to the present analysis.

The videos were in English: most participants were not English native speakers but are used to be taught in English. The first video included 13 slides and was 11 minutes 52 seconds long. The second video included 8 slides and was 10 minutes 7 seconds long. The content of slides was usually displayed in 3 to 4 steps. This increase conceptual with me ness since, at step 3 for instance student could look at objects that appeared on the slides in steps 1 and 2, but not yet at the object that will appear in step 4 and 5. The content of the slides was mostly text: sentences, lines of code and mathematical formulas. The content was mostly in black and white. Globally, slides were very light, with an average of only 55 words per slide.

Expertise and performance levels are given according the participants' scores from the pretest and the posttest respectively.

Results

Controlled Variables

We observe no significant relation between the three variables. There is no significant relation between pretest score and posttest score. There is no significant relation between pretest score and learning strategy: the student with good pre-requisites, are not necessarily deep learners. More surprisingly, there is no significant relation between learning strategy and posttest score: deep learners do not learn more. This may certainly feed the debate on deep versus surface learning. However, our goal was not to enter into this debate but simply to control that there is not another variable, pre-requisite level or learning strategy, which would interfere massively with our key variables, perceptual and conceptual with-me-ness.

Pretest Score and With-Me-Ness and Learning Strategy and With-me-ness

We did not observe any significant relation between pretest score and the two levels of with-me-ness. We also did not observe any significant relation between learning strategy and the two levels of with-me-ness.

Posttest Score and With-Me-Ness

We observed significant correlations for the two different levels of with-me-ness and the posttest score.

Entry Time: We observe no correlation between entry time and the posttest score (Spearman's correlation = 0.1, $p > 0.5$, Figure 2(b)). This can be explained using the saliency of the teacher's pointer. When a moving object appears on the screen, it constitutes a salient visual feature to which gaze is always attracted. This attraction does not reflect a deeper cognitive process and this is probably why it is not predictive to learning.

First Fixation Duration: We observe a significant correlation between the posttest score and the time spent for the first time the student look at the referred site (Spearman's correlation = 0.35, $p < .05$, Figure 3(a)). The students who scores high in the posttest were paying more attention to the teacher's pointers. This behavior is indicative of more attention during the moments of deictic references.

Number of revisits: We observe a significant correlation between the posttest score and the number of times the student look at the referred site (Spearman's correlation = 0.31, $p < .05$, Figure 3(b)). The students who scores high in the posttest came back to the referred sites more often than the students who scored less in the posttest. Having more revisits also resulted in having more fixations and thus more aggregated fixation duration as well. The revisiting behavior is indicative of rereading. Moreover, having more overall fixation duration on the referred sites is indicative of more reading time.

Conceptual with-me-ness: We observe a significant correlation between the posttest score and the time spent by the student following teachers' dialogues on the content of the slide (Spearman's correlation = 0.36, $p < .05$, Figure 2(a)). The students who scores high in the posttest were paying more attention to the teacher's dialogue. This behavior is indicative of more attention during the whole video lecture.

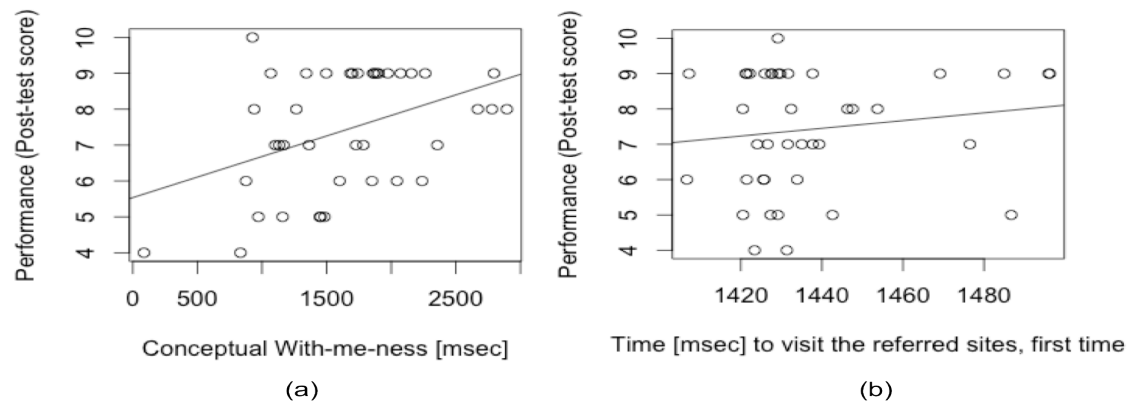


Figure 2. Posttest score and the With-me-ness. (a) Conceptual with-me-ness (x-axis) and posttest score (y-axis). (b) Entry time for perceptual with-me-ness (x-axis) and posttest score (y-axis).

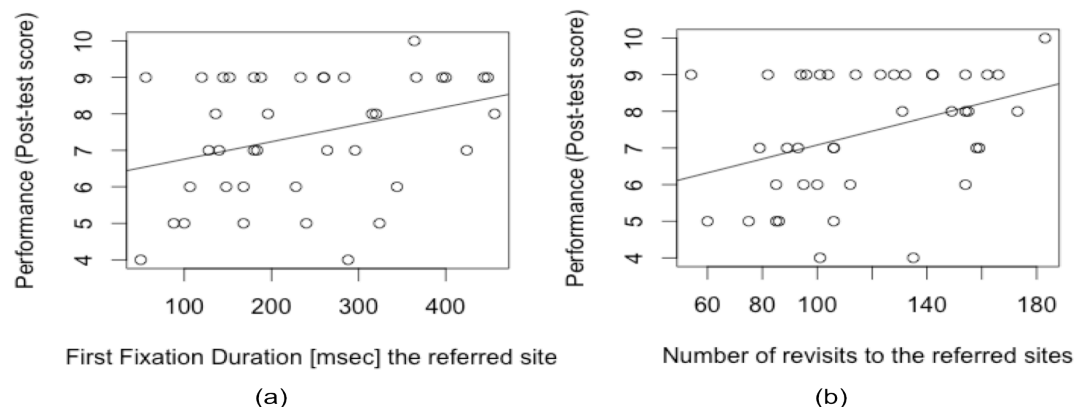


Figure 3. Posttest score and the With-me-ness. (a) First fixation duration for perceptual with-me-ness (x-axis) and posttest score (y-axis). (b) Revisits for perceptual with-me-ness (x-axis) and posttest score (y-axis).

Discussion

The entry-time component of the perceptual with-me-ness can be seen as the gaze behavior when there is a salient element present on the visual stimulus (Parkhurst et. al., 2002). The pointer of the teacher appears only a few times on the screen during the video lecture. We observe no correlation between the entry-time and the posttest scores. This can be explained by the fact that the pointer of teacher introduces a salient feature on the stimulus to which gaze is attracted. It does not reflect cognitive processing.

However, once the pointer is on the screen, the first fixation duration on the referred site is correlated with the posttest scores. The good-performers (those who scored high in the posttest) have more first fixation duration on the referred sites than the poor-performers. This is a typical situation during the moments of deictic references. Jermann and Nüssli (2012), in a pair-programming task, showed that better performing pairs have more concurrent gaze patterns during the moments of deictic references. Dale et. al. (2011), in listening comprehension task, showed that the pairs having more concurrent gaze during the period of references performed better than the other pairs.

The revisit component of the perceptual with-me-ness can be seen as rereading behavior. We observe a positive and significant correlation between the number of revisits to the referred sites and the posttest scores. The participants scoring high in the posttest have higher number of revisits to the referred sites. Mills and King (2001) showed in their studies that rereading improves the comprehension. In the present study, the scenario is somewhat different than Mills and King (2001). In the present study, the students did not reread the study material. Instead, the students referred back to the previously seen content again in the duration the slide was visible to them. Thus the relation between rereading of the same content and the performance should be taken cautiously, clearly further experimentation is needed to reach a causal conclusion.

The conceptual with-me-ness corresponds to a deeper form of attention, in terms of both the temporal scale and the cognitive effort “to be with the teacher”. We observe a positive and significant correlation between the conceptual with-me-ness and the posttest scores. The conceptual with-me-ness can be explained as a gaze-measure for the efforts of the student to sustain common ground within the teacher-student dyad. Dillenbourg and Traum (2006) and Richardson et. al. (2007) emphasized upon the importance of grounding gestures to sustain mutual understanding in collaborative problem solving scenarios. A video is not a dialogue; the learner

has to build common grounds, asymmetrically, with the teacher. The correlation we observed between conceptual with-me-ness and the posttest score seems to support this hypothesis.

Conclusions

We found interesting relationships between gaze patterns and indicators for performance i.e., posttest scores. Those who achieved high scores in the posttest had more with-me-ness. In other words, from a teacher's perspective, students who scored high in the posttest were "with the teacher" for longer period of time than those who scored low in the posttest. They have more perceptual as well as conceptual with-me-ness than the poor-performers. However, the results reported are only correlations between variables; hence there is no causality claimed in the present contribution. Our conclusion is that we have identified indicators that, on the one hand, can be captured by technology, and, on the other hand, related to learning performance.

The results also contribute towards our long-term goal of defining the student profiles based on their performance using the gaze data. The gaze-measure of students' attention can serve the purpose of a delayed feedback to the students based on their attention span. Perceptual with-me-ness can be used to give feedback to students about where they start lagging behind in the lecture. Moreover, revisits can be used to give feedback to students about their rereading behavior. The conceptual with-me-ness can be used to give feedback to the students about what they missed. Although, the results reported here are to be taken cautiously and certainly more experimentation are needed to find any causality.

References

- Astin, A. W. (1971). Predicting academic performance in college: Selectivity data for 2300 American colleges.
- Biggs, J., Kember, D., & Leung, D. Y. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, 71(1).
- Dale, R., Kirkham, N. Z., & Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Frontiers in psychology*, 2.
- Dillenbourg, Pierre, and David Traum. "Sharing solutions: Persistence and grounding in multimodal collaborative problem solving." *The Journal of the Learning Sciences* 15.1 (2006).
- Finn, J. D. (1989). Withdrawing from school. *Review of educational research*, 59(2).
- Grabe, M., & Latta, R. M. (1981). Cumulative achievement in a mastery instructional system: The impact of differences in resultant achievement motivation and persistence. *American Educational Research Journal*, 18(1).
- Holmes, R. M., Pellegrini, A. D., & Schmidt, S. L. (2006). The effects of different recess timing regimens on preschoolers' classroom attention. *Early Child Development and Care*, 176(7).
- Jermann, P., Nüssli, M. A., & Li, W. (2010). Using dual eye-tracking to unveil coordination and expertise in collaborative Tetris. In *Procs. of the 24th BCS Interaction Specialist Group Conference*.
- Jermann, P., & Nüssli, M. A. (2012). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. In *Procs. of the ACM 2012 conference on CSCW*. ACM.
- Juola, J. F., Bouwhuis, D. G., Cooper, E. E., & Warner, C. B. (1991). Control of attention around the fovea. *Journal of Experimental Psychology: Human Perception and Performance*, 17(1).
- Kaplan, R. M. (1982). Nader's raid on the testing industry: Is it in the best interest of the consumer?. *American Psychologist*, 37(1).
- Mayer, R.E., and Moreno, R., (2003) "Nine ways to reduce cognitive load in multimedia learning." *Educational psychologist* 38.1.
- Millis, K. K., & King, A. (2001). Rereading strategically: The influences of comprehension ability and a prior reading on the memory for expository text. *Reading Psychology*, 22(1).
- Paivio, A. (1991). "Dual coding theory: Retrospect and current status." *Canadian Journal of Psychology*.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision research*, 42(1).
- Prinzmetal, W., Presti, D. E., & Posner, M. I. (1986). Does attention affect visual feature integration? *Journal of Experimental Psychology: Human Perception and Performance*, 12(3), 361.
- Schwartz, D.L., and Hartman, K. (2007). "It is not television anymore: Designing digital video for learning and assessment." *Video research in the learning sciences*.
- Sharma, K., Jermann, P., Nüssli, M.-A., & Dillenbourg, P., (2011). Gaze Evidence of Different Activities of Program Understanding. In *Procs. of the 24th Psychology of Programming Interest Group Conference*.
- Sharma, K., Jermann, P., Nüssli, M. A., & Dillenbourg, P. (2013). Understanding Collaborative Program Comprehension: Interlacing Gaze and Dialogues. In *Procs. of 10th International conference on CSCL*.
- Wiley, D. E. (1976). Another hour, another day: Quantity of schooling, a potent path for policy. *Schooling and achievement in American society*.
- Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human perception and performance*, 16(1).

A Design Inquiry: Bridging Assessment and Curriculum Frameworks to Engage Students in Science Practices

Angela Haydel DeBarger, SRI International, angela.haydel@sri.com

Erika Tate, bluknowledge, erika@bluknowledge.com

Yves Beauvineau, Culturally Responsive Science Pedagogies, beauvineau.yves@gmail.com

Mingyu Feng and Patricia Schank, SRI International, mingyu.feng@sri.com, patricia.schank@sri.com

Tamara Heck and Michelle Williams, Michigan State University, heckt@msu.edu, mwilliam@msu.edu

Abstract: This reflective inquiry delves into the collaborative design process of a technology-based middle school genetics unit. With a lens on science practices, we describe how bridging assessment and curriculum design frameworks can inform decisions in an iterative co-design process. The paper explores the role of these design frameworks in articulating learning goals and outcomes, designing the learning experience to motivate student inquiry, and identifying features of technology tasks and tools to elicit artifacts and evidence. A pilot study with four teachers and 435 students revealed how to enhance elements of the design approach to strengthen the learning goals, improve learning experiences with visualizations, and streamline the flow of the curriculum to emphasize productive patterns of inquiry. Implications of this integrated approach for informing the design of learning environments, the scalability of designs, and teacher practice are discussed.

Introduction

Extensive research situates science education as a means for broadening participation in the practice and culture of science. Researchers have designed learning environments that present science as an important part of everyday life using relevant science dilemmas, engage learners in authentic scientific practices such as modeling and evidence-based explanations, and incorporate technology that helps students visualize scientific phenomena and monitor understanding (e.g., Bell, 2004; Lee, Linn, Varma & Liu, 2010; McNeil & Krajcik, 2011).

Despite the preponderance of evidence that suggests students learn science best through these rich experiences, school science has been reported to decontextualize science and require less rigorous learning performance. Limited resources, such as poor access to technology, too few professional development opportunities, and policies that endorse fragmented, low-level content standards, have challenged teachers' ability to offer coherent learning experiences for all students (Kali, Linn, & Roseman; 2008; Lee et al., 2010).

Well-designed, coherent materials have the potential to address the aforementioned challenges. With a lens on science practices, this paper describes how bridging assessment and curriculum design frameworks can inform key decisions in the iterative design process of a technology-based middle school genetics curriculum. We report emergent design knowledge aimed to help designers and researchers leverage resources in the (re)design, study, and scale of science learning environments that are accessible to schools and promote deep and meaningful science learning and teaching.

Theoretical Approach

Two frameworks guide our approach: Knowledge Integration and Evidence Centered-Design.

Knowledge Integration (KI)

Knowledge Integration (KI) offers a learning perspective and design framework to explain how students develop a deep understanding of science in everyday life and designed learning environments. KI recognizes that students maintain a repertoire of ideas about scientific topics, practices, and disciplines (Bransford, Brown, & Cocking, 2000; Linn, Davis, & Bell, 2004). As students learn through a variety of experiences in and outside school, they revise their repertoire by adding new ideas or changing the relation among new and existing ideas.

The KI design framework, comprised of *principles*, *processes* and *patterns*, guides the design of coherent science instruction, which provides students multiple opportunities to consider all their ideas, add scientifically sound new ideas, develop relations among these ideas, and promote an integrated understanding (Kali, Linn, & Roseman, 2008; Lee et al., 2010; Linn & Eylon, 2006). The *principles* (make science accessible, make thinking visible, help students learn from each other, and promote autonomous and lifelong learning) communicate the nature and quality of learning experiences that promote knowledge integration. The *process* (elicit ideas - add ideas - distinguish ideas - sort ideas) and *patterns* (e.g., predict, observe, create a model, explain) help designers coordinate learning activities that develop a deep understanding of science.

Evidence-Centered Design (ECD)

Evidence-centered design (ECD; Mislevy & Haertel, 2006) provides tools for creating valid assessments within learning environments. ECD involves an analysis of the substantive domain; the construction of an assessment argument; specification of tasks, rubrics and psychometric models; and the implementation and delivery of tasks within an operational assessment. A critical contribution of ECD is its provision of processes and structures to describe epistemic practices within disciplines. ECD facilitates the articulation of (1) learning goals (knowledge, practices, abilities and the integration thereof), (2) evidence produced in the form of actions by and among students, and (3) features of environments to elicit the desired evidence and learning goals. By making the underlying evidentiary argument for an assessment explicit, ECD facilitates coherence in assessment design. As technologies (e.g., visualizations and simulations) become further integrated in learning environments, ECD is critical for communicating decisions among the experts involved the design of these assessments.

Coordinating and Integrating KI and ECD

Our design approach aims to attend carefully to three dimensions: (1) learning goals that integrate core ideas, practices, and cross-cutting concepts within the discipline of science; (2) clear articulation of evidence, artifacts, and the ways students should engage to produce these; and (3) the features and flow of activity designs and participation structures. See Table 1. While KI prioritizes deepening students' understanding in a discipline through practice, ECD heightens awareness of assessment design within learning environments—how we know learning is taking place. Along these lines, KI and ECD function at different grain sizes in the design space.

Table 1: Mapping of KI and ECD Theoretical Approaches to Design Dimensions for Learning Environments.

	Design Dimensions		
	Learning Goals	Evidence, Artifacts and Engagement	Features and Flow of Activity and Task Designs
Knowledge Integration (KI)	Focus on connections between key ideas that promote lifelong learning	Learning as engagement through eliciting, adding, distinguishing, sorting ideas	Application of design patterns to promote knowledge integration within and across activities
Evidence-Centered Design (ECD)	Knowledge, practices, and abilities and combinations thereof that will be the target of assessments	Clear articulation of evidence produced by students to indicate progress toward and attainment of learning goals	Specification of activity and task design features to elicit desired evidence

We openly explored how KI and ECD approaches inform the design of a genetics unit, specifically addressing two questions: (1) What is the role of KI and ECD in articulating learning goals and outcomes? (2) How can KI and ECD help to define the learning experience and artifacts and elicit evidence of learning?

Project Context

The project context is the development and testing of a 5-week technology-based middle school genetics unit. Our goal is to develop students' understanding of inheritance through science practices of constructing models and explanations (see NRC, 2012). "*How can you use genetics to feed the world in 2052?*" is the driving question. The unit introduces a situation where the world will be running short of food and fossil fuels in 2052. Students inquire about how to selectively breed for more nutritious rice and higher endurance horses. Student pairs complete 10 activities with frequent opportunities for whole class discussion facilitated by the teacher using the open-source WISE 4.0 platform. Notable advantages of the WISE platform include tools such as *Idea Manager*, which supports sorting ideas and constructing explanations (McElhaney et al., 2012), and *WISE Draw*, a tool that students can use to draw models that illustrate a mechanism or process.

Initial Design Approach and Implementation

Our co-design process involved experts in curriculum design, assessment design, science content, science teaching and software design. We engaged teachers as co-designers, and these teachers served as mentor teachers to their colleagues during implementation of the unit. A key structure in facilitating conversation about flow was the unit template, which documents for each activity (1) learning goals addressed, (2) anticipated student problematic ideas, (3) learning experiences needed to address problematic ideas, (4) assessment opportunities, and (4) the KI design pattern sequence to motivate student inquiry in the activity.

Articulation of the Learning Goals and Outcomes

The design team referred to the *Framework for K-12 Science Education* (NRC, 2012) to identify disciplinary core ideas that comprise a deep understanding of inheritance for seventh grade students. The *Framework* also helped to unpack two science practices targeted by the unit: *Developing and Using Models* and *Constructing Explanations*. Here, we followed the ECD framework to specify abilities targeted for each practice as Knowledge Skills, and Abilities (KSAs; e.g., Ability to construct a causal explanation; Ability to construct a model and use the model to explain a phenomenon). The KI perspective helped establish 12 measurable learning goals that integrated content. The learning goals began with a broad statement about the core idea and science practice (e.g., *Students will be able to generate explanations that link the macro and micro structures/processes/functions relating to genetic expression*) with additional details about the content to be addressed (e.g., *An organism's characteristics can be expressed in different versions or variations called traits*).

Defining the Learning Experience, Artifacts and Evidence

KI prompted us to think about how to scaffold students to write integrated explanations of heredity and develop models. The design team identified WISE steps that engaged students in the knowledge integration process (i.e., elicit, add, distinguish, and sort ideas). For example, the Explanation Builder step guides students to *distinguish* which ideas help explain the genetic expression process. Using ECD, evidence for each of the practices was broadly specified in terms of *potential observations* we might expect to see if students are engaged in constructing models or explanations (e.g., application of science concepts to reason about the phenomenon).

Task Features and Flow

The KI design principles (e.g., make science visible) and patterns (e.g., *Predict-Observe-Explain* [POE]) informed an activity flow to scaffold coherent explanations of heredity. To *make the practices of scientists visible*, the activities require students to conduct selective breeding experiments using interactive visualizations. WISE Reflection Notes were then used to elicit students' predictions and post-observation explanations to *make students' scientific ideas visible*. ECD served to define specific design features of assessment opportunities as *characteristic features* or *variable features*. We considered which features needed to be present in tasks (e.g., All items that prompt for a scientific explanation will include data or evidence in stimulus materials) and documented how tasks might vary (e.g., the complexity of data/evidence).

Pilot Study

Four seventh grade science teachers implemented the unit in their classes for five weeks in Spring 2013; two teachers were from a school in a Midwest suburban district, and two were from a Southern suburban district. The Midwest school district student body is approximately 60% Caucasian, 18% African American, 9% Asian, 7% Hispanic, 5% Multi-racial and 1% American Indian and Pacific Islander. Twenty-five percent of the students are on free or reduced price lunch. The student body of the school district in the South is approximately 64% African American, 28% Hispanic, and 8% Caucasian. Sixty-one percent of the students are on free or reduced price lunch. 435 students from 19 classes participated in the study and worked in pairs. Prior to implementation, all teachers participated in a 2-day professional development workshop to discuss the unit learning goals, the activities and web-based tools, and the teacher interface to manage student data. Student work was logged by WISE system. The team also conducted regular classroom observations, and mentor teachers facilitated conversations with their colleagues during the implementation of the unit.

Sources of Data to Inform Revisions

Student work on assessments provided evidence of their thinking and reasoning. Memos from classroom observations and conversations with teachers provided insight into the student and teacher reactions to particular activities and factors that hinder implementation.

Findings

An analysis of student work revealed that on average, groups submitted four to five ideas while viewing visualizations and selected between three to seven ideas to support explanations. One fourth (24.7%) of the ideas were problematic and non-normative. For example, students struggled to link ideas about macro-level observed traits and micro-level cellular-level processes. Observations revealed that teachers' facilitation of prediction, modeling, and explanation steps was uneven. Some teachers engaged students in productive discourse around focal ideas in the unit (e.g., "Why do we want rice with high starch and horses with high endurance?"), but based on the observations, there was no evidence that these rich conversations occurred frequently. These findings demonstrated the need for additional support for teachers and students to promote knowledge integration. Revisions to the design approach are described below.

Revised Design Approach

Articulation of Focused Knowledge-In-Use Learning Goals and Evidence

We believed that students struggled to construct explanations and models that link ideas about macro-level observations and micro-level processes in part because our learning goals needed to be much more explicit in this regard. Using ECD, we refined the original 12 learning goals into five learning performances that more explicitly consider targeted applications of the science practices within genetics: (1) *Ability to use a model of trait expression to explain the observed traits in an organism*; (2) *Ability to construct a model of processes within a cell and use the model to explain how traits are expressed*; (3) *Ability to evaluate different models of trait expression*; (4) *Ability to construct a scientific explanation about how sexual reproduction can affect genetic variation*; and (5) *Ability to explain how sexual reproduction and expression results in trait variation*. For each learning performance, we crafted statements that clarify the evidence students would need to produce to demonstrate proficiency. See Table 2. The KI lens with ECD helped to outline the macro (e.g., identify observed trait) and micro (e.g., use the model to explain the mechanism of trait expression) elements.

Table 2: Example of Refined Learning Goal and Anticipated Evidence from Student Artifacts.

Learning Goal	Evidence in Student Artifacts
Ability to use a model of trait expression to explain the observed traits in an organism	<p>Appropriate use of visualization (model) of gene expression to explain why an organism has a particular trait. Macro and micro elements are connected in the explanation.</p> <ul style="list-style-type: none"> • <i>Macro</i>: Identification the observed characteristic and trait • <i>Micro-structural</i>: Identification of the location of the relevant component (s) of the cell • <i>Micro-functional</i>: Description of the function of relevant component (s) of the cell • <i>Micro-mechanistic</i>: Description of the mechanism of to explain trait expression

Designing a More Consistent and Coherent Learning Experience

The rearticulated learning goals and evidence statements prompted the design team to bring practices to the forefront and further realize the design principle, *make the practices of scientists more visible*. We first made the alignment of activity steps to KI patterns more transparent. Activity steps now more explicitly map on to aspects of science practices. For example, for the *POE* pattern, a Brainstorm step is used to elicit students' predictions in conjunction with the Idea Manager Tools (Idea Basket and Explanation Builder) to support students in documenting observations and constructing explanations. Applying ECD, we refined features of the Idea Basket steps to promote more active and focused observations of visualizations, as shown in Figure 1. Explanation Builder question prompts elicit more directly explanations that require links between micro- and macro-level ideas (e.g., *Explain what happens inside a rice plant that makes it high nutrition and how Farmer Wilder can check that the rice is very nutritious.*) For some activities we also extended the basic *POE* pattern to incorporate modeling (*Predict-Observe-Model-Explain [POME]* pattern). In the *POME* pattern, students create a dynamic visual model using the WISE Draw tool.

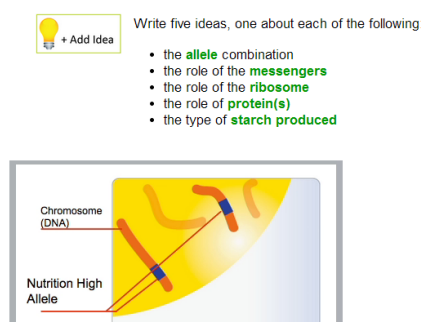


Figure 1. Screenshot of Expression Visualization with Idea Basket Prompts

Conclusions and Implications

Findings from this design study demonstrate that KI and ECD are complementary design perspectives. KI offers strong grounding around patterns to guide the flow of curricular experiences as well as a lens that focuses on designs promoting connected conceptual understanding in models/explanations. ECD bolsters design practices to support coherence in the rendering and application of tools to elicit intended knowledge and skills within a curriculum and facilitates the articulation of evidence to look for in student-generated artifacts.

The preliminary nature of our study limits our ability to generalize beyond the context of this unit. Yet, our early attention to emergent design knowledge positions us for a more rigorous investigation of student engagement and learning outcomes during a larger scale implementation study in the 2013-2014 school year. This paper lends insights to practical approaches for designing knowledge-in-use learning environments that engage students in important epistemological practices and incorporate valid assessments.

This approach also supports scalability and reuse of design principles. Generating design solutions not only involved conversations among the co-design team, but also clear documentation of decisions (e.g., new KI patterns and the unit template). These types of design documents afford the potential reuse of designs and scalability of assessments and tasks for related purposes. By explicating these decisions using shared schemas, other learning environments (e.g., game-based) that incorporate these science practices can subsequently be generated more easily by designers without having to retrace decision paths.

While teacher practice is not the focus of this paper, we highlight some implications of our approaches for teachers. We envision the curriculum and assessments as a starting point for discourse on constructing models and explanations. Thus, successful implementations of this unit and similar learning environments require supports for establishing norms to promote equitable participation (e.g., Hudicourt-Barnes, 2003), eliciting student reasoning and promoting discussion (e.g., Penuel, Beauvineau, DeBarger, Moorthy, & Allison, 2012), and using assessments to provide feedback to students (e.g., Ruiz-Primo & Furtak, 2007).

References

- Bell, P. (2004). The educational opportunities of contemporary controversies in science. In M. Linn, E. Davis & P. Bell (Eds.), *Internet environments for science education* (pp. 233-260). Mahwah: Lawrence Erlbaum Associates.
- Bransford, J., A. Brown, & R. Cocking. (2000). *How people learn: Brain, mind, experience, and school*. Washington: National Academy Press.
- Hudicourt-Barnes, J. (2003). The use of argumentation in Haitian Creole science classrooms. *Harvard Educational Review*, 73(1), 73-93.
- Kali, Y., Linn, M., & Roseman, J. E. (2008). *Designing coherent science education*. New York: Teachers College Press.
- Lee, H. - S., Linn, M. C., Varma, K., & Liu, O. L. (2010). How do technology-enhanced inquiry science units impact classroom learning? *Journal of Research in Science Teaching*, 47(1), 71-90.
- Linn, M. C. & Eylon, B-S. (2006). Science education: Integrating views of learning and instruction. In P.A. Alexander and P. H. Winne (Eds.), *Handbook of Educational Psychology* (2nd Ed., pp. 511-544). Mahwah, NJ: Lawrence Erlbaum Associates.
- McElhaney, K., Matuk, C., Miller, D., & Linn, M. (2012). Using the idea manager to promote coherent understanding of inquiry investigations. In van Aalst, J., Thompson, K., Jacobson, M. J., & Reimann, P. (Eds.) (2012). *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012) – Volume 1, Full papers*. International Society of the Learning Sciences: Sydney, NSW, AUSTRALIA.
- McNeill, K. L., & Krajcik, J. S. (2011). *Supporting grade 5-8 students in constructing explanations in science: The claim, evidence, and reasoning framework for talk and writing*. New Jersey: Pearson.
- Mislevy, R. J., & Haertel, G. D. (2006). Implications for evidence-centered design for educational testing. *Educational Measurement: Issues and Practices*, 25, 6-20.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.
- Penuel, W. R., Beauvineau, Y., DeBarger, A. H., Moorthy, S., & Allison, K. (2012). Fostering teachers' use of talk moves to promote productive participation in scientific practices. 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) – Volume 2, short papers, symposia, and abstracts*. International Society of the Learning Sciences: Sydney, NSW, AUSTRALIA.
- 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.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant Number DRL-1119055. 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 National Science Foundation.

Design Principles for Motivating Learning with Digital Badges: Consideration of Contextual Factors of Recognition and Assessment

Cathy Tran, University of California, Irvine, ctran27@gmail.com
Katerina Schenke, University of California, Irvine, kschenke@uci.edu
Daniel T. Hickey, Indiana University, dthickey@indiana.edu

Abstract: There is broad interest in the use of open digital badges to enhance learners' motivation. These web-enabled tokens of learning and accomplishment have the potential to induce excitement and elicit powerful forms of engagement and learning. Researchers and developers, however, appear divided on the role of digital badges in motivating learners. Our paper addresses the skepticism and promise that surrounds badges and presents the design principles for motivating learning found among 30 digital badge projects and aligns them with research. In doing so, we consider how contextual factors—such as how badges are used to recognize learning and how that learning is assessed—may play out to influence learner motivation in badge-oriented learning ecosystems.

Project Purpose

Traditional physical badges have been used for many years by organizations such as the Boy Scouts of America to acknowledge skills from archery to first aid. Now, so called “open digital badges” have become popular in a variety of learning environments. The MacArthur and Gates Foundations recently invested more than \$4 million to fund projects to design and implement digital badge systems. These newer badges are web-enabled tokens of accomplishment, skill, quality, or interest (Casilli & Knight, 2012). Unlike, grades, transcripts, or certificates, they can contain specific claims, detailed evidence supporting those claims, and links to additional claims and evidence; they can also be readily shared over social media & email and annotated & accumulated in standalone “backpacks.” While these features have obvious motivational potential, they are quite new. Little is known about how these features are being used individually or as part of larger educational ecosystems, and there has been little systematic consideration of actual or potential implications for motivation. We tackle one piece of the puzzle by focusing on how contextual factors such as the way badges are used to recognize and assess learning might influence motivation. We investigated the following research questions: (1) Which motivational design principles emerged from the specific practices we extracted from 30 projects? (2) What implications do recognition and assessment practices have on those motivation principles? (3) What is the likely motivational impact of recognition and assessment practices in a typical badges project?

Theoretical Framework

Researchers and developers appear divided on the role of digital badges in motivating learners. Reflecting longstanding concerns over extrinsic rewards, skeptics of badges “worry that students will focus on accumulating badges rather than making connections with the ideas and material associated with the badges” (Resnick, 2012). Badge evangelists find promise in having a new way to assess learners apart from the “current multiple-choice form of testing (that) doesn’t measure all that is being learned and de-motivates true curiosity” (Davidson, 2012). Our search for appropriate practices for motivating learning with badges is informed by well-known motivational theories (Dweck & Leggett, 1988; Eccles, 1983; Ryan & Deci, 2002). Our search is further informed by sociocultural views that consider motivation in the context in which it operates (Goodenow, 1992; Hickey, 1997, 2003). Sociocultural views consider motivation primarily in terms of the larger social and technological context and only secondarily in terms of individual differences that learners are presumed to bring to those contexts. Rather than embracing one side or the other on the enduring debate over incentives and learning, we instead documented the emerging practices for using badges to recognize and assess learning, as well as the deliberate ways projects were using badges in attempt to motivate. We then considered the potential positive and negative consequences of those practices and their interactions for the engagement (and potential disengagement).

Data Sources and Analytic Methods

Data came from awardees in the 2012 *Badges for Lifelong Learning* initiative. Thirty educational programs were funded to develop digital badge systems using the Open Badges Infrastructure developed by the Mozilla Foundation. Awardees range from after-school programs aligned with the Common Core State Standards (*Pathways for Lifelong Learning*) to teacher professional development programs (*Who Built America?*) and skill-based digital practice apps (*BuzzMath*). The research project set out to document the practices for using

digital badges for assessing, recognizing, motivating, and studying learning that emerged across these 30 projects. It did so by capturing “practical wisdom” as projects moved from intended practices (in their original proposal) to enacted practices (enacted in the actual badge system).

The research project organized the practical wisdom across programs as general design principles and project specific practices (The Design-Based Research Collective, 2003). The project assumed that local theories are built in the context in which they are intended to be used where insights can then be transferred to similar situations (Cobb et al., 2003). To identify the motivational practices in each project, we analyzed project proposals to identify the *intended* practices that system designers expected to motivate learning. In doing so, we documented and interpreted key design decisions related to learner motivation and design rationales for these decisions. A design rationale framework provides an account of the decisions teams make and the reasons for their decisions (Jarczyk, Loeffler, & Shipman, 1992; Lee & Lai, 1991). Employing this, our team asked project staff, through phone and in-person interviews, about design decisions they made to motivate learners to generate a list of intended practices. In addition, based on their grant proposals, we flagged other practices that may have unintended motivational consequences based on motivation research. After identifying the intended practices in each project, we categorized practices into larger design principles by dynamically sorting and re-sorting the practices into different groups of principles. These principles were vetted for feedback and revised both by representatives from the badge projects as well as attendees at the Digital Media and Learning and the Games, Learning, and Society conferences.

Results

RQ 1: Which Motivational Design Principles Emerged from Practices We Extracted from the 30 Projects?

Eleven overarching design principles with examples of practices for motivating learning are in Table 1. These principles aren't meant to be prescriptive—the process of designing a learning environment is not an exact science. Our goal is to provide perspectives and resources for educators and badge system developers to consider as they design badge ecosystems and figure out which badge design elements work best within their context to motivate learners.

RQ 2: What Implications Do Assessment and Recognition Practices Have on those Motivation Principles?

To illustrate the implications of assessment and recognition practices on motivation, we will focus on the recognition practice “Providing Privileges,” which was particularly prevalent across the projects. Privileges ranged from internship opportunities for youth who had earned particular badges to the receipt of a physical prize such as robots or entrance to a museum. The categories of providing privileges that emerged from our analyses were: 1) tangible prizes unrelated to the subject domain; 2) peer mentorship positions; 3) new related activities inside the program; and 4) access to outside internships.

The contingencies for receiving the badges as well as the types of reward the badges provide reflect different patterns of assessment and recognition that impact the motivational implications of providing privileges. The ways in which students receive those badges are assessment practices. The four categories of privileges constitute recognition practices because they illustrate ways in which badge achievements are acknowledged. By analyzing the different ways that privileges are associated with badges, we show how assessment and recognition practices have implications for motivation design principles. Below, we first outline the motivational principles and then discuss assessment and recognition principles.

Tangible prizes unrelated to subject domain. In some projects, learners are awarded physical prizes when they earn a badge. Students are recognized for their achievements by the receipt of these prizes. In such an environment, the assessment can therefore be viewed as accomplishing any means to receive those prizes. If rote memorization and repeating the easy activities is what boost points to attain badges for prizes, those may be the strategies that are likely to be employed by the learners.

Peer mentorship positions. As a privilege of collecting specific badges, some projects allow learners to be peer mentors. Research has suggested that teaching is motivating for learners because students feel in charge and are eager to help their tutees improve and as a byproduct put forth more effort and learn more themselves (Chase, Chin, Oppezzo, & Schwartz, 2009). This recognition of being a peer mentor is dependent on meeting assessment requirements such as collecting a series of leveled badges to demonstrate expertise in a domain to be allowed to assist peers at lower levels. Inherent in this recognition practice of providing peer mentorship privileges is also peer assessment which have different implications from computer and human expert (e.g., teacher) assessment.

New related activities inside the program. Giving learners access to new activities within the program is a prize awarded in several badge projects. In one example, learners are able to gain access to math contests

within the community and get exclusive access to “problem solving” missions. The recognized privilege and the assessment criteria are therefore linked to the same learning objectives, building on one another.

Table 1: Design principles for motivating learning

Principle name	Explanation	Example
Recognizing identities	Badges can recognize a learner’s role within the badging system such as recognizing their specialization in journalism, engineering, or peer mentoring. Badges can also recognize learner’s identities by being incorporated into projects that themselves target specific groups.	<i>S2R Medals</i> awards learners badges for their journalism and live reporting skills.
Engaging with communities	Some learners are able to earn badges for their involvement in their communities both at the physical and digital level.	<i>Planet Stewards</i> awards learners badges for engaging with their online community and acting as science communicator and collaborator.
Display badges to the public	Some projects give earners the option of displaying badges themselves, while other projects automatically display badges.	<i>Mouse Wins!</i> automatically displays learner's badges on their website.
Outside value of badges	Some projects integrate practices to give badges value outside of the badge system. These include having badges count as academic or course credit, showing badges to outside agencies, and/or documenting the link between the badges and real life applications.	Earners of <i>4H-USDA Robotics</i> badges have the opportunity to earn internships with partner institutions such as NASA.
Setting goals	Badges allow for learners to set goals and visualize the previous goals that they’ve accomplished. Badge systems can use goal setting in many different ways.	<i>BuzzMath</i> provides learners with clear learning pathways of the badges they have earned.
Collaboration	Some projects award group badges for group accomplishments as well as personal badges for having a role in a collaboration.	<i>Robotics and STEM Badges using NASA Content</i> awards badges to groups of learners.
Competition	Scarcity of badges and use of a point system are two ways that we have seen projects contribute to competition among badge earners.	<i>S2R Medals</i> limits the number of badges awarded to learners.
Recognizing different outcomes	The type of learning that a badge recognizes and the way that recognition is managed has profound implications for motivation.	<i>Design for America</i> awards learners badges for roles such as "peer mentor" and "project leader".
Utilizing different types of assessments	Projects utilize different types of assessments for learning such as computer, peer, expert, or self assessment.	<i>Sweet Water Aquapons</i> allows peers to award badges.
Providing privileges	Learners are awarded a variety of privileges in response to earning a badge such as prizes, the opportunity to take part in new activities, and access to internships.	Earning badges with <i>Design Exchange</i> allows learners access to internships.

Note. These principles will continue to evolve as projects move from intended to enacted practices.

Access to outside opportunities. Badge earners have the opportunity to serve as interns for some programs. Evidence from studies using an expectancy-value theory (Eccles, 1983) to understand learner choices would say that if the privileges that badges provided are interesting or exhibit high utility for learners, relative to cost, learners are more likely to initiate and sustain engagement with the subject matter. The assessment is therefore linked to the skills that these outside employers value and recognize so that they are willing to provide these internships.

One of the most salient points of these examples is the difficulty of making assumptions about the outcomes of providing privileges without understanding the context in which these privileges are used. The contingency of the privilege (assessment) and the privilege itself (recognition) are both contexts that matter. Some may undermine motivation while others may support the motivational impact of providing privileges. Incentives in themselves do not necessarily have positive or negative motivational influences; rather, it is studying them in the contexts in which they operate that provide valuable insight about their impact.

RQ3: What Is the Likely Motivational Impact of Assessment and Recognition Practices in a Typical Badges Project?

To answer our third research question, we turn to a case study of the Supporter to Reporter (S2R) Medals project. *S2R Medals* gives youths a glimpse at what it is like to be a journalist in the sports reporting world. More than 2,000 individuals have developed their reporting skills through S2R and reported at more than 1,000 events including the 2012 Olympic Games.

Assessment practices used by S2R Medals include leveled badges from bronze to gold that are aligned to standards that teachers are using in school. Leveled badges allowing learners to set goals for themselves and visualize those goals is an important strategy for self-regulated learning in which learners plan and monitor their learning (Zimmerman, 2000). S2R employs different types of assessments including computer scoring systems, peers, and experts. While computer assessment may benefit from being more efficient and free of social judgment, peer or expert assessment may be more meaningful and therefore increase the quality of work put into earning the badge. Relationships with peers are also implicated in the designer's intention of the badges to guide students along the path from novice to mentor, enabling advanced students to become a source of peer assessment for newer students. Such relationships within a learning community can help learners feel more connected and therefore persist within that learning environment (for an example, see Summers, Svinicki, Gorin, & Sullivan, 2002).

Recognition practices in S2R include allowing students to report at real sports events once they reach an appropriate level of expertise based on their accumulation of badges. Allowing students to report at real sports events illustrates the motivation design principle of providing privileges of access to outside opportunities. Privileges that badges provide that are interesting or exhibit high utility for learners, relative to cost, are more likely to initiate and sustain learner engagement with the subject matter (Eccles, 1983). As such, for those who are interested in journalism and find the project to be highly relevant for their goals are more likely to be positively influenced by these recognition practices. This case study illustrates the importance of acknowledging the interactive influences among assessment, recognition, and motivation within specific learning contexts.

Looking beyond the individual, the S2R case study also revealed ways that badges can help motivate *connected learning* (Ito et al., 2013). Much of the impetus behind the MacArthur Foundation's focus on digital badges follows from the assumption that they can help bring together and integrate spheres of knowledge, culture, and social practice that are normally very disconnected for most young people. Specifically, the case study of S2R medals uncovered the various ways that badges were used to recognize learning helped motivate connections between knowledge and abilities that were interest-driven, related to academic pursuits, and related to peer-culture and social networking. For example, examining one participant's S2R Medals home page displays the badges that individual earned, the number of peers who have "friended" him, and the various artifacts he has produced (1). Clicking on those badge reveals academic *and* professional competencies that the earner had to develop to earn the badge and the specific evidence of those competencies. Finally, the actual interest-driven (i.e., sports-related) artifact that the earner developed included familiar social networking features to make it easy for the earners' peers to "like" and post comments upon the artifact. Arguably, the digital badges motivated the kind of self-directed activity that Getzels and Csikszentmihalyi (1976) said was seldom possible in highly organized school activities but that was necessary for developing problem-finding skills and creativity.

Theoretical and Developmental Significance

Educational technologies are advancing at a much faster rate than research around those technologies; the recent surge of digital badges is no exception. As projects develop badge ecosystems, research on the development of these systems becomes increasingly important. In this paper we have outlined badge design principle that badge developers consider for motivating learners. Design principles need to be considered in the context of the

learning environment and individual differences among learners as well as in connection with the other principles that we have derived.

For developers of educational environments, we have offered badge design principles that they consider in educational settings. For motivation researchers, we have provided an analysis of the motivational impact of different badge designs. Our analysis illustrates how the same badge design may be motivationally adaptive in one situation but not the other. Understanding how badges look in practice as well as badge developers' initial instincts in designing badge systems is the next step in evaluating badging practices on learner motivation as well as in gaining insight on ways in which motivation theories extend to or are limited by new contexts in educational technology.

Endnotes

- (1) An example S2R homepage is at <https://www.makewav.es/r/glennwheeler>; one of the badges located on that home page can be viewed at <https://www.makewav.es/story/565038> and the story associated with that badges is at <https://www.makewav.es/story/569437/title/interviewwithhannahcockroftmbe>

References

- Casilli, C. & E Knight, E. (2012, June 11). 7 things you should know about badges [Web log post]. Retrieved from <http://www.educause.edu/library/resources/7-things-you-should-know-about-badges>
- Chase, C. C., Chin, D. B., Oppezzo, M. A., & Schwartz, D. L. (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of Science Education and Technology*, 18(4), 334-352.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32(1), 9-13.
- Davidson, C. (2012, February 21). Can badging be the zipcar of testing and assessment? [Web log post]. Retrieved from <http://dmlcentral.net/blog/cathy-davidson/can-badging-be-zipcar-testing-and-assessment>
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256.
- Eccles, J. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: Freeman.
- Getzels, J. W., & Csikszentmihalyi, M. (1976). *The creative vision: A longitudinal study of problem finding in art*. New York: Wiley.
- Goodenow, C. (1992). Strengthening the links between educational psychology and the study of social contexts. *Educational Psychologist*, 27, 177-196.
- Hickey, D. T. (1997). Motivation and contemporary socio-constructivist instructional perspectives. *Educational Psychologist*, 32(3), 175-193.
- Hickey, D. T. (2003). Engaged participation versus marginal nonparticipation: A stridently sociocultural approach to achievement motivation. *The Elementary School Journal*, 401-429.
- Ito, M., Gutiérrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K.,... & Watkins, S. C. (2013). *Connected learning: An agenda for research and design*. University of California Digital Media and Learning Research Hub. http://dmlhub.net/sites/default/files/ConnectedLearning_report.pdf
- Jarczyk, A., Loeffler, P., & Shipman, F. (1992). Design rationale for software engineering: A survey. *25th Hawaii International Conference on System Sciences Proceedings* (pp. 577-586).
- Lee, J., & Lai, K. Y. (1991). What's in design rationale? *Human-Computer Interaction*, 6(3-4), 251-280.
- Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28(1), 129.
- The Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5-8.
- Resnick, M. (2012, February 27). Still a badge skeptic [Web log post]. Retrieved from <http://hastac.org/blogs/mres/2012/02/27/still-badge-skeptic>
- Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 33). Rochester, NY: University of Rochester.

What's Happening in the "Quantified Self" Movement?

Victor R. Lee, Utah State University, 2830 Old Main Hill, Logan, UT 84322-2830, USA, victor.lee@usu.edu

Abstract: Rapid adoption of wearable tracking devices and motion sensitive apps has led to the development of the "Quantified Self" movement (QS). Some in the learning sciences community have begun to take notice and incorporate ideas from QS into the research and design of new learning environments. Yet the QS movement is still new enough that very little is known about it, and there are many open questions about how QS might be of value to the learning sciences. This paper provides some history of the movement and through a qualitative analysis of a public video corpus of QS presentations, identifies the variety of participants, the reported motivations driving individuals to self-quantify, and the data analysis activities of some individuals active in the movement. Opportunities for future research and design efforts in the learning sciences are also discussed.

Introduction

The reduced cost and increased availability of wearable devices and tracking apps has led to the development of what the popular press has called the "Quantified Self" (QS) movement. QS centrally involves extended tracking and analysis of personally relevant data. Given the emphasis on data and new technologies, QS is now starting to be discussed with research associated with the learning sciences. For example, QS tools and techniques appear in new frameworks for supporting learners in their own reflections on their learning processes (e.g., Rivera-Pelay, Zacarias, Müller & Brown, 2012), as a personal and consequential data source for input into educational games (Ching & Junicke, 2013) and in new designs for learning activities that turn everyday school experiences, such as recess or a walk to the library, into objects for students to study (e.g., Lee & Thomas, 2011; Lee & Drake, 2013).

These are all new developments. If QS is to continue to grow and overlap with work in the learning sciences, there is a fundamental question that should be asked of a movement that is still so new. Namely: What is happening in the Quantified Self movement? More specifically, who is involved in QS, what leads to them being involved, and what are some of the specific data collection and analysis activities undertaken by these "Quantified Selfers" (QSers)? Secondarily, what should the LS community seek to gain from understanding the QS movement? One goal of this paper is to identify a few features of the QS movement that we may wish to examine, leverage, or modify so that we can better understand and design for technology-supported learning that involves people having frequent encounters with data.

A Brief Background on QS: Origins, Growth, and Interactions

While there have been longstanding efforts to develop and promote wearable technologies, the past decade has seen a more aggressive effort among businesses to produce and distribute consumer-level wearables that emphasized physical activity tracking (Lee & DuMont, 2010). For example, in 2006, Apple announced an activity sensor and iPod kit that ultimately led to the Nike+iPod system. This incorporated a sensor embedded in the sole of a running shoe that could communicate wirelessly with an iPod. Around the same time (2007, to be exact), Kevin Kelley and Gary Wolf, two editors with *Wired* magazine who were credited with coining the name "quantified self", were meeting informally with a number of entrepreneurs and technology enthusiasts in California's Silicon Valley. Kelley and Wolf proceeded to report on new technologies and opportunities being explored through major periodicals (e.g., *The New York Times*) and presentations (e.g., TED Talks). As word about QS spread, informal local QS interest groups comprised of technology enthusiasts, hobbyists, and others interested in personal data began to coordinate face-to-face, regionally-based "meetups".

Rapidly growing popular interest in self-quantification led to the formation of a support organization, Quantified Self Labs (QS Labs). QS Labs launched a website (quantifiedself.com) with online guides and recommendations for how people in even remote locations could start their own QS meetup group. QS Labs also established an online discussion forum with over 2300 registered accounts, as of late 2013. The number of QS meetup group members expanded rapidly from one California-based group to 96 groups distributed across six continents. In early 2013, membership in QS meetup groups was already over 10,000 (Lee, 2013).

Face-to-face meetups typically occur every 1-2 months and feature volunteers giving ten-minute "Show & Tell" (S&T) talks, followed by a brief interactive discussion. These talks usually feature reporting of a personal self-data project followed by a personal reflection on what the presenter learned as a result. A number of these S&Ts are voluntarily recorded and uploaded to an online video repository, where they are then curated by QS Labs and posted publicly on the main quantifiedself.com page. In 2011, an international conference series was established. In its upcoming sixth iteration, the conference features keynote addresses related self-quantification, volunteers presenting refined versions of their regional S&Ts, and informal "breakout" sessions

for QSers to meet together in a separate room to discuss common interests and help push the QS movement further.

Theoretical Perspectives

While QS can be understood from a number of theoretical perspectives. I opted to view participation in QS as an instance of participation in an affinity space (Gee, 2005). An affinity space refers to a social affiliation organized around a common endeavor that involves variable degrees of discretionary participation and highly distributed, multidirectional exchanges of knowledge across various media. The canonical example of an affinity space is the social affiliation associated with a video game. Video games do not exist in isolation but rather are part of a set of practices that involve knowing and doing in particular ways with particular tools. As such, a number of ways exist for one to become an active participant in the affinity space associated with a video game that go beyond strictly playing the game. This can include posting in message boards, reading published strategy guides, or even getting verbal referrals and recommendations related to the game from knowledgeable peers. These pathways are designated as *portals*. In the case of QS, the *quantifiedself.com* website, the discussion forums, the meetup groups, public news media, and the new tracking devices and apps that one can purchase are all portals into QS.

Also important to the maintenance of an affinity space are *generators*. Generators are the entities that produce content and representations against which participants create meanings. In a video game affinity space, the game itself is a generator, as it has embedded in it many signs and representations that participants interpret and act upon. Portals can also be generators as well, as would be the case with a message board about a video game that mediates participants' understandings of what representations are meaningful in the game while also serving to generate new representations itself. Within QS, the S&T talks and the international conference are generators, as are discussion boards, and various other tools created and shared by QSers for producing novel data representations.

Self-quantification can take the form of a highly tailored set of data practices that build upon (and are also constrained by) the specifics associated with one's lives and circumstances (Azevedo, 2011). The canonical QS device may be something like a FitBit activity tracker or Zeo sleep monitor and involve tracking activity or sleep in the interest of improving health, but QSers also engage with self-data that collected in a range of ways from diverse activities. One goal of this paper is to illustrate some of that diversity.

Research Approach, Data Sources, and Analysis Process

In an ongoing effort to understand some of activities and participants in the QS affinity space, I have been engaged in roughly a year's worth of participant observation. This included interacting with QSers in a number of capacities, including attending meetups in my region, speaking on the phone or in person with personnel at QS Labs about their backgrounds and their current activities, attending the annual conference and talking with attendees and presenters, participating in QS discussion boards, and purchasing and using self tracking devices to examine data about my own activities. From these activities, I have obtained a number of artifacts for analysis, such as conference programs, field notes, message board exchanges, and transcriptions. For the current paper, I am drawing primarily the user-contributed online videos of S&R presentations (N = 220 and spanning from 10/2008-09/2013). This particular data source has inherent limitations and biases associated with how the videos were captured and which ones were ultimately selected for public dissemination (Hall, 2000). Still, it served as a useful and public source for understanding who participates in QS and what activities QSers pursue.

Demographic information such as gender, approximate age range, and ethnicity were identified based on appraisal of immediate information such as short video snippets, names of the participants, and comments posted with all the videos. A subset of 12 videos (a little over 5% of the entire set of videos) representing a range of topics and self-data projects was selected for transcription and further analysis. The transcripts of these videos were coded iteratively. The first round of codes served to flag all statements in the presentations that involved descriptions or motivators driving the collection or analysis of data. These codes were then refined by cross-case comparisons and through code frequency analysis. The videos were then re-reviewed and then re-coded. Following that step, codes were organized into major categories induced from this second review and analysis of the coding scheme. This process yielded a set of common features associated with the use of self-data and a set of specific examples to share, presented below.

Findings from Show and Tell Video Analysis

The age of the presenters appeared to be normally distributed, based on visual appraisal. There were clear majorities for specific genders and ethnicities. With respect to age, there were estimated to be 58 presenters (26.9%) in their mid-20s to mid-30s, 93 (43.1%) in their mid-30s to mid-40s, 49 (22.7%) in their mid-40s to mid-50s, 16 (7.4%) in their mid-50s to mid-60s, and 4 who were unclassified. For gender, there were 174 (82.5%) male presenters, 37 (17.5%) female presenters, and 9 joint presentations that had a male and female.

For ethnicity, there were 185 presenters (84.1%) who were categorized as White or of European descent, 24 (10.9%) Asian, 6 (2.8%) Hispanic or Latino, and 5 (2.2%) who were unidentified.

Through the coding process described above, I settled upon a coding scheme with 50 codes, from which 7 categories emerged. For this paper, I discuss three of those categories: *Motivations* for pursuing self-quantification, *Enabling conditions* that could build on initial motivation and supported the presenter's ability to pursue self-quantification and *Analyses of self-data* that had been collected.

Motivations

The primary identified *motivations* among the analyzed presentations included ties to ongoing personal interests (9 out of 12), a need articulated in part by a family member (3 out of 12), a concern for personal health (5 out of 12) and, a professional or commercial interest (3 out of 12). In addition, there were ways in which being a participant or attendee at QS gatherings helped to provide inspiration for doing some self-quantification. To illustrate, consider how Rajiv described at the beginning of his Show & Tell presentation (entitled "Papa, what should I read?") what led to his quantification of reading:

Rajiv: Ok so I love to read, I've read all my life...I have all these memories of leaving the library with armloads of books...A few years ago I started tracking my books, not for a project, but simply because I couldn't remember. My little girl is growing up and interested in reading my books, not just hers. She always asks me, "what can I read?" and I couldn't remember the name of the title or the author. So I started a simple spreadsheet of the name, the author, when I started, when I finished it, I gave it a star rating based on how much I liked it...A few years ago at a QS [meetup] a person gave a presentation on an information diet. It struck me, maybe there is something worth understanding from my tracking of books.

There are several features to note from this short excerpt. One is that Rajiv characterized reading as a lifelong pursuit tracing back to his youth, which suggests that while the quantification of reading was a relatively new endeavor, it was also built upon a longstanding preference (Azevedo, 2011) he had maintained throughout his life. His reading was done for his enjoyment but was not something he formally quantified until his daughter asked for recommendations on what to read, and he discovered that he could not recall what books he had read. That ultimately led him to produce a spreadsheet to keep track of what he had read and when. Later in his presentation, Rajiv discussed that he had kept these records for four years. He then saw another self-quantification project at a meetup, and was intrigued enough to see if he could find out anything new from his own records of his reading. This appropriation of how others work with self-data into his own project is an instance of knowledge being distributed and appropriated by individuals (Gee, 2005). Taken together, the convergence of personal interest, a need stated by a family member, and a model from a more experienced QSer all jointly supported the pursuit of his self-quantification project to examine his own reading patterns.

As a second example of convergence, consider how Jules explained his motivation to analyze snoring.

Jules: I am a snorer. This didn't really bother me because after all I couldn't hear my snoring. The same wasn't true for my wife who would let me know by means of an elbow in my ribs...So I went and took my phone and thought surely there is an app for that and looked and there were a few snoring apps, but there was nothing [good]... Hence, as I was looking at all these not very good apps, I could envision something much better which might be able to quantify things...to track your snoring through time and then be able to test if certain things affected your snoring...I've had loads of bad business ideas, and I thought, "hold on, maybe this is actually a good one!"...So I went to my wife and I said "Darling, darling I've got an amazing idea. You know all that money we've got saved up for a house deposit, why don't we employ a developer to build an app to measure snoring!"

In this example, Jules was both responding to a need to respond to the concerns of a family member (his wife, who was being awoken), and he also had hopes for developing a new app for a mobile phone as a business idea. In fact, one of the things Jules was doing through his S&T was advertising the app he had developed, "SnoreLab", while also describing his experiences with different snoring aids and how they affected his snoring. Also, as suggested by his past of having 'loads of bad business ideas' the decision to develop this app also seemed to flow from an ongoing preference for entrepreneurial pursuits. The convergence of this entrepreneurial streak along with an immediate need to deal with a snoring problem both help to explain what led to his engagement with the QS affinity space and why it took the form that it did.

Examples like those of Rajiv and Jules suggest that the motivations behind pursuing a self-data project are multi-faceted. It is not simply that someone abruptly decides to track some aspect of their life. Rather there can be several influences from many different spheres of one's life that leads to self-quantification. Among

QSers, it is well known that many people new to the space will excitedly initially try a tracking app or device for a few days but then cease using it. The exact reasons for this discontinuation has not been discussed heavily, but the experiences of those individuals who have gone beyond initial excitement with a tool and pursued a self-data tracking project suggest that motivation for self-quantification relies on the convergence of multiple personal and social forces, rather than general capabilities built into a tracking tool.

Enabling Conditions

QS participation is enabled by meetup groups and other physical and virtual *QS knowledge exchanges* (mentioned by 6 of 12), as well as the increased commercial availability of mobile and wearable technologies. The *ease of use* of a given tool or device or the *automatic data transfer* of data to another service were cited as being important for enabling self analysis (4 and 2 out of 12, respectively). Also of note was that half of the presenters highlighted how important *seamless integration* of data collection into daily activities was to their projects. While *ease of use* and *automatic data transfer* both helped in this regard, the high degree of integration was not an inherent feature of any given technology. Rather, *seamless integration* built upon *ease of use* or *automatic data transfer* along with some very small modification to existing routines. For example, in Matthew's comments during his presentation on how he lost 50 pounds by self-tracking illustrated how important he perceived that small modification could be:

Matthew: I picked up a Wi-Fi Withings scale and started stepping on it everyday. Everyday, before even saying "hi" to my wife in the morning, I just stepped on the scale... This [accompanying app data display] is a screen that I interact with a lot every day. Obviously the scale feeds into the weight...I can't say enough about the scale. It is the beginning of my day.

While Matthew ultimately tracked a number of other things, such as some of his exercise and caloric intake, the scale and his ritualized use of it was key. The perceived prominence of ritualization appeared in how he titled his entire presentation (i.e., "One small step on a scale"), and also when he explicitly stated at the end of his presentation how important it was for him to take that initial 'step' and look at that data as a form of feedback so that he could successfully lose that amount of weight. The way in which tracking technologies are typically promoted is in a way that suggests the user needs to do little to nothing to obtain their data. Ultimately, experiences like Matthew's helps show that QSers are interested and willing to do just a little bit of extra work to obtain data about their activities, but it the amount of work really has to be very little. It appears from this and related cases that for QS activities to take hold, the data collection technology needs to fit easily into existing routines of everyday life rather than demanding some routine in life adapts to fit with the technology.

Analyses of Self Data

While there was one individual who used a moderately advanced statistical technique (i.e., performing linear regressions), the approach to analysis as shown in the presentations most often involved *identifying changes over time* (7 out of 12), *finding central tendencies* (6 out of 12), and *comparing across conditions* (5 out of 12). Change over time was an important feature for presenters like Matthew who had a goal in mind and needed to see how well they progressed on their goal. This often tied into *finding central tendencies*, although central tendencies could also operate outside of an immediate goal related to the modification of some behavior. For example, Robby was interested in seeing how much time and money he spent on various forms of transportation. He logged a month of his transportation using his mobile phone with some hacks he developed and then determined average speeds and costs to facilitate *comparing across conditions* (including walking, car rental, bus, driving his own vehicle, taxi, and Zipcar car sharing).

In discussing his results, Robby had some suspicions confirmed, such as the relatively high average cost of a taxicab and also that cabs tended to be the fastest mode of transportation. However, he was surprised by how much it cost to use Zipcar and how slow it was, especially relative to other means of automotive transit he had used. To explain this, he proceeded to verbally reconstruct the situation that generated the data:

Robby: The least represented thing is the Zipcar...because it makes it look horrible with this data. I rented a Zipcar right when they had Hemp Fest downtown, so traffic was just awful and I got stuck in it. I was just trying to get away from it, and I ended up sitting in traffic for like 30 minutes and...I just took it back so I have a whole bunch of travel time that I was paying for. If you just look at Zipcar's rates by the time you are traveling it looks terrible.

Engaging in this kind of *data reconstruction* when looking at a data representation appeared in half of the presentations. This may ultimately speak to one of the unique and attractive features of participation in the QS affinity space: there is an inherently intimate relationship with the data that are collected because they are about 'the self'. As such, the ability to look at some of the resulting representations, tendencies, or changes over time

leverages familiarity. The patterns or trends that are identified are not always taken immediately at face value. Rather, they are subject to additional scrutiny because the history of the data are known and the meanings of the data can also be consequential for how people live their lives because the data already come from their lives.

Parting Thoughts: What might LS Take from a Snapshot of QS?

The work in this short report is preliminary, but as learning scientists work on issues of practice, design, participation, and learning, there are some aspects of the QS movement in its current form that could stir future research interest and pose new research questions of consequence to LS. These include:

- *QS models how technical and social infrastructure can leverage personal knowledge and interest related to the self to motivate learning.* Among the things that LS can uncover through study of QS practices and participation is how the emergent combination of infrastructure, physical and virtual knowledge support, and space for individuals to customize and share their knowledge pursuits all work together to mobilize intimate knowledge and curiosity about the self in service of learning. What combinations of support, technical infrastructure, and know-how are necessary for QS to work?
- *QSers encounter disciplinary content and participate in versions of valued disciplinary practices.* QS represents a unique, emerging hobby in which QSers identify trends over time, look for central tendencies, and compare differences across conditions. They touch upon mathematical, statistical, and scientific ideas such as variability, correlation, and experimental design. What is the nature of QSers understanding of this content? Is QS an effective model for how learners in other settings could meaningfully participate in scientific and mathematical inquiry practices?
- *QS currently lacks heterogeneity in its participants.* The bulk of the QS participants appear to be middle-aged white males of European descent. While LS is often recognized for its work in areas of design, cognition, and modeling of learning, issues of equity still figure prominently in the field. If QS is a promising model for designing learning experiences, could the model work with different focal populations (e.g., young women from underrepresented ethnic groups)? If not, what supports would be needed to enable equitable participation? Are the gains from doing self-quantification with populations different from that most commonly associated worth the investment?

References

- Azevedo, F. (2011). Lines of practice: A practice-centered theory of interest relationships. *Cognition & Instruction*, 29(2), 147-184.
- Ching, C. C., & Hunicke, R. (2013). *GETUP: Health Gaming for "the Rest of Your Life"*. Paper presented at Games, Learning & Society 9.0, Madison, WI.
- 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). Cambridge: Cambridge University Press.
- Hall, R. (2000). Videorecording as theory. In R. Lesh & R. Lehrer (Eds.), *Handbook of Research Design in Mathematics and Science Education* (pp. 647-664). Hillsdale, NJ: Lawrence Erlbaum.
- Lee, V. R. (2013). The Quantified Self (QS) movement and some emerging opportunities for the educational technology field. *Educational Technology*, 53(6), 39-42.
- Lee, V. R., & Drake, J. (2013). Quantified recess: Design of an activity for elementary students involving analyses of their own movement data. In J. P. Hourcade, E. A. Miller & A. Egeland (Eds.), *Proceedings of the 12th International Conference on Interaction Design and Children 2013* (pp. 273-276). New York, NY: ACM.
- Lee, V. R., & DuMont, M. (2010). An exploration into how physical activity data-recording devices could be used in computer-supported data investigations. *International Journal of Computers for Mathematical Learning*, 15(3), 167-189. doi: 10.1007/s10758-010-9172-8
- Lee, V. R., & Thomas, J. M. (2011). Integrating physical activity data technologies into elementary school classrooms. *Educational Technology Research and Development*, 59(6), 865-884. doi: 10.1007/s11423-011-9210-9
- Rivera-Pelayo, V., Zacharias, V., Müller, L., & Braun, S. (2012). Applying Quantified Self approaches to support reflective learning. In S. Dawson & C. Hathornwaite (Eds.), *LAK '12 Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 111-114). New York, NY: ACM.

Acknowledgments

Ralph Trumble, Kylie Williamson, Mary Briggs, and Michelle Berry all provided valuable support. Deborah Fields offered thoughtful comments on earlier drafts and championed 'affinity spaces'. This work was supported by NSF grant DRL-1054280. The opinions in this paper are those of the author and not necessarily of NSF.

The Nature of Student Thinking and Its Implications for the Use of Learning Progressions to Inform Classroom Instruction

Alicia Alonzo, Michigan State University, alonzo@msu.edu
Andrew Elby, University of Maryland, elby@umd.edu

Abstract: Underlying much of the work on learning progressions (LPs) is a strong though often tacit assumption that student thinking is theory-like and context-independent. In this work-in-progress, we use both theoretical perspectives on the nature of novices' knowledge and empirical evidence of the context-dependent variability of students' reasoning in physics to question this assumption and to argue that characterizing students in terms of LP "levels" inadequately captures their understanding of force and motion. We then analyze one teacher's use of LP-based data to reason about student thinking and instructional responses. While the teacher reasoned fluidly using LP levels, he more frequently used finer-grained knowledge elements and contextual factors to interpret student thinking; and these finer-grained interpretations led to more actionable instructional implications. Thus, while recognizing LPs as models (imperfect representations) of student thinking, we argue that their assumption of "levels" of student understanding may limit their utility for classroom decision-making.

Introduction

Learning progressions (LPs)—“descriptions of the successively more sophisticated ways of thinking about a topic” (National Research Council [NRC], 2007, p. 219)—have been touted as having the potential to influence educational policies and practices including standards, curriculum, instruction, and assessment (e.g., Black, Wilson, & Yao, 2011; National Assessment Governing Board, 2008; NRC, 2007, 2012). Although the NRC (2007, 2012) has tended to emphasize large-scale applications and LPs that span many years of instruction, LPs have also been promoted as tools to support teachers' instructional decision-making (Alonzo, 2011; Black et al., 2011; Furtak, 2012). Finer-grained LPs may be needed for this purpose (Alonzo & Gearhart, 2006; Gotwals, 2012). In this paper, we focus on the use of these finer-grained LPs to inform teachers' classroom work.

Underlying much of the work on LPs is the strong assumption that a student's thinking is reasonably consistent with a particular LP level (Alonzo, 2012). By providing generalized descriptions of students' thinking at each level and categorizing students as being “at” a level, LPs infer that, while some contexts may be more difficult than others, students apply the same thinking somewhat consistently. This assumption mirrors the claim that children's thinking is internally consistent and theory-like (e.g., Ionnides & Vosniadou, 2001). However, some researchers argue that students' prior knowledge consists partly of *misconceptions*, each with its own internal coherence, but that a student's misconceptions might not cohere into a theory-like structure (e.g., Carey, 1986; Clement, 1982). Still others argue that novices' intuitive knowledge is even more fragmented, consisting of abstractions from experience, such as *closer is stronger*, with different networks of these knowledge elements activated in different contexts (e.g., diSessa, 1993). Indeed, evidence suggests that student thinking in the “messy middle” (Gotwals & Songer, 2010, p. 277)—the conceptual territory between students' initial ideas and scientific concepts—may be particularly fragmented and context-dependent (Steedle & Shavelson, 2009).

Although some work on learning trajectories in mathematics has explored underlying assumptions about the nature of student thinking and learning (e.g., Battista, 2011), LP work in science has rarely questioned the assumption that students' thinking displays level-based consistency. However, this assumption is crucial for teachers' use of LPs. If teachers do not find levels-based information about students' understanding useful for classroom-level decision-making, LPs will not have their anticipated impact on instruction.

In this work-in-progress, we question the assumption that students' thinking displays theory-like consistency, and we explore implications of this assumption for teachers' decision-making. Relying on both theoretical work on the nature of novices' knowledge and empirical evidence of the context-dependent variability of students' reasoning, we argue that characterizing students in terms of LP “levels” inadequately captures their understanding of force and motion (FM). While recognizing LPs as models (imperfect representations) of student thinking, we use one teacher's interactions with LP-based score reports to argue that the LP model does not provide him with instructionally actionable information. Although the teacher was able to reason fluidly about LP-based information, he tended to use finer-grained knowledge elements and contextual factors to reason about students' responses to assessment items and to formulate instructional responses.

The Nature of Students' Thinking

In this section, we briefly review different models of novices' knowledge about the physical world. We then argue that most LP researchers at least tacitly assume theory-like consistency in student thinking at each “level,” although this assumption is problematic in light of recent empirical work.

Theoretical Perspectives on the Form of Students' Intuitive Knowledge in Physics

Some researchers assume that novices' knowledge consists of alternative or intuitive/naïve theories, similar in cognitive structure but different in substance from the theories held by experts (e.g., Ionnides & Vosniadou, 2001). For example, McCloskey (1983a, 1983b) has argued that many students initially hold a theory of motion similar to the one held by natural philosophers in the Middle Ages. Alternative theories about a topic such as FM, once ranked in order of increasing sophistication, map neatly onto the levels of a LP.

Other accounts of students' knowledge map less well onto LP levels. Some researchers characterize students' prior knowledge as consisting partly of misconceptions, such as *force is required for motion*, even motion at constant velocity (e.g., Carey, 1986; Clement, 1982). Each misconception is assumed to drive a student's reasoning about all or most situations in which the misconception is relevant, but a student's various misconceptions are not assumed to cohere into a theory-like structure. By this account, since a student's conceptions at a given moment could consist of many combinations of correct conceptions and misconceptions, it may be empirically inadequate to describe students' knowledge in terms of a small number of LP "levels."

Advocates of the knowledge-in-pieces (KiP) or "resources" perspective go a step farther from theory-like structures, by assuming that students' intuitive knowledge of physics consists largely of abstractions from experience, such as *closer is stronger* or *balancing* (Elby, 2000; diSessa, 1993; Hammer, 2000; Sherin, 2006). By this account, students' intuitive knowledge elements are neither correct nor incorrect, but are activated in context-dependent ways, sometime productively and sometimes not. For instance, students' intuitive sense of *balancing* can lead to correct conclusions about the forces at play in a tug-of-war between equally-matched teams, while leading to incorrect conclusions about the forces exerted upon a ball at its peak after being thrown straight up. By this account, the "misconceptions" documented in the literature often consist of networks of intuitive knowledge elements, which can be unstable in response to changes in context.

Empirical Evidence of Context-Dependent Variability in Students' Reasoning

Empirical evidence is consistent with a context-dependent view of student thinking. Finegold and Gorsky (1991) found that, while more than two-thirds of college and advanced high school students used one of 11 frameworks to describe forces acting on objects in motion, for objects at rest, "specific rules exist for specific situations: a force law for objects at rest on surfaces, for objects suspended from strings, etc." (p. 103). When they tried to identify frameworks that could account for both conditions, there were "almost as many models as there were students" (p. 109). Halloun and Hestenes (1995) found a smaller number of theories of motion (only three); however almost none of the college students they studied applied the same theory across different problem situations. Researchers working from a KiP perspective have also provided convincing evidence to support their accounts of student thinking (e.g., Clark, D'Angelo, & Schleigh, 2011). Thus, there is reason to question the extent to which student thinking can be considered to be theory-like or context-independent.

Learning Progression Perspective: "Levels" Assumes Theory-Like Consistency

Most LPs assume that students will use knowledge at a particular level to reason about phenomena across contexts. For example, a LP for carbon cycling (Mohan, Chen, & Anderson, 2009) expects students to reason similarly about *carbon generation* (plant growth), *carbon transformation* (e.g., human growth), and *carbon oxidation* (e.g., burning). Indeed, the assumption of consistency across contexts is required to diagnose a student as being at a particular level (e.g., Stevens, Delgado, & Krajcik, 2010). This assumption is present even in work that builds from a misconceptions perspective. Alonzo and Steedle (2009) describe how a LP for FM was developed in part by "grouping similar sets of ideas together into a single level" (p. 393) such that student thinking is treated as a predictable web of context-independent misconceptions.

The iterative design of LPs and associated assessments often entails evaluating the consistency of students' responses to assessment tasks. Thus, even when researchers do not explicitly describe students' thinking as context-independent and consistent, they nonetheless use a theory-like perspective as they seek reliable means of sorting students into context-independent, somewhat consistent "levels" of thinking (e.g., Alonzo & Steedle, 2009; Duncan, Rogat, & Yarden, 2009). Indeed, the consistency of students' reasoning at each level is central to the "conceptual coherence" criterion (Anderson, 2008, p.4) for validation of LPs. Although researchers acknowledge that particular contexts may be more or less difficult, most LP work assumes that student thinking is sufficiently context-independent and consistent to be characterized in terms of levels.

One Teacher's Reasoning about LP-Based Evidence of Student Thinking

Data Collection Context

Our work-in-progress uses data collected as part of a study of teachers' interpretations and use of LP-based formative assessment information. Seven physics teachers, all recommended as employing consistent and high-quality formative assessment practices, participated in a series of three interviews. Before the second interview, teachers read a brief description of the LP construct in general and a FM LP (Alonzo & Steedle, 2009) in

particular. During the second and third interviews, teachers interacted with score reports based on the FM LP and a set of 16 associated ordered multiple-choice (OMC; Briggs, Alonzo, Schwab, & Wilson, 2006) items. The score reports provided 1) LP diagnoses for each student and 2) item text and information about students' responses to each item. Teachers were asked to "think aloud" (Ericsson, 1993) as if the score reports described their own students. Afterwards, the interviewer asked follow-up questions, both for clarification and to probe teachers' reactions to particular features of the score reports.

We have begun analyzing "Tim's" second and third interviews. We chose Tim for this analysis because, from our prior three years of work with him, we knew him to be thoughtful and familiar with the idea of "learning progressions." Indeed, some of our work with him involved the FM LP. Although Tim represents a "best case" (strong formative assessment practices and familiarity with the LP construct), his interactions with the score reports displayed similar patterns to those of the other six teachers in the larger study.

Analysis

In this work-in-progress, we drew upon the theoretical perspectives described above to explore two approaches to characterizing Tim's interactions with the LP-based score reports. One researcher started by looking for evidence (and coded the transcripts in terms) of Tim's use of a theory-like, a misconceptions, and a KiP view of student thinking. The other researcher started with a narrative analysis of how Tim was reasoning about both student thinking and instructional responses. After this initial pass through the data, both researchers independently coded the transcripts using the two approaches. These codes were used as the basis for consensus conversations about the nature of Tim's diagnoses of students' thinking and associated instructional responses.

Results

Rather than holding one view of the nature of students' thinking, Tim expressed all three perspectives described above. In both interviews, he switched back and forth between thinking about students' ideas in terms of a) the provided LP levels and b) (mis)conceptions and even finer-grained knowledge elements and contextual factors that he thought affected students' reasoning about the assessment items.

When working with summary data that aggregated a student's or the class's responses to all items, Tim sometimes offered interpretations consistent with a LP perspective. For example, when interpreting summary data for an individual student, Tim used a LP level description to characterize the student's thinking:

It seems like he's fighting between [level] 2 or [level] 3, so I'd go down to [the description of level] two and double check that. [Reading the LP document] "Motion implies a force in the direction, non-motion implies no force. Conversely, student believes that force implies motion in the direction." So it seems like he's caught up with this idea between force being necessary for motion... you know, we can justify having... no net force when the object is standing still but not necessarily when the object is moving.

However, in both interviews, Tim focused primarily upon item-level data, using the item text and student response data to make finer-grained interpretations of student thinking. Consistent with the misconceptions perspective (that students' thinking is predictable and context-independent but not theory-like), he decomposed the LP levels into descriptions of more specific student ideas. For example, he questioned the scoring scheme for one of the items because two different misconceptions were mapped onto the same LP level:

Those two level 3 [option]s address different... issues. One, the rocket will move at a constant speed... the force is equal to the motion. The other one is saying... it will move until it reaches the maximum speed. There was another question about maximum speed...

Similarly, he differentiated between students' ideas about "maintaining motion" versus "speeding up motion," even though the LP levels do not differentiate between these two types of scenarios.

He also offered even finer-grained interpretations, looking to contextual features to provide him with important information about students' thinking. At times, he was explicit about the information provided by items set in different physical contexts. For example, after comparing two items, Tim commented:

So it's interesting that when [the object] is on a table, they... don't recognize that gravity is acting on it. But when it's in the air, they recognize that gravity's acting on it. So maybe just associating gravity with falling and only with falling, not as a force that's always there. So that gives me some information there.

The idea that gravity is a special force, eliciting different student reasoning than is elicited by similar items foregrounding other forces, was a prevalent theme across both interviews. Rather than considering students'

responses in terms of LP levels, Tim found the specific context of gravity to be a more relevant lens through which to view the assessment items and diagnose the students' difficulties: "I mean, it's a [Newton's] third law question. But it's got that gravity thing in it, and they struggle with gravity."

At the end of the third interview, Tim seemed to question the utility of the LP levels. After noticing that a student who provided level 3 responses to most of the questions had struggled with items involving gravity, he observed that "Kids have issues with gravity, and I'm noticing that there aren't any specific forces [in the LP]"; he wondered "how gravity fits into" the LP and "if it's a special case." While he highlighted the student's ideas about gravity as being central to the student's understanding of FM, the LP did not specifically mention gravity and, thus, did not allow Tim to identify the student's conceptual difficulties using LP levels.

LP researchers (e.g., Alonzo, 2011) have argued for the use of differences between levels, along with diagnoses of students' LP levels, to make instructional decisions. Tim demonstrated this use of LPs when reasoning about a summary of students' LP levels. After concluding that he would "focus on getting them out of the level 3 area," Tim read the descriptions of levels 3 and 4 of the FM LP and concluded that "this idea that force causes motion as opposed to force changing motion... would be where the focus of the class would be."

However, when reasoning about specific, context-dependent student ideas, Tim provided a much more detailed discussion of instructional implications. For example:

This idea that gravity is this holding force that doesn't let something move... just keeps popping up. So we would definitely have to talk about gravity... When we start talking about the normal force of interaction... they have a good intuitive sense of heavier things are affected more by friction. So I would try and base off of that idea... They're telling me they think that gravity, heavier things affect... friction, which is a good place to start.

In addition, in this example, Tim treated student ideas as sensible foundations on which to build. In contrast, when reasoning from the LP levels, Tim focused on the differences between students' ideas and correct scientific ideas but not on how students' ideas could play a role in bridging the gap.

Conclusions

Like Tim, we question the utility of LP levels for informing classroom practice. Theoretically and empirically, we have evidence that student thinking does not follow the neat patterns codified by LP levels. While recognizing LPs as models (imperfect representations) of student thinking, we argue that their assumption of context-independent, quite consistent "levels" of student thinking may limit their utility for classroom decision-making. We echo and empirically support Shavelson and Karpus' (2012) concern about the danger of cubbyholing student ideas into a LP model.

While not questioning the utility of multi-year LPs to inform standards and curriculum development, we argue that the finer-grained LPs intended to guide teachers' decision-making may still be too coarse-grained for this purpose. Previous work suggests that teachers do not always use LPs as intended. For instance, Furtak (2012) found that some teachers used a LP unproductively, to identify student misconceptions that need to be suppressed. This paper makes a different point. We found that Tim, at times, used the LP-based score reports exactly as the LP designers intended, attending to students' "levels" of thinking and how to transition students to the next higher level. We argue, however, that Tim's analysis of the LP-based score reports provided more actionable interpretations when he did *not* focus on the LP levels. Specifically, his finer-grained interpretations of student thinking, consistent with misconceptions and KiP perspectives, led to more specific, detailed instructional ideas. We present this work-in-progress as an existence proof, advancing the possibility that, while LPs may be useful in supporting view of student thinking that is more nuanced than "gets it"/"doesn't get it," they may not be well-suited for day-to-day instructional decision-making.

References

- Alonzo, A. C. (2011). Learning progressions that support formative assessment practices. *Measurement: Interdisciplinary Research and Perspectives*, 9, 124–129.
- Alonzo, A. C. (2012). Eliciting student responses relative to a learning progression: Assessment challenges. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 241–254). Rotterdam, The Netherlands: Sense Publishers.
- Alonzo, A. C., & Gearhart, M. (2006). Considering learning progressions from a classroom assessment perspective. *Measurement: Interdisciplinary Research and Perspectives*, 14, 99–104.
- Alonzo, A. C., & Steedle, J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education*, 93, 389–421.
- Anderson, C. W. (2008, February). *Conceptual and empirical validation of learning progressions: Response to "Learning progressions: Supporting instruction and formative assessment."* Retrieved from: http://www.cpre.org/ccii/images/stories/ccii_pdfs/learning%20progressions%20anderson.pdf

- Battista, M. T. (2011). Conceptualizations and issues related to learning progressions, learning trajectories, and levels of sophistication. *The Mathematics Enthusiast*, 8, 507-569.
- Black, P., Wilson, M., & Yao, S.-Y. (2011). Roadmaps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research and Perspectives*, 9, 71-123.
- Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment*, 11, 33-63.
- Carey, S. (1986). Cognitive science and science education. *American Psychologist*, 41, 1123-1130.
- Clark, D. B., D'Angelo, C. M., & Schleigh, S. P. (2011). Comparison of students' knowledge structure coherence and understanding of force in the Philippines, Turkey, China, Mexico, and the United States. *The Journal of the Learning Sciences*, 20, 207-261.
- Clement, J. (1982). Student preconceptions in introductory mechanics. *American Journal of Physics*, 50, 66.
- diSessa, A. A. (1993). Towards an epistemology of physics. *Cognition and Instruction*, 10(2-3), 105-225.
- Duncan, R. G., Rogat, A. D., & Yarden, A. (2009). A learning progression for deepening students' understandings of modern genetics across the 5th-10th grades. *Journal of Research in Science Teaching*, 46, 655-674.
- Elby, A. (2000). What students' learning of representations tells us about constructivism. *Journal of mathematical behavior*, 19, 481-502.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Finegold, M., & Gorksy, P. (1991). Students' concepts of force as applied to related physical systems: A search for consistency. *International Journal of Science Education*, 13, 97-113.
- Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, 49, 1181-1210.
- Gotwals, A. W. (2012). Learning progressions for multiple purposes: Challenges in using learning progressions. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 461-472). Rotterdam, The Netherlands: Sense Publishers.
- 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, 259-281.
- Halloun, I. A., & Hestenes, D. (1985). Common sense concepts about motion. *American Journal of Physics*, 53, 1056-1065.
- Hammer, D. (2000). Student resources for learning introductory physics. *American Journal of Physics, Physics Education Research Supplement*, 68(S1), S52-S59.
- Ioannides, C., & Vosniadou, S. (2001). The changing meanings of force: From coherence to fragmentation. *Cognitive Science Quarterly*, 2(1), 5-62. Retrieved from the University of Athens website: <http://www.cs.phs.uoa.gr/el/staff/vosniadou/force.pdf>.
- McCloskey, M. (1983a). Intuitive physics. *Scientific American*, 249, 122-130.
- McCloskey, M. (1983b). Naive theories of motion. In D. Gentner & A. Stevens (Eds.), *Mental Models* (pp. 299-324). Hillsdale, NJ: Lawrence Erlbaum.
- 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, 675-698.
- National Assessment Governing Board. (2008). *Science framework for the 2009 National Assessment of Educational Progress*. Retrieved from: <http://www.nagb.org/publications/frameworks/science-09.pdf>.
- National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: The National Academies Press.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts and core ideas*. Washington, DC: The National Academies Press.
- Shavelson, R. J., & Kurpius, A. (2012). Reflections on learning progressions. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 13-26). Rotterdam, The Netherlands: Sense Publishers.
- Steedle, J. T., & Shavelson, R. J. (2009). Supporting valid interpretations of learning progression level diagnoses. *Journal of Research in Science Teaching*, 46, 699-715.
- Sherin, B. (2006). Common sense clarified: The role of intuitive knowledge in physics problem solving. *Journal of Research in Science Teaching*, 43, 535-555.
- 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, 687-715.

Acknowledgments

Funding for this project was provided by NCS Pearson, Inc. We acknowledge the assistance of Elizabeth Xeng de los Santos, who conducted the interviews reported here.

Teaching about Confidence Intervals: How Instructors Connect Ideas Using Speech and Gesture

Elise Lockwood, Oregon State University, Department of Mathematics, elise314@gmail.com

Amelia Yeo and Noelle Crooks, University of Wisconsin–Madison, Department of Psychology,
ayeo@wisc.edu, nmcrooks@wisc.edu

Mitchell J. Nathan and Martha W. Alibali,
University of Wisconsin–Madison, Wisconsin Center for Education Research,
mnathan@wisc.edu, mwalibali@wisc.edu

Abstract: This work investigated how university-level statistics instructors use gestures along with speech to connect ideas when teaching about confidence intervals. We identified segments of classroom discourse in which instructors connected ideas, and coded the modalities they used to communicate those links. Most linked ideas were expressed multimodally, typically in speech and gesture.

Introduction

Statistical reasoning is a critical skill for many career paths in STEM fields, and it is also essential in many other aspects of life for citizens in modern society. As in any mathematical domain, one hallmark of conceptual understanding in statistics is understanding *connections* among ideas. Course work in statistics at the high school and college levels seeks to foster students' building a rich, interconnected knowledge base to support their reasoning and problem solving.

But, what factors affect whether students grasp connections among ideas? Recent research has targeted teachers' instructional communication as one important factor. In mathematics and statistics instruction, teachers frequently highlight relationships among inscriptions, concepts, and procedures. In doing so, teachers typically use speech, and along with that speech they often produce spontaneous *gestures*, defined as movements of the hands and arms that are part of speaking. Teachers' gestures contribute to instructional communication in a variety of ways; they express emphasis, guide students' attention to features of the instructional context, and convey information (Alibali, Nathan & Fujimori, 2011). Gestures are ubiquitous in instructional communication, and research on gesture in instructional settings, including both naturalistic and experimental settings, is burgeoning (e.g., Flevares & Perry, 2001; Goldin-Meadow, Kim & Singer, 1999; Rasmussen, Stephan & Allen, 2004; Richland, Zur & Holyoak, 2007; Roth, 2001).

In this study, we investigate how university-level statistics instructors use gestures along with speech to connect ideas in introductory statistics courses. In particular, we focus on instruction about *confidence intervals* (CIs). There have been many calls for an increased focus on confidence intervals, both in statistics instruction and in reporting of research findings (e.g., APA, 2009). However, CIs are difficult to understand, and statistics students often hold misconceptions about CIs (e.g., Garfield & Ben Zvi, 2008). The research presented in this paper highlights potentially effective ways to make conceptual connections about CIs using speech and gesture.

Students' Difficulties with Confidence Intervals

Undergraduate students perform poorly on assessments of CI understanding (e.g., Castro Sotos, Vanhoof, Van den Noortgate & Onghena, 2007). Additionally, knowledge of CIs does not improve significantly over the course of a typical introductory statistics course (e.g., Delmas, Ooms, & Chance, 2007). Difficulties with CIs persist even beyond college; researchers in many domains have trouble making inferences from CIs (Cumming, Williams, & Fidler, 2004), and often incorrectly interpret CIs (e.g., Coulson, Healey, Fidler & Cumming, 2010). One reason for these difficulties may be inadequate or incorrect links between CIs and other statistical concepts and representations (e.g., Cumming et al., 2004).

Teachers Use Gestures to Connect Ideas in Instructional Communication

In this work, we explore how statistics instructors connect ideas in lessons about CIs. We focus not only on the connections that instructors make verbally, but also on the connections they make using spontaneous gestures. A large body of research demonstrates that gestures contribute to communication (for reviews, see Hostetter, 2011; Kendon, 1994). Moreover, experimental studies suggest that students benefit more from lessons with gestures than from lessons without gestures (e.g., Church, Ayman-Nolley & Mahootian, 2001; Cook, Duffy & Fenn, 2013; Singer & Goldin-Meadow, 2005; Valenzano, Alibali & Klatzky, 2003). Students demonstrate better uptake of instructional information, more generation of new forms of reasoning, more generalization to new problem types, and greater retention of knowledge from lessons with gestures. Thus, teachers' gestures can have a substantial influence on students' learning. There is also evidence that teachers regularly express links

between ideas multi-modally. In a recent analysis of 18 middle school mathematics lessons, drawn from 6 teachers, it was reported that when teachers linked ideas, they expressed both (or all) of the linked ideas multi-modally in 90% of cases (range 65%-100%) (Alibali, et al., 2014).

Research Questions

The primary aim of this research was to characterize how statistics instructors connect ideas during instruction about CIs. We investigated whether instructors typically used multiple modalities (e.g., speech, gesture or writing/drawing) to express the linked ideas, or whether they sometimes expressed linked ideas in a single modality. In light of past research showing an abundance of multi-modal linking in mathematics lessons (Alibali, et al., 2014), we hypothesized that instructors would tend to express linked ideas multi-modally.

Note that we construed the notion of “linked ideas” broadly, to include inscriptions that are traditionally used in statistics (e.g., equations, graphs of distributions), and also to include verbal expressions and pictorial and gestural depictions of mathematical or statistical entities (e.g., a diagram drawn on the board or in the air). We focused not only on links across different external representations (e.g., between an equation and graph), but also on links within a single representation (e.g., links between two elements within the same graph).

A secondary aim of this research was to investigate the range of ways instructors used gestures to connect ideas. We focused on the *types* of gestures (in particular, pointing vs. depictive gestures). People often use gestures to indicate objects, inscriptions, or locations in the physical world via pointing. They also commonly use gestures to depict or represent actions, objects or events. We predicted that instructors would use both of these types of gestures in communicating key ideas in statistics instruction; for example, they might point to elements of a graph when referring to those elements, or they might represent the boundaries of a CI using depictive gestures, to evoke a mental image of an interval in their students’ minds.

In focusing on teacher’s communicative behaviors in statistics instruction, this work addresses authentic practices in a key domain of student learning. Past research suggests that teachers use gestures to help scaffold students’ understanding, particularly for material that is challenging for students (e.g., Alibali, et al., 2013). We suggest that better understanding of instructor practices in connecting ideas will yield insights into how best to organize such practices for optimal student learning.

Methods

We videotaped four university-level instructors as they taught about confidence intervals in introductory statistics courses. Three of the instructors were employed at a large public university in a mid-size Midwestern US city, and one instructor was employed at a mid-sized public university in a large Midwestern US city. The videotapes were transcribed and analyzed using Transana video analysis software. Although other parts of the lessons were viewed for context, for the purposes of this study, only segments of the lessons explicitly addressing CIs were analyzed. Since different instructors organized the lessons in different ways, the overall duration of the CI portions varied. We recorded the word count and length of each of these sections.

Identifying Linking Episodes

To analyze the data, one member of the research team first viewed the video segments and wrote narrative descriptions of what occurred. Next, this researcher and a second member of the research team went through the transcripts and video (using Transana) and identified *linking episodes*. Linking episodes were defined as segments of discourse in which instructors explicitly made connections between two or more different ideas or representations. These ideas or representations included representations of key statistical concepts, such as *confidence interval*, *sample mean*, and *population distribution*. A linking episode was a segment of discourse in which links between such statistical representations were set up or described. Within each linking episode, the researchers identified the *target link(s)*. A *target link* was the portion of the discourse in which the link between the ideas or representations was directly expressed. Some linking episodes contained more than one target link, either because the instructor explicitly stated the link multiple times, or because the instructor stated links among multiple ideas or representations in varying subsets (e.g., for a linking episode that connected ideas A, B, and C, the instructor might first state the target link between A and B, and then that between B and C, for a total of two target links).

The two researchers worked together to come to consensus on the portions of the lesson that were considered linking episodes and which segments within those episodes were considered target links. Three potential linking episodes were excluded from analysis because the researchers could not come to consensus, either that a particular episode contained a link or that a segment of a broader linking episode was a target link.

Identifying the Modalities in which Linked Ideas Were Expressed

We recorded the modalities instructors used to express each of the linked ideas in each target link. Each linked idea could be expressed in speech, gesture, writing or drawing, or in multiple modalities. Gestures were defined as movements of the hands or arms that co-occurred with speech (e.g., spreading both hands apart while saying

“the confidence interval is a range”). For target links that contained at least one idea expressed in gesture, we coded the gestures using an adaptation of McNeill’s (1992) gesture coding scheme. Gestures were classified as points, depictive gestures, or writing gestures. Points were defined as hand movements that indicated a referent, usually with the index finger. Depictive gestures were defined as gestures that represented semantic features of the referents via handshape or motion trajectory (e.g., holding both hands apart with palms facing one another to show the range of a CI). Writing gestures occurred when the instructor made markings on the board *while speaking* that either directed attention to particular piece of information (e.g., using chalk to mark a dot repeatedly on the symbol μ) or emphasized the relationship between two or more things that were on the board (e.g., drawing a line from the mean (μ) to the upper or lower limit of the CI). We included writing gestures as a separate category because these actions seemed similar to hand gestures that occurred with speech and qualitatively different from writing to put content (e.g., equations or figures) on the board.

Results

How Many Links Did Instructors Produce?

There was substantial variability across instructors, both in the amount of time they spent on CIs and in the number of links they produced during instruction on CIs. Time spent on CIs ranged from 16:54 (min:sec) to 37:59, with an average of 25:03. The number of links expressed ranged from 5 to 27 links, with an average of 15.5. Because different instructors spent different amounts of time on CIs, we also consider these findings in terms of the rate of links per 10 minutes of instruction. The instructors produced from 2.96 to 11.70 links per 10 minutes of instruction, with an average of 6.57. By definition, each link connected two or more ideas; some links incorporated as many as 5 distinct ideas. Averaged across instructors, the mean number of ideas expressed per link was 2.68 (range 2.5 – 3.0).

Did Instructors Express Most Links Multi-Modally?

We next examined whether instructors tended to use multiple modalities to express the linked ideas, or whether they sometimes expressed linked ideas in a single modality. Given past findings on multi-modal linking in middle-school teachers (Alibali, et al., 2014), we hypothesized that statistics instructors would tend to express linked ideas multi-modally. Table 1 presents the distribution of links for each instructor, classified according to the modalities in which each of the linked ideas was expressed. The upper portion of the table includes links in which at least one of the linked ideas was expressed in a single modality, and the lower portion includes links in which both (or all) of the linked ideas were expressed multi-modally. Overall, in 55% of links, *all* of the linked ideas were expressed multi-modally. Note that this value is much lower than the comparable value of 90% for the middle-school lessons (Alibali, et al., 2014). There was substantial variation across instructors in the proportion of links they expressed multi-modally, from a low of 40% to a high of 100% (mean = 62%).

Table 1: Target links produced by Instructors A through D, classified in terms of the modalities in which the linked ideas were expressed.

<i>Modalities in which linked ideas expressed</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Both/all in speech alone	0	2	0	0
All in either speech alone or gesture alone	1	0	0	0
Speech alone, gesture alone <i>and</i> speech+gesture	0	1	0	0
Speech alone <i>and</i> speech+gesture	11	6	0	2
Speech alone <i>and</i> speech+gesture <i>and</i> speech+writing <i>or</i> speech+gesture+writing	0	2	0	0
Gesture alone <i>and</i> speech+gesture <i>or</i> speech+gesture+writing	2	1	0	0
Total uni-modal links	14	12	0	2
Both/all in speech+writing	1	2	1	0
Both/all in speech+gesture	10	4	5	2
Speech+gesture <i>and</i> speech+writing	2	2	0	1
Speech+gesture <i>and</i> speech+gesture+writing	0	0	4	0
Total multi-modal links	13	8	10	3

What Types of Gestures Did Instructors Use to Express Linked Ideas?

Instructors used different types of gestures to express linked ideas. Depictive gestures were most common overall (58% of the 253 coded gestures), followed by points (32%) and writing gestures (10%). Instructors often

mixed gesture types within the same link; 45% of target links included gestures of multiple types, while 23% included pointing gestures only, 19% included depictive gestures only, and 3% included writing gestures only. The four instructors varied in their gesture patterns; one produced more than four times as many pointing gestures than depictive gestures, while the other instructors produced more depictive gestures than pointing gestures. These differences in gestural styles could be related to individual differences in the instructors' use of other visual representations in instruction.

As an example of a linking episode in which the linked ideas were expressed in depictive gestures, consider Figure 1, which presents a target link in which the instructor connected point estimates to interval estimates. The specific words that accompanied each individual gesture are marked in square brackets in the utterances quoted below. Just prior to this segment, the class had worked through an example in which the mean value was 21.73. As the instructor began the discussion of CIs, he explained that the point estimate was this mean, saying, "And so our *point estimate* in this case is going to be the mean of the sample. So that's where you're going to start with the confidence interval, giving that [point estimate]. That's your best guess." With this utterance, he put his fingers (of both hands) together in front of him, representing a point (Figure 1, first panel). He went on to note, however, that the mean is an overly precise estimate, and that another sample would not necessarily yield the same value (21.73) again. Then he introduced the idea of a CI, saying, "And so what we do with the confidence interval [is just to widen that], [make] a little bracket. And that becomes our confidence interval." With this utterance, he simultaneously extended both hands downward at 45-degree angles, ending by representing an interval "bracket" (Figure 1, second and third panels). Thus, in this example, the instructor linked point estimates and interval estimates using depictive gestures that expressed each concept, and that also expressed the notion of *widening*. This series of gestures could help students understand the quantitative relations between point and interval estimates, by presenting a visuospatial representation of the key ideas being linked. The visual representations the instructor expressed in gestures could clarify the verbal labels and support students' understanding of how point estimates and interval estimates are related.

This instructor returned repeatedly to the idea captured by these gestures throughout the lesson, and in fact, he repeated this basic set of gestures three additional times, producing what McNeill and Duncan (2000) termed a gestural *catchment* – a set of gestures in which some physical features of the gestures, or the entire gestures, are repeated. Catchments function to increase cohesion of the discourse—in, this case, perhaps, to foster a cohesive understanding of the relations between point estimates and interval estimates.

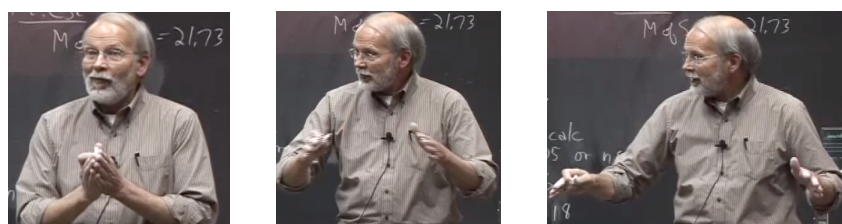


Figure 1. Target link in which both linked ideas (point estimate and interval estimate) were expressed in depictive gestures.

Discussion

Statistical knowledge is indispensable in many ways in modern society, from helping researchers in the learning sciences to conduct sound studies, to making culturally and politically responsible citizens. Among statistical topics, CIs are of central importance, but students often have misconceptions about them. This research bridges statistics education, embodied cognition, and gesture studies, and in so doing contributes to a broader perspective on teaching practices, and a specific focus on aspects of such practices that might affect student learning. Specifically, this work documents that instructors use gesture, as well as speech and writing, to connect ideas multi-modally in instruction on CIs.

Past research in other mathematical domains has suggested that teachers' gestures serve to *scaffold* students' understanding of central concepts in the lessons (e.g., Alibali et al., 2013). One reason gestures may support students' comprehension and learning of instructional material is because they offer an alternative, visuo-spatial format for expressing ideas. Statistics involves many visuo-spatial representations (such as number lines and graphs of distributions) so gesture is a potentially valuable communicative modality in this domain.

Although the statistics instructors in this study expressed a majority of links multi-modally, they used many fewer multi-modal links than the middle school teachers studied in previous work (55% of links for statistics instructors vs. 90% for middle school teachers) (Alibali, et al., 2014). Many factors might account for this difference. Younger students may benefit more from teachers' gestures than older students (Hostetter, 2011), and teachers may recognize this implicitly and adjust their use of gestures to match their students' needs. Alternatively, differences in content may account for some of the difference.

In sum, we have shown that gesture is pervasive in college-level instruction on CIs, and that it is used to convey important information about connections among ideas. This work paves the way for future studies that will test the implications of variations in instructors' linking for students' learning about CIs. It is our hope that ultimately, this line of inquiry will lead to empirically based recommendations regarding effective modes of communicating links among ideas in this important STEM domain.

References

- Alibali, M. W. & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from students' and teachers' gestures. *Journal of the Learning Sciences*, 21, 247-286.
- Alibali, M. W., Nathan, M. J., Church, R. B., Wolfgram, M. S., Kim, S., & Knuth, E. J. (2013). Gesture and speech in mathematics lessons: Forging common ground by resolving trouble spots. *ZDM – International Journal on Mathematics Education*, 45, 425-440.
- Alibali, M. W., Nathan, M. J., & Fujimori, Y. (2011). Gestures in the mathematics classroom: What's the point? In N. L. Stein & S. W. Raudenbush (Eds.), *Developmental cognitive science goes to school* (pp. 219-234). New York: Routledge, Taylor & Francis.
- Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Church, R. B., Johnson, C. V., Jacobs, S. A., & Knuth, E. J. (2014). How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction*, 32(1), 65-100.
- American Psychological Association (2009). *Publication Manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- Castro Sotos, A. E., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review*, 2, 98-113.
- Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism*, 7, 303-319.
- Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, 84(6), 1863-1871.
- Cumming, G., Williams, J., & Fidler, F. (2004). Replication and researchers' understanding of confidence intervals and standard error bars. *Understanding Statistics*, 3(4), 299-311.
- Coulson, M., Healey, M., Fidler, F., & Cumming, G. (2010). Confidence intervals permit, but do not guarantee, better inference than statistical significance testing. *Frontiers in Psychology*, 1, 1-9.
- Delmas, R., Garfield, J., Ooms, A., & Chance, B. (2007) Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 62(2), 28-58.
- Flevaris, L. M., & Perry, M. (2001). How many do you see? The use of nonspoken representations in first-grade mathematics lessons. *Journal of Educational Psychology*, 93, 330-345.
- Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Emeryville, CA: Key College.
- Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teachers' hands tell the student's mind about math. *Journal of Educational Psychology*, 91, 720-730.
- Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, 137(2), 297-315.
- Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, 27, 175-200.
- McNeill, D. (1992). *Hand and mind*. Chicago: University of Chicago 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.
- Rasmussen, C., Stephan, M., & Allen, K. (2004). Classroom mathematical practices and gesturing. *Journal of Mathematical Behavior*, 23, 301-324.
- Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science*, 316, 1128-1129.
- Roth, W-M. (2001). Gestures: Their role in teaching and learning. *Review of Educational Research*, 71, 365-392.
- Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science*, 16, 85-89.
- Valenzeno, L., Alibali, M. W., & Klatzky, R. L. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, 28, 187-204.

Acknowledgments

This research was funded by a grant #0909699 from the National Science Foundation. We thank Steve Jacobs and Breckie Church for their contributions to this research.

Studying Students' Early-Stage Software Design Practices

Brian A. Danielak, University of Wisconsin-Madison, danielak@wisc.edu
William E. J. Doane, Institute for Defense Analyses, wdoane@ida.org

Abstract: This paper describes research on undergraduate engineering students designing computer programs. We do three things. First, we explain what early-stage software design practices are and argue why we should look for students' productive capacities in them. Second, we outline the basics of a developing research program that aims at carefully documenting students' capacities for doing software design. Third, we relate early results from that research.

Introduction and Overview

In what follows, we do three things.

1. We outline a gap in existing Computer Science Education (hereafter CSEd) research: we lack a detailed understanding of how students design software. With a few notable exceptions (Harel & Papert, 1992; Papert, 1980; Peppler & Kafai, 2008; Turkle & Papert, 1990), most CSEd research does not view students as designers. Consequently, there is a paucity of research on how students understand the task of software design, what guides their moment-by-moment activity as they work on a design, and how the final form of code reflects their design knowledge. Emerging research shows what professional software design looks like in practice, but that only fleshes out the “expert” end of a novice-expert continuum. We're still left with an asymmetrical model of how software design expertise—as opposed to simply programming expertise, debugging practices, etc.—develops: we know far more about the target than we do possible origins or the paths between. That problem carries over to instruction. To answer questions of what students *should* do as a practice in software design (cf., Shaffer & Resnick, 1999), we need to first explore what they *can* do.
2. We propose *how* researchers can study students—specifically high school and university students—as *designers* of computer programs. Studying students as designers means looking not just at final code submissions but at what happens *between* compilations. It also means looking at students' early-stage activity: the interactions and artifacts that can emerge before a single line of source code is ever written.
3. We explain *why* pursuing this research matters. We underscore the importance of treating students *as designers* and show how such an orientation helps us make progress in both theory-building and instructional reform.

We Lack Accounts of Students' Software Design Practices

In 2010, the journal *Design Studies* devoted an entire issue to how professional software engineers design complex systems (Petre, van der Hoek, & Baker, 2010). That issue's editors argued fields of design studies, interaction analysis, and human-computer interaction don't know enough about how software engineers use representations and collaborative exchanges to organize the beginning phases of a design:

During formative design, software engineers spend a great deal of time engaging in creative, exploratory design thinking using pen and paper or a whiteboard—whether alone or in a small group. However, not enough is known about how software designers work in such settings. What do designers actually do during early software design? How do they communicate? What sorts of drawings do they create? What kinds of strategies do they apply in exploring the vast space of possible designs? (Petre et al., 2010, p. 533)

These questions were reprised in a 2012 issue of *IEEE Software* (Baker, van der Hoek, Ossher, & Petre, 2012), but we are still in the early stages of answering them.

Careful study of how experts design software is starting to reveal the complex disciplinary practices that comprise early-stage work (Ball, Onarheim, & Christensen, 2010; Jackson, 2010; Rooksby, 2010). For example, Rooksby and Ikeya (Rooksby & Ikeya, 2012; Rooksby, 2010) analyzed video recordings of three professional software engineer teams in the early stages of a software design task. Their research reveals that whiteboards—common artifacts in software design work—serve a much more complex purpose than simply to “log ideas and decisions” (Rooksby & Ikeya, 2012, p. 58). Moreover, the inscriptional content of the whiteboard is only a small part of design-work's complexity. Also crucial are the practices—such as conveying *epistemic uncertainty* about design solutions (Ball et al., 2010)—that help designers “remain coordinated and focused while collaborating” (Rooksby & Ikeya, 2012, p. 56). Ultimately, students need to learn as much about these

practices “as they do about procedures and technologies” when learning software design (Rooksby & Ikeya, 2012, p. 60)

The above results from software engineering echo findings from deep studies of engineers and scientists working and designing in practice. Practicing professionals make complex use of talk, gesture, and representational artifacts in physics (Gupta, Hammer, & Redish, 2010; Kaiser, 2005; Ochs, Gonzales, & Jacoby, 1996); chemistry (Stieff, 2007); field biology (Hall, Stevens, & Torralba, 2002); civil engineering (Stevens & Hall, 1998); structural engineering (Gainsburg, 2006); mechanical and electrical engineering (Bucciarelli, 1994; Henderson, 1999); and architectural design (Hall et al., 2002). Given that:

1. Science and engineering education research has made progress by looking for continuities in how novice learners develop disciplinary practices (Gupta et al., 2010; Hall & Stevens, 1995; Stevens & Hall, 1998), and
2. Emerging research on software engineering reveals that early-stage software design involves complex inscriptional, discursively, and epistemic practices,

it seems striking that there is no contemporary body of research, comparable to studies of expert practice, that looks at students’ software design practices. In other words, we have every reason to believe that expertise in software design involves complex practices, but little (if any) research on what productive capacities students have for them. Finding those productive design capacities requires a shift from questions such as *how can we assess and mitigate students’ difficulty in programming?* toward questions such as *how do students learn and display evidence of design thinking in programming?*

That difference is subtle, but it bears repeating. If we rephrase the quoted questions above (Petre et al., 2010, p. 533) and treat *students* as designers, we find the following questions that we think can drive a program of CSEd Research:

- What do students actually do during early stage software design work?
- What does students’ exploratory design thinking look like?
- How do students communicate?
- What sorts of drawings do students create when they design software?
- What kinds of strategies do students apply in exploring the vast space of possible software designs?

We think such research is both possible and potentially fruitful. We can do it using methods that already exist. And, it stands to greatly inform how we build theory about students’ capacities for design in programming.

Studying Students’ Design Practices Means Analyzing the Inscriptional, Gestural, Discursive, and Computational Artifacts They Create When Building Programs

The methods below form the core of our developing program to study students as software designers. None of the methods below are new; all have been used in prior educational research. What *is* new, we believe, is the opportunity to combine them all under the umbrella of understanding what happens when students design software.

We should collect students’ *code history*. By this, we mean we should preserve frequent snapshots in time of what students’ code looks like. Research has already shown code snapshotting to be a useful method for understanding large-scale patterns of student error (Jadud, 2006; Rodrigo, Tabanao, Lahoz, & Jadud, 2009; Spacco et al., 2006). And, the resolution of snapshots is extremely fine: Spacco et al. are able to capture the contents of a file each and every time a student saves it to disk. But, that research takes an aggregate view: it identifies large-scale error patterns at the expense of detailed naturalistic understandings of why students make those errors. Moreover, it’s primarily used to identify what *mistakes* students make, which is distinctly different from a research orientation that considers the negative *and* positive consequences of students’ design choices. A currently untapped advantage of collecting code history data, then, is that while we historically use it to aggregate *across programmers* it nevertheless also gives us fine-grained individual or team-based histories of how designs evolve.

We should conduct *clinical interviews* with students. Clinical interviews have proven historically useful in understanding the substance of students’ knowledge and the nature of conceptual change (diSessa & Sherin, 1998; Duckworth, 2006; Ginsburg & Oppen, 1988; Ginsburg, 1997). Crucially for CSEd Research, clinical interviews can tell us about students’ epistemologies—how they view knowledge and knowing in a discipline (Hofer & Pintrich, 1997)—which can affect how they approach and adopt that discipline’s practices (Hammer, 1989, 1994; Lising & Elby, 2005). Moreover, when interviews are videotaped and analyzed from perspectives of interaction analysis (Goodwin, 2000; Jordan & Henderson, 1995), interviews can offer rich evidence of the substance of students’ design practices

We should analyze students’ in-interview *inscriptions* when they design — what they write, how they write it, and how those writings get used. Evidence from both science studies (Hall et al., 2002; Henderson, 1999; Hutchins, 1995; Kaiser, 2005; Latour, 1990; Ochs et al., 1996) and educational research (Hall & Stevens, 1995; Lehrer, Schauble, Carpenter, & Penner, 2000; Stevens & Hall, 1998) highlights the centrality of

inscriptions to disciplinary practice in science and mathematics. Ethnomethodological data from studies of professional software engineers supports the same result: inscriptional practice is central to how professional engineers design software (Rooksby & Ikeya, 2012; Rooksby, 2010). Because part of that inscriptional environment is the computer itself, we should strive to capture and analyze what happens on-screen as students design programs.

We should pay close attention to students' in-interview *gestures*. Perspectives of gestural analysis hold that gestures can not only support or extend thinking, they can also communicate entirely different information than what's being said (Goldin-Meadow, 2003). Moreover, perspectives from cognitive anthropology and embodied cognition studies argue that bodily motion *is itself* cognition (Hall & Nemirovsky, 2012; Hutchins, 1995; Nemirovsky, Rasmussen, Sweeney, & Wawro, 2012). For example, when students describe a part of code by moving their hand across an imaginary row of items and tapping each item, their bodies convey information we can interpret about how they understand iteration. Whether or not one subscribes to the strongest tenets of embodied cognition, it still urges us to uncover the role of the physical environment in students' software design thinking.

Figure 1 presents a visual overview of some of these methods and modes of analysis. In particular, the first (top-left) panel depicts how we deploy these data collection methods during a clinical interview:

- a voice recorder captures speech (often a redundancy in case another recording device fails)
- a LiveScribe Pulse pen digitizes written inscriptions, allowing us to play back what was written in time
- a videocamera records data for knowledge analysis, interaction analysis, and gestural analysis,
- an in-interview computer tracks code history and screen-capturing software records real-time activity.

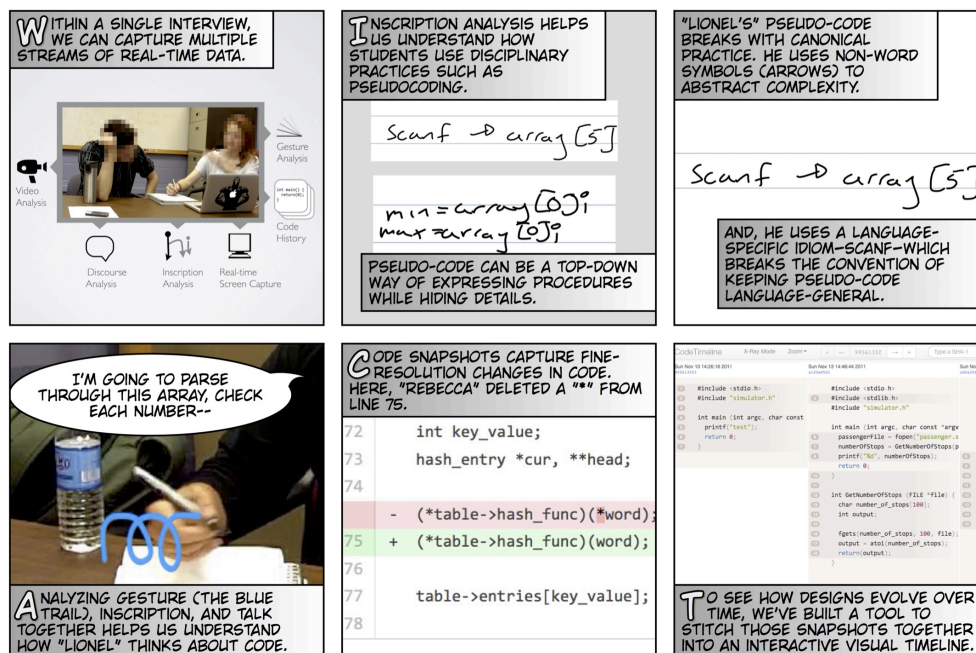


Figure 1. An overview of select strategies for capturing students' design practices.

Early Results Suggest Students View and Enact Disciplinary Practices with Remarkable Diversity

This paper reflects only one aspect of a larger study reported in Danielak (2014) to motivate a conversation between the learning sciences community and the computer science education research community. As an example of what this research orientation can provide, one strand of our research asks whether and how students use pseudocode in their design work. Pseudo-code is a high-level, abstract way of expressing procedures while backgrounding the details of how those procedures are carried out. It is also meant to be language-neutral, such that procedures expressed in pseudocode should be implementable in a variety of specific programming languages. We found that both the form pseudocode takes and the uses students put it to vary widely. One student, "Lionel," writes pseudocode as a prospective design tool. In Lionel's design process, writing pseudocode before writing language-specific program code is an absolutely essential part of convincing himself he understands the top-level view of what his program will do. Moreover, Lionel makes opportunistic use of language-specific idioms. Using such language-specific programming constructs would seem to clash with professional guidelines for writing pseudocode. Nevertheless, Lionel makes unproblematic use of his hybrid

syntax pseudocode to successfully solve problems. Another student, “Rebecca,” uses pseudocode sparingly. Rebecca writes pseudocode when she feels unsure of how to write a procedure properly in the language she’s using. For Rebecca, pseudocode becomes a fallback; a place-holder for the code she thinks belongs in a section of her program, but does not currently know how to write. Crucially, our data suggest when Rebecca uses pseudocode it seems to coincide with her feeling diffident about her programming ability. For Rebecca, pseudocode becomes a concession.

The contrast between how Lionel and Rebecca write and view pseudocode is striking. While Lionel uses pseudocode deliberately to plan and understand, Rebecca uses it hesitantly, viewing it as non-working code she must ultimately replace with working code. Even if Rebecca and Lionel both produce working programs as a design product, their trajectories look markedly different. Moreover, an instructor’s field of options in helping each student varies, not because of conceptual knowledge students have or don’t have, but because of differing design practices they’ve internalized.

Ultimately, attending to students’ design processes opens up new possibilities for research and practice. Sensitive research can help us formatively understand students’ productive capacities for design. We can conduct more powerful studies of design knowledge *in transition* (Smith, diSessa, & Roschelle, 1993). As argued, we need to understand the knowledge and experiences students already have for design if we aim to model the path toward expertise using constructivism. In a parallel way, responsive and informed instruction depends on a deep understanding of what students think they’re doing when they’re doing software design.

References

- Baker, A., van der Hoek, A., Ossher, H., & Petre, M. (2012). Guest Editors’ Introduction: Studying Professional Software Design. *Software, IEEE*, 29(1), 28–33. doi:10.1109/MS.2011.155
- Ball, L. J., Onarheim, B., & Christensen, B. T. (2010). Design requirements, epistemic uncertainty and solution development strategies in software design. *Design Studies*, 31(6), 567–589. doi:10.1016/j.destud.2010.09.003
- Bucciarelli, L. L. (1994). *Designing engineers*. Cambridge, Mass: MIT Press.
- Danielak, B. A. (2014). *How electrical engineering students design computer programs*. University of Maryland, College Park.
- diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education*, 20(10), 1155–1191. doi:10.1080/0950069980201002
- Duckworth, E. R. (2006). *“The having of wonderful ideas” and other essays on teaching and learning* (3rd ed.). New York: Teachers College Press.
- Gainsburg, J. (2006). The mathematical modeling of structural engineers. *Mathematical Thinking & Learning*, 8(1), 3–36. doi:10.1207/s15327833mtl0801_2
- Ginsburg, H. P. (1997). *Entering the child’s mind: the clinical interview in psychological research and practice*. Cambridge, MA; New York: Cambridge University Press.
- Ginsburg, H. P., & Oppen, S. (1988). *Piaget’s Theory of Intellectual Development* (3rd ed.). Englewood Cliffs, N.J: Prentice-Hall. Retrieved from <http://lccn.loc.gov/87017353>
- Goldin-Meadow, S. (2003). *Hearing gesture: how our hands help us think*. Cambridge, Mass: Belknap Press of Harvard University Press.
- Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32(10), 1489–1522. doi:10.1016/S0378-2166(99)00096-X
- Gupta, A., Hammer, D., & Redish, E. F. (2010). The Case for Dynamic Models of Learners’ Ontologies in Physics. *Journal of the Learning Sciences*, 19(3), 285. doi:10.1080/10508406.2010.491751
- Hall, R., & Nemirovsky, R. (2012). Introduction to the Special Issue: Modalities of Body Engagement in Mathematical Activity and Learning. *Journal of the Learning Sciences*, 21(2), 207–215. doi:10.1080/10508406.2011.611447
- Hall, R., & Stevens, R. (1995). Making space: A comparison of mathematical work in school and professional design practices. In S. L. Star (Ed.), *The cultures of computing* (pp. 118–145). Oxford, UK: Blackwell Publisher.
- Hall, R., Stevens, R., & Torralba, T. (2002). Disrupting representational infrastructure in conversations across disciplines. *Mind, Culture & Activity*, 9(3), 179–210.
- Hammer, D. (1989). Two approaches to learning physics. *The Physics Teacher*, 27(9), 664–670. doi:10.1119/1.2342910
- Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction*, 12(2), 151–183. doi:10.2307/3233679
- Harel, I., & Papert, S. (1992). Software design as a learning environment. *Learning to Design, Designing to Learn: Using Technology to Transform the Curriculum*, 35–70.
- Henderson, K. (1999). *On line and on paper: visual representations, visual culture, and computer graphics in design engineering*. Cambridge, Mass: MIT Press.

- 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. doi:10.3102/00346543067001088
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, Mass: MIT Press.
- Jackson, M. (2010). Representing structure in a software system design. *Design Studies*, 31(6), 545–566. doi:10.1016/j.destud.2010.09.002
- Jadud, M. C. (2006). Methods and tools for exploring novice compilation behaviour. In *Proceedings of the 2006 international workshop on Computing education research - ICER '06* (p. 73). Canterbury, United Kingdom. doi:10.1145/1151588.1151600
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4(1), 39. doi:10.1207/s15327809jls0401_2
- Kaiser, D. (2005). *Drawing theories apart: the dispersion of Feynman diagrams in postwar physics*. Chicago: University of Chicago Press.
- Latour, B. (1990). Drawing things together. In M. Lynch & S. Woolgar (Eds.), *Representation in Scientific Practice* (1st MIT Press ed., pp. 19–68). Cambridge, Mass: MIT Press.
- Lehrer, R., Schauble, L., Carpenter, S., & Penner, D. (2000). The interrelated development of inscriptions and conceptual understanding. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing and Communicating in Mathematics Classrooms: Perspectives on Discourse, Tools, and Instructional Design* (pp. 325–360). Mahwah, N.J: Lawrence Erlbaum Associates.
- Lising, L., & Elby, A. (2005). The impact of epistemology on learning: A case study from introductory physics. *American Journal of Physics*, 73(4), 372. doi:10.1119/1.1848115
- Nemirovsky, R., Rasmussen, C., Sweeney, G., & Wawro, M. (2012). When the Classroom Floor Becomes the Complex Plane: Addition and Multiplication as Ways of Bodily Navigation. *Journal of the Learning Sciences*, 21(2), 287–323. doi:10.1080/10508406.2011.611445
- Ochs, E., Gonzales, P., & Jacoby, S. (1996). “When I come down I’m in the domain state”: grammar and graphic representation in the interpretive activity of physicists. In *Interaction and Grammar* (pp. 328–369). Cambridge: Cambridge University Press.
- Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas*. New York: Basic Books. Retrieved from <http://lccn.loc.gov/79005200>
- Peppler, K. A., & Kafai, Y. B. (2008). Youth as media art designers: workshops for creative coding. In *Proceedings of the 7th international conference on Interaction design and children* (pp. 137–140). New York, NY, USA: ACM. doi:10.1145/1463689.1463740
- Petre, M., van der Hoek, A., & Baker, A. (2010). Editorial. *Design Studies*, 31(6), 533–544. doi:10.1016/j.destud.2010.09.001
- Rodrigo, M. M. T., Tabanao, E., Lahoz, M. B. ., & Jadud, M. C. (2009). Analyzing Online Protocols to Characterize Novice Java Programmers. *Philippine Journal of Science*, 138(2), 177–190.
- Rooksby, J. (2010). “Just try to do it at the whiteboard”: Researcher-participant interaction and issues of generalisation. In *Proceedings of the Studying Professional Software Design (SPSD) Conference*. San Diego, CA, USA.
- Rooksby, J., & Ikeya, N. (2012). Collaboration in Formative Design: Working Together at a Whiteboard. *IEEE Software*, 29(1), 56–60. doi:10.1109/MS.2011.123
- Shaffer, D. W., & Resnick, M. (1999). “Thick” authenticity: New media and authentic learning. *J. Interact. Learn. Res.*, 10(2), 195–215.
- Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *The Journal of the Learning Sciences*, 3(2), 115–163.
- Spacco, J., Hovemeyer, D., Pugh, W., Hollingsworth, J., Padua-Perez, N., & Emad, F. (2006). Experiences with Marmoset: Designing and Using an Advanced Submission and Testing System for Programming Courses. In *ITiCSE '06: Proceedings of the 11th annual conference on Innovation and technology in computer science education*. ACM Press.
- Stevens, R., & Hall, R. (1998). Disciplined perception: Learning to see in technoscience. In *Talking Mathematics in School: Studies of Teaching and Learning* (pp. 107–150). Cambridge, U.K: Cambridge University Press.
- Stieff, M. (2007). Mental rotation and diagrammatic reasoning in science. *Learning and Instruction*, 17(2), 219–234. doi:10.1016/j.learninstruc.2007.01.012
- Turkle, S., & Papert, S. (1990). Epistemological Pluralism: Styles and Voices within the Computer Culture. *Signs*, 16(1), 128–157.

Acknowledgments

This project is supported in part by grants from the National Science Foundation (NSF): DRL 0733613 and EEC 0835880. The views expressed are those of the authors and not necessarily those of the NSF.

Can Scaffolds from Pedagogical Agents Influence Effective Completion of Sub-Goals during Learning with a Multi-Agent Hypermedia-Learning Environment?

Michelle Taub, Roger Azevedo, Nicholas Mudrick and Erika Clodfelter, North Carolina State University
François Bouchet, Sorbonne Universités
mtaub@ncsu.edu, razeved@ncsu.edu, nvmudric@ncsu.edu, eeclodfe@ncsu.edu, francois.bouchet@lip6.fr

Abstract: Research on Self-Regulated Learning (SRL) in hypermedia-learning environments revealed that students are unable to engage in effective use of SRL, which have important implications for designing hypermedia-learning environments. 60 undergraduate students interacted with MetaTutor, an intelligent, multi-agent hypermedia-learning environment, with the overall goal of learning about the human circulatory system. There are two experimental conditions; prompt and feedback and control. We investigated if there were significant differences on the performance on and use of sub goals among the different experimental conditions. Results indicated significant differences between experimental conditions, such that participants in the prompt and feedback condition had higher sub goal quiz scores, spent more time engaging in each sub goal, had larger page relevant ratios, and had lower irrelevant page ratios, compared to participants in the control condition. These results suggest the importance of scaffolding in promoting the effective use of SRL strategies while interacting with a hypermedia-learning environment.

Introduction

Self-Regulated Learning (SRL) research has investigated students' use of Cognitive, Affective, Metacognitive, and Motivational (CAMP) processes as they interact with hypermedia-learning environments (e.g., Azevedo et al., 2013; Kinnebrew et al., 2013; Lester et al., 2013). According to Winne and Hadwin's (1998) model, learning occurs through different cyclical phases, during each of which information processing occurs. Unfortunately, students typically do not regulate their learning and therefore several researchers have incorporated Pedagogical Agents (PAs) to model and scaffold self-regulatory processes in Computer-Based Learning Environments (CBLEs). Research has investigated how effective the agents are in fostering effective learning (Veletsianos & Russell, 2014). Such research has examined agent aesthetics (Mayer & DaPra, 2012) and agent presence (Park & Catrambone, 2007), with respect to learning; and the influence of PAs on learning and motivation (Baylor, 2009), but has not investigated how Pedagogical Agents can influence students' engagement in self-regulated learning processes and activities, such as goal-setting, taking quizzes, and visiting content, which is relevant to the current sub-goal being worked on.

This study investigated how students completed their sub-goals in order to complete the learning session with MetaTutor, a hypermedia-learning environment. More specifically, we investigated if there were significant differences in students' performance on sub-goal quizzes, time spent on sub-goals, and navigation to relevant and irrelevant pages to the particular sub-goal, among experimental conditions, which differ on the feedback provided from the PAs. We expected that: (1) participants in the prompt and feedback condition would obtain significantly higher sub-goal quiz scores than participants in the control condition; (2) participants in the prompt and feedback condition would spend significantly more time engaging in each of their sub goals than participants in the control condition; and (3) participants in the prompt and feedback condition would have significantly higher relevant page ratios and significantly lower irrelevant page ratios than participants in the control condition.

Methods

Participants

Sixty undergraduate students (60% female) from two large universities in North America participated in this study. Participant ages ranged from 18-38 years old; $M = 21.3$, $SD = 2.66$. They were compensated with \$10 per hour for their participation, up to a total of \$40.

Research Design

This study was conducted with an experimental design; participants were randomly assigned either to the prompt and feedback condition ($n = 29$) or control condition ($n = 31$). These conditions vary with respect to the role of the Pedagogical Agents (PAs) in the system. Briefly, in the prompt and feedback condition, students are provided with scaffolding from the PAs, who prompt students to engage in several cognitive (e.g., Taking Notes, Summarizing, and Prior Knowledge Activation) and metacognitive (e.g., Judgment of Learning, Feeling

of Knowing, and Content Evaluation) SRL strategies, and provide feedback on the use of these strategies. Thus, in this condition, participants were not expected to learn independently for they are provided with assistance from the PAs, who are there to guide them to engage in effective SRL, and furthermore to learn successfully about the human circulatory system. In the control condition, students were not provided with any prompts or scaffolds from the PAs. Students were, therefore, free to navigate the system as they choose to, and can engage in the SRL strategies, but will not be prompted to do so.

MetaTutor: A Multi-Agent Hypermedia-Learning Environment

MetaTutor is an intelligent, multi-agent hypermedia-learning environment that provides learners with an overall goal of learning everything they can about the human circulatory system (Azevedo et al., 2013). In this environment, there are 48-pages of text and diagrams, which students can navigate to as they complete their sub goals. The sub goals, which students can set for learning are: (1) Path of Blood Flow, (2) Heartbeat, (3) Heart Components, (4) Blood Vessels, (5) Blood Components, (6) Purposes of the Circulatory System, and (7) Malfunctions of the Circulatory System. Each student sets two sub-goals, and can choose to add more during the learning session. The experimental session lasted for two days. On day one, the session lasted thirty minutes, and involved the student completing several demographic self-report questionnaires, and the pre-test, which included 30 multiple-choice questions on the human circulatory system. On day two, students returned to the lab, where they were instructed to use the MetaTutor system and its hypermedia content for 90-minutes, followed by the completion of a post-test (which also included 30 multiple-choice question test on the circulatory system). During the 90-minute session, Gavin the Guide (one of the four PAs embedded in the environment) administered the AEQ self-report emotions questionnaire (e.g., Pekrun et al., 2011), in order to assess how the student was currently feeling at that point during the session. Gavin administered this measure every 14 minutes, for a total of six times throughout the learning session. There are seven sub goals, which students can set at the beginning of the learning session.

There are four Pedagogical Agents (PAs) embedded within the system, who represent the different facets of SRL. *Gavin the Guide* introduces the learner to the session, and administers the pre-test, post-test, demographic questionnaires and emotion questionnaires throughout the session. *Pam the Planner* assists in planning, and helps students set sub goals to work through during the session. *Sam the Strategizer* assists in the use of learning strategies, and encourages students to create summaries on the content. *Mary the Monitor* focuses on monitoring strategies, and administers the sub goal quiz when a student has completed working on his or her current sub goal.

The screenshot displays the MetaTutor (version 1.2.8) interface. At the top, a 'Time Left' indicator shows 42:53. The 'Learning Goal and Subgoals' section states the goal is to learn about the Circulatory System. Below this, 'Your current subgoals are' lists 'Path of blood flow' and 'Blood components'. A 'Table of Contents' on the left lists various topics, with 'Metabolism' highlighted. The main content area, titled 'Other aspects of Circulatory System: Metabolism', provides a text description of metabolism and its importance in cellular respiration. To the right of the text is a diagram titled 'How our metabolism works' showing the flow of nutrients (Amino acids, Carbohydrates, Fatty acids) into the cell, their conversion to Acetyl-CoA, and their entry into the Citric Acid Cycle within the mitochondria. On the far right, a 'Learning Strategies' panel offers options like 'Tell you what I already know about this', 'Assess how well I understand this', 'Evaluate how well I already know this content', 'Evaluate how well this content matches my current subgoal', 'Take notes', 'Make an inference', and 'Summarize'.

Figure 1. Screenshot of the MetaTutor interface

The MetaTutor interface consists of multiple parts. See Figure 1 for an image of the MetaTutor interface. On the top left corner, there is a clock, which displays the time remaining in the session. Just below the timer, there is a table of contents, which displays the title of each page, and so students can choose which page to navigate to. On the bottom, there is a text box, where students can type in information, and can respond to the questions, which respond to the metacognitive SRL strategy questions. On the right hand side, there is the SRL palette, which lists the cognitive and metacognitive SRL processes, such as taking notes, content evaluation, and judgment of learning, which students can click on, if they choose to engage in one. The Pedagogical Agent is located in the top right corner. One agent is displayed at a time, and is displayed according to the scaffold and activity the student is engaging in. On the top of the interface, the student's overall learning goal, to learn everything about the human circulatory system, is listed, with a list of the sub goals set, which are placed underneath it. There is a status bar for each sub goal, which indicates how far along the student is in completing the sub goal; and students can click to complete the current sub goal if he or she feels it is complete.

Experimental Procedure

During learning with MetaTutor, we collected multi-channel data, which provides a wide array of product and process data related to cognitive and metacognitive SRL processes and learning activities. As students learn, we collected: (1) log-file data, which include time-stamped information, at the millisecond level, of every mouse click and keyboard entry the student makes into the system, which includes when the student engages in an SRL strategy, and when a PA intervenes; (2) eye-tracking (Tobii T60) data, which measures students' gaze behavior and fixations as he or she navigates through the system; (3) physiological data, which measured students' arousal based on electro-dermal activity to events (e.g., agent feedback) during learning; (4) video recordings, which captured the students' facial recording and allows for us to detect their emotional states during learning; and (5) video and audio recordings of the learners' concurrent think-aloud protocols and interactions with the system. For this study, we only analyzed data from the log-files.

Coding and Scoring

Several variables were extracted from the multi-channel data. However, in this study we only report on a subset of the variables extracted from the log-files. For example, mean sub-goal quiz scores and time spent engaging in the current sub goal were obtained from the log-files. Quiz scores ranged from 0-9; $M = 6.16$, $SD = 2.33$, and mean time spent on each sub-goal was calculated in seconds. A graduate student, with the assistance of two undergraduate students read through the printed log-files and identified the sub-goals participants set (e.g., Path of Blood Flow), how long students spent engaging in each sub goal, and which pages students navigated to as he or she completed the sub goal.

Sub-goals were coded into two groups, based on level of difficulty. Group 1 consisted of the less difficult sub goals: Path of Blood Flow, Heart Components, Blood Vessels, and Blood Components; and Group 2 consisted of the more difficult sub goals: Heartbeat, Purposes of the Circulatory System, and Malfunctions of the Circulatory System. Sub-goals were categorized into these categories based on the number of pages associated with the sub-goals, such that sub-goals with more relevant pages were deemed less difficult, and categorized into Group 1, whereas sub-goals with fewer relevant pages were deemed more difficult, and were categorized into Group 2. In order to determine page relevancy ratios, each page number was marked as relevant or irrelevant, based on an excel sheet containing information regarding each sub-goal, including which pages of content and which diagrams are relevant to each of the seven sub-goals, and so we were able to mark the pages relevant to each sub goal. Total numbers of relevant and irrelevant pages were calculated, which were then calculated as ratios using the following formulas: $[\text{Total Relevant Pages Viewed} : \text{Total Pages Viewed}]$ and $[\text{Total Irrelevant Pages Viewed} : \text{Total Pages Viewed}]$ for relevant page ratio and irrelevant page ratio, respectively. Participants set at least two sub goals each, and so although we had 60 participants, we obtained 183 data points, which represented the total amount of sub goals that were set among all 60 participants.

Results

In order to address the posited research questions, we ran several *t*-tests, in order to test for significant differences between the experimental conditions on their performance and use of the seven MetaTutor sub goals.

Research Question 1: Are there Significant Differences between Conditions on Sub Goal Quiz Scores?

An independent samples *t*-test was used to determine if there was a significant difference between experimental groups on their scores on Sub Goal quizzes. We extracted quiz scores from the log-files, which captured time-stamped information during the learning session.

Results from the t -test revealed that there were significant differences between sub goal quiz scores among experimental groups; $t(40.688) = 2.693, p = .01, \eta^2 = 0.11$. More specifically, students in the prompt and feedback condition obtained higher mean sub goal quiz scores ($M = 6.94, SD = 1.20$) than students in the control condition ($M = 5.43, SD = 2.87$). This result indicates that participants who were provided with prompts and feedback during the learning session performed greater and achieved higher mean sub-goal quiz scores than participants in the control condition, suggesting that scaffolds from the Pedagogical Agents were effective in influencing how students performed on sub-goal quizzes.

Research Question 2: What Are the Effects of Prompts and Feedback on Time Spent on Sub Goals?

To address this research question, we performed an Independent Samples t -test, with time spent on sub goals as our independent variable. We calculated the time spent per sub goal by scanning through the log-files and calculating duration by obtaining the time started and time completed for each sub goal the participant worked on. We obtained data from 60 participants, however each participant completed at least two sub goals (range = 2 to 7; Mean number of sub-goals per participant = 3.05, $SD = 1.28$), and so we obtained 183 time data points.

Results from the Independent Samples t -test revealed a significant difference on time spent per sub goal among experimental conditions; $t(181) = -2.246, p = .03, \eta^2 = 0.03$. Specifically, participants in the prompt and feedback condition spent significantly more time working on each of their sub goals ($M = 37.54$ minutes, $SD = 26.54$ minutes) than participants in the control condition ($M = 28.55$ minutes, $SD = 27.05$ minutes). This finding indicates that when participants interacted with the Pedagogical Agents during learning, they spent significantly more time working on the sub goals that they set, which suggests that these participants made more valuable use of their learning time, compared to students who did not interact with the Pedagogical Agents during learning.

Research Question 3: Are there Significant Differences between Condition and Relevant Pages?

In order to test this research question, we performed two Independent Samples t -tests, with relevant page ratio and irrelevant page ratio as the independent variable for each test. The page numbers, which participants navigated to were obtained from the log files, and were then assessed as being relevant or irrelevant to the current sub goal. The page ratio was then calculated by dividing the total relevant pages/total pages and the total irrelevant pages/total pages for the relevant ratio and irrelevant ratio, respectively.

Results from these analyses indicated two sets of significant results. We found a significant difference in the ratio of relevant pages visited among experimental conditions; $t(181) = -5.442, p = .00, \eta^2 = 0.14$. Participants in the prompt and feedback condition had significantly greater relevant page visited ratios ($M = .62, SD = .26$) than participants in the control condition ($M = .40, SD = .31$). Furthermore, we found a significant difference in the ratio of irrelevant pages visited among experimental conditions; $t(174.838) = 2.853, p = .012, \eta^2 = 0.04$. Participants in the prompt and feedback condition had significantly lower irrelevant page visited ratios ($M = .35, SD = .24$) than participants in the control condition ($M = .47, SD = .32$). These results indicate that if students interacted with the Pedagogical Agents during learning, this allowed for them to spend more time reading pages, which are relevant to the current sub-goal, and reading less pages that are irrelevant to the current sub goal. This suggests that students contributed effectively to completing their sub-goals, in terms of visits to relevant pages and less visits to irrelevant pages.

Discussion and Educational Implications

The results obtained from these analyses revealed the importance of providing prompts, scaffolds, and feedback to students as they learn with multi-agent, adaptive hypermedia-learning environments. In addition, these results have important implications for Self-Regulated Learning because they inform us of the usefulness of setting sub-goals, and of including Pedagogical Agents in hypermedia-based learning systems, in order to provide prompts and feedback to students to ensure effective SRL.

Moreover, it is important to consider the educational implications from these results. First, the role of the teacher is crucial in the classroom because teachers can provide the appropriate scaffolds and feedback to his or her students to ensure that each student understands the content, and can apply these constructs to every day life. If teachers were able to collect multi-channel data on each student, this would allow for the teacher to gather reliable data, as opposed to making inferences, about each student's progress during learning. In addition, teachers can use methods, such as having students set sub-goals, in order for them to grasp an optimal understanding of the topic they are learning. Future work can, therefore, implement using multi-channel data in the classrooms, which would allow for accurate measures of student performance, and thus promote optimal learning in the classroom. The agents can be adaptive based on students' pre-test scores, sub-goal quiz scores, page quiz scores, use of SRL strategies (e.g., Taking Notes, Judgment of Learning, and Content Evaluation), emotional states, and think-aloud data, which they can assess and provide the appropriate prompts and feedback,

based on performance on all of these aspects. If we were to design adaptive multi-agent hypermedia systems, this would allow for us to cater to individual students' learning needs, and will promote optimal self-regulated learning when interacting with CBLEs.

Overall, results from this study indicated that there is support for incorporating Pedagogical Agents in hypermedia-learning environments, as well as having students set sub-goals to work on as they navigate through the system. The PAs can be an important asset to include in Computer-Based Learning Environments because they have been shown to be effective in providing prompts, scaffolds, and feedback, which have assisted learners in obtaining higher mean sub-goal quiz scores, spending more valuable time accomplishing sub-goals, visiting more pages, which were relevant to their sub-goals, and visiting fewer pages, which were irrelevant to their sub-goals, compared to students who were not provided with this assistance from the PAs. Therefore, the PAs were effective in fostering and promoting effective completion of sub-goals, and thus effective self-regulated learning. It is advantageous, therefore, to maintain two experimental conditions in our set-up, in order to compare and determine how students interact with the PAs the most effectively. In addition, the results confirmed that setting sub-goals can be a useful tool for learning, and that when students are scaffolded appropriately, they can make effective use in completing the sub-goals they set.

References

- Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R. (2013) Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 427-449). Amsterdam, The Netherlands: Springer.
- Baylor, A. L. (2009). Promoting motivation with virtual agentes and avatars: Role of visual presence and appearance. *Philosophical Transactions*, 364, 3559-3565. doi: 10.1098/rstb.2009.0148
- Kinnebrew, J., Biswas, G., Sulcer, B., & Taylor, R. (2013). Investigating self-regulated learning in teachable agent environments. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 451-470). Amsterdam, The Netherlands: Springer.
- Lester, J.C., Mott, B.W., Robinson, J.L., & Rowe, J.P. (2013). Supporting self-regulated science learning in narrative-centered learning environments. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 471-483). Amsterdam, The Netherlands: Springer.
- Mayer, R. E., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology, Applied*, 18(3), 239-252. doi: 10.1037/a0028616
- Park, S., & Catrambone, R. (2007). Social facilitation effects of virtual humans. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(6), 1054-1060. doi: 10.1518/001872007X249910.
- Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Perry, R. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36, 36-48.
- 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 York, NY: Springer.
- 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., & Hadwin, A. F. (2013). nStudy: tracing and supporting self-regulated learning in the Internet. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 293-308). Amsterdam, The Netherlands: 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.

Civilian Analogs of Army Tasks: Supporting Pedagogical Storytelling Across Domains

Andrew Gordon, Mark Core, Sin-Hwa Kang, Catherine Wang, and Christopher Wienberg
ICT, University of Southern California, 12015 Waterfront Drive, Los Angeles, CA 90094
Email: gordon@ict.usc.edu, core@ict.usc.edu, kang@ict.usc.edu, catherhw@usc.edu, cwienberg@ict.usc.edu

Abstract: Storytelling is the most basic means by which people learn from the experiences of others. Advances in educational technologies offer new opportunities and experiences for learners, but risk losing the natural forms of pedagogical storytelling afforded by face-to-face teacher-student discussion. In this paper, we present a technology-supported solution to the problem of curating and algorithmically delivering relevant stories to learners in computer-based learning environments. Our approach is to mine public weblogs for textual narratives related to specific activity contexts, both inside and outside the domain of the target skillset. These stories are then linked directly to task representations in the learner model of an intelligent tutoring system, and delivered to learners along with other tutoring guidance. We demonstrate our approach to curating stories by creating collections of narratives that are analogous to tactical tasks of the U.S. Army, and evaluate the difficulty of incorporating these stories into intelligent tutoring systems.

Introduction

In George Lucas's classic 1977 film, *Star Wars*, protagonist Luke Skywalker considers the rebels' plan to destroy the enemy's Death Star super-weapon by dropping a bomb down a ventilation shaft: "I used to bull's-eye womp rats in my T-16 back home. They're not much bigger than two meters." Here Luke draws an analogy between a military task (precision bombing) and a civilian pastime (womp rat hunting), which he uses to persuade others of the technical feasibility of the plan. With this dialogue, George Lucas establishes that a young moisture farmer from the planet Tatooine could believably pilot a rebel fighter in space combat, appealing to the audience's commonsense intuitions about the applicability of experiences across domains.

The science of transfer learning, however, is often counter to these commonsense intuitions. In their survey of research in educational psychology on transfer learning, Day and Goldstone (2012) note that the spontaneous transfer of solutions across situations is difficult for learners. Much of this difficulty may be attributed to the challenges that learners have in recognizing that past situations have solutions that are relevant to current problems. In an often-repeated result, Gick and Holyoak (1983) found that subjects perform poorly on a classic insight problem, even when directly analogous problems and solutions were presented immediately prior. However, explicitly noting the relevance of the previous case to the current problem led to dramatic improvements, implicating retrieval and correspondence as the cognitive challenges in the application of experience to new situations. In human-to-human tutoring, the tutor often performs these two cognitive tasks, retrieving stories from their own past experiences and explaining the correspondence between these experiences and learners' impasses. As researchers seek to model effective tutoring in intelligent tutoring systems, we must ask: Can the process of pedagogical storytelling be automated as well?

In this paper, we present a technology-supported solution to the problem of curating and algorithmically delivering pedagogically relevant stories from inside and outside of the target skill domain, for use in intelligent tutoring systems. Our solution is to harvest collections of stories that are relevant to specific task domains from public weblogs, and to provide system developers with the tools needed to connect these stories to the skill representations that drive intelligent tutoring behaviors. We describe an application of our approach, the Civilian Analogs of Army Tasks, consisting of a hundred collections of personal stories from public weblogs that are analogously related to tactical tasks performed by the United States Army. We describe the technical methodology used to curate these collections, and an approach to automated pedagogical storytelling in intelligent tutoring systems that links individual stories to explicit task models. Finally, we describe an online experiment to assess people's abilities to draw connections between stories and tasks, comparing stories both inside and outside the task domain.

Civilian Analogs of Army Tasks

Officers, enlisted soldiers, and government civilians in the United States Army are routinely tasked with missions that are far more diverse than those depicted in Hollywood war movies. Along with the skills of armed combat, Army personnel in recent deployments have conducted bilateral negotiations with local government officials, provided preventative dentistry services, constructed educational facilities, contracted covert informants, produced mass media programming, and razed fields of illicit agricultural crops. The need to execute tasks such as these often arises during deployments based on changing conditions in the operational

environment. As a consequence, soldiers often are faced with tasks for which they have received little institutional training. The United States Army, like other very large organizations that must be highly adaptable in conducting operations, alleviates some of the need for more training by capitalizing on the wisdom of the organization itself, i.e. the experiences of other soldiers who have previously executed similar tasks. To enable learning from these experiences, effective knowledge management continues to rise as a major priority within the Army. Sharing knowledge in online discussion forums is one solution, such as those integrated into the organization's Army Knowledge Online (AKO) web portal. An example is the popular Company Command web forum (Dixon et al., 2005), where company commanders effectively bridge the gap between people who have experiences to share and those that can most benefit from receiving them.

Our belief is that organizational learning inside the Army could be improved by looking outwards, expanding the pool of experiences from which it draws to include those from analogous civilian domains. The vast majority of tasks assigned to soldiers are clearly analogous with one or more activities that are skillfully executed by people outside of the Army. Some analogies are extremely close. School construction in the Army requires nearly exactly the same skills as those employed by professional construction workers. Razing fields of crops in Afghanistan has a lot in common with razing fields of crops in Arkansas. Other analogies share fewer surface features but share deep features. Bilateral negotiations between soldiers and government officials pose many of the same concerns found in international business negotiations. The production of mass media material in Army psychological operations is not unlike the work done by brand managers at advertising agencies. The stories told by these construction workers, crop farmers, business executives, and advertising agents could be useful to soldiers in the acquisition of Army skills.

To investigate the broader applicability of civilian experiences to Army tasks, we conducted an analysis of a large-coverage taxonomy of tasks assigned to Army units. The Army Universal Task List: Field Manual 7-15 (United States, 2012) lists official designations for hundreds of tasks that are assigned to Army units at the tactical level (as opposed to the operational or strategic level), to include tasks such as Purify Water (ART 4.1.3.11.1) and Support Famine Prevention and Emergency Food Relief Programs (ART 7.3.3.5). For each of these tasks, we brainstormed the activities in the civilian world that were most analogous, e.g. working at a water treatment facility and charitable food drives. Although many Army tasks lacked obvious civilian analogs, we found that correspondences could be made at higher levels of abstraction. We created a notional list of one hundred and two civilian activities that were broadly relevant to Army tasks, organized into eight high level Army concerns. Table 1 lists these categories, with examples of civilian tasks that we saw as a potentially relevant source of stories from large numbers of civilian practitioners.

Table 1: Eight high-level Army concerns and examples of analogous civilian tasks

High-level Army category	Example civilian tasks
Training	Going to shooting ranges, Backpacking, and Triathlons
Education	Tutoring, Taking exams, Study abroad
Leadership	Coaching sports teams, Refereeing sporting events
Battle Command	Firefighting, Wedding planning, Poker
Stability Operations	Prison operations, Political campaigns, Fundraising campaigns
Support Operations	Charitable food drives, Health examinations, Emergency room care
Signals Intelligence	Hackathons, Computer repair, Fireworks
Logistics	Emergency evacuations, Class field trips, Truck driving

Retrieving Narratives From Public Weblogs

Millions of people chronicle the events in their personal lives online in public weblogs, creating opportunities to automatically amass large collections of stories about specific civilian activities. However, Swanson (2011) estimated that only 5.4% of all non-spam English-language weblog posts are personal stories, defined as non-fictional narrative discourse that describes a specific series of causally related events in the past, spanning a period of time of minutes, hours, or days, where the storyteller or a close associate is among the participants. Using supervised machine learning techniques, Swanson constructed a text classifier to identify these personal stories in streams of weblog posts (precision 59.1%, recall 41.4%), which was subsequently used by Gordon et al. (2012) to construct a story extraction pipeline from data provided by a commercial weblog aggregator. Running since January of 2010, this pipeline has collected over 30 million stories thus far.

This story repository included numerous narratives for each of the 102 activities that we identified as civilian analogs of Army tasks. In order to curate collections of stories for each activity, we used two prototype search technologies. The first, *StoryUpgrade* (Gordon et al., 2012), is an activity search tool that incrementally builds a statistical topic model for use as a textual search query through the use of relevance feedback provided by the user. In this approach, users begin a search for stories by authoring a paragraph-sized "boring story" of the desired activity: a fictional past-tense narrative describing the activity that includes as much of the domain-specific vocabulary as possible, but avoids specific terminology unrelated to the activity, e.g. proper names of

people and places. Terms in this initial query are weighted and elaborated by hand-annotating the relevance of top search results, which are then used to iteratively query the collection until an adequate number of relevant stories are identified, or when the top search results consistently show irrelevant stories. The second search tool we used, *PhotoFall* (Wienberg & Gordon, 2012), capitalizes on the finding that 82% of photographs in narrative posts were taken during the course of events described in the surrounding narrative text. The PhotoFall search tool exploits this close connection between photographs and the narrative text to provide search users with a fast relevance feedback mechanism. Photographs from the top one thousand results of the current query are extracted and shown to the search user as a proxy for the full posts, allowing the user to quickly guess their relevance from the image alone. This feedback is again used to iteratively weight and elaborate the terms in a textual query, improving the learned topic model for the activity for both retrieval tools. In both of these search tools, story collections are created only from stories judged as relevant, a process that overcomes the less-than-perfect precision of both the learned topic model and the initial story classifier.

Using these search tools, we found that it required roughly three person-hours to curate modest-sized story collections for each of our 102 activities. On average, we identified 14.25 relevant stories for each activity, and a total of 1454 narratives of civilian analogs of Army tasks.

Integration into an Intelligent Tutoring System

We sought to integrate pedagogical storytelling into technology-based immersive training environments where learners acquire skills through deliberate practice, and where instructional strategies are partially automated in software through Intelligent Tutoring Systems (ITS). The central problem of pedagogical storytelling in these environments is to deliver just the right story at just the right time in the course of a specific learner's acquisition of target skills. As a first attempt to integrate pedagogical storytelling into immersive training environments, our approach was to piggyback on top of user-modeling technologies used by an ITS to critique and provide feedback for a learner's performance.

Our integration approach is best suited for ITSs that model learners by aligning their behavior in an immersive training environment with an explicit task model of expert performance of a skill. This ITS design is seen in the BiLAT training system, a game-based environment for practicing negotiation skills in a cross-cultural context (Kim et al., 2009). Built for training U.S. Army personnel, learners in BiLAT prepare for and execute face-to-face negotiations with virtual characters in order to improve conditions in a fictional Iraqi city. Using a menu-based dialogue system, learners select conversational moves in an evolving dialogue context. The associated ITS in BiLAT (Lane et al., 2013) compares these selections to an explicit task model of expert performance, formulated through a cognitive task analysis of the skills of expert negotiators. This explicit task model is organized as a hierarchy of required steps, optional steps, rules of thumb, and actions to be avoided in expert performance of negotiation skills in a cross-cultural context. In developing BiLAT's ITS, every possible dialogue action in BiLAT was associated positively or negatively with one or more items in this hierarchy. When using BiLAT, a learner's abilities are assessed by aggregating evidence throughout the interaction, enabling the ITS to provide targeted feedback and hints during and after each training session.

Pedagogical storytelling can be integrated into an ITS of this sort by linking stories to specific items in the explicit task model. When the learner demonstrates poor performance on a particular task, the ITS can present one or more of the stories linked to the task, determined by the nature of the performance problem. In the BiLAT domain, there are three specific causes of poor performance that storytelling may address: 1) Learners may ignore the advice given about negotiation and cultural awareness thinking they know better. 2) The advice is necessarily general and real world problems require determining specific words to say. 3) The domain is inherently unpredictable; correct actions do not always bring success, and incorrect actions do not necessarily hurt performance. To enable an ITS to select the most appropriate story for a learner, we developed a taxonomy of relations between task-model elements and stories (Table 2). To address motivational issues, an ITS could use stories labeled with *evidence*, *warning* and/or *motivation* relations to convince the learner to heed the advice given. Stories labeled with *explanation* could teach learners causal relationships to help them adapt to the variety in real world situations. *Background* stories would contain concrete examples allowing learners to practice matching general advice to specific situations. Stories about *exceptions* could promote adaptive thinking to deal with real world uncertainty and help prevent misconceptions.

Table 2: Task model / story relationship taxonomy

Relation name	Relation to task
Evidence	An example of correct behavior resulting in a positive outcome.
Warning	An example of incorrect behavior resulting in a negative outcome.
Explanation	An explanation of causal information underlying correct behaviors.
Exception	An exception to the rules due to luck or special circumstances.
Motivation	Personal and/or emotional motivations for performing the correct behavior.
Background	A real world example.

Evaluation of Authoring Feasibility

The ITS-based pedagogical storytelling approach described in the previous section requires that instructional designers can reliably make the relational connections between stories and task model elements. Given a collection of stories, instructional designers must read and comprehend each one, identify the elements of the task model to which they are related (if any), and characterize the relationship between these elements and the stories. Our hypothesis was that instructional designers are capable of performing these tasks, but we were concerned that stories from analogous domains might be more difficult than those directly related to the training domain. To investigate these concerns, we conducted an experiment to assess people's ability to link stories to task models, comparing performance on within-domain stories to analogous-domain stories via measures of inter-coder agreement and self-reported confidence.

For the purpose of this experiment, we selected the training domain of bilateral negotiations as conducted by personnel in the U.S. Army. This choice allowed us to use the explicit task model from the BiLAT ITS as a basis for our experimental task. The complete task model of BiLAT is a large collection of actions and rules of thumb organized into a hierarchy defining different levels of granularity. We focused on the top-level of this hierarchy and identified seven crucial cross-cultural negotiation skills (Table 3).

Table 3: Simplified list of skills for bilateral negotiations

Skill	Description
Gather background information	Gather background information about your meeting partners and their possible actions as well as your possible actions and their impacts.
Respect culture of partners	As possible, learn and follow customs of your meeting partners, and follow their lead (if they are hosting).
Present yourself well	Given the culture of your meeting partners and the situation, present yourself in the best light. For example, some cultures value the concept of "face" and you should do nothing to lose face or cause your partners to lose face.
Use caution with interpreters	If you have the option to choose an interpreter, you may want to use one that you are already comfortable with, or who is a relevant specialist. Otherwise, be careful of bad translations.
Master small talk	Avoid sensitive topics, focus instead on neutral topics, or topics of interest to your meeting partners.
Build relationships	As culturally appropriate, get to know your meeting partners and build relationships.
Use a win/win negotiation strategy	Where possible, avoid a win/lose strategy. The negotiation should support a long term relationship. The idea is that everyone is able to present their interests and reach a middle ground which everyone is comfortable with.

Using the *StoryUpgrade* and *PhotoFall* weblog story search tools described above, we gathered eight nonfiction stories related to the skills in Table 3. Four of these stories (*domain stories*) were narrations of experiences of soldiers during recent military deployments in Afghanistan (i.e., negotiations and culture and language issues). The four remaining stories (*analogous stories*) were from the analogous domain of businessmen relating experiences in foreign countries (i.e., buying and selling in marketplaces, dinner table etiquette, corporate culture). For the purpose of our experiment, each of these eight stories was abridged to a single page of text. In several cases, we wrote an introductory paragraph for the story that provided context to the readers that would normally be presented directly on the weblogs from which these stories were gathered.

As a proxy for a population of training developers, we recruited 202 college-educated Americans to participate as subjects in a 35-minute online experiment. Using a within-subjects experimental design with randomized ordering of trials, we tasked subjects to read each of the eight stories, identify the most relevant skill from Table 3, and identify the relation from Table 2 that best describes the nature of this relevance. For each selection task, subjects were given the option to select "None of the above," and were additionally asked to judge their confidence in their selection on a 7-point Likert scale.

To compare inter-coder reliability, we calculated Krippendorff's alpha on the nominal data from this study. Table 4 shows low levels of agreement overall among subjects for both skill and relation types. Agreement is better for domain stories than for analogous stories, and better for skill selection than for relation selection.

Table 4: Krippendorff's Alpha for Skill Annotation and Relation Annotation, with Observed Disagreement (*Do*) and Expected Disagreement (*De*)

	Domain stories			Analogous stories		
	<i>Alpha</i>	<i>Do</i>	<i>De</i>	<i>Alpha</i>	<i>Do</i>	<i>De</i>
Skill	.448	.133	.241	.156	.192	.228
Relation	.260	.181	.245	.132	.213	.245

To compare confidence scores, we analyzed the data with the Repeated Measures ANOVA, a variation of ANOVA used when the same subjects participate in all conditions of an experiment. Table 5 shows significantly higher confidence when linking domain stories versus analogous stories, both for skill selection and relation selection.

Table 5: Effect of the type of a story on the level of subjects' confidence for their decisions regarding the *Skill* and the *Relation*.

	Domain stories		Analogous stories		<i>F</i>	η^2	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Skill	5.59	.89	5.05	.99	157.14	.44	<.001
Relation	5.15	1.07	4.91	1.12	27.71	.12	<.001

Overall, the results of this evaluation indicate to us that instructional designers will, indeed, find it more difficult to link stories from analogous domains to task models than those from within the domain of the task. As a consequence, we expect that utilizing the story collections in the Civilian Analogs of Army Tasks will place additional burdens on training developers over the use of stories from within the domain of the training application. The additional challenges of cross-domain story linking should be considered when developing authoring tools for intelligent tutoring systems and assessing learning effectiveness.

Conclusions

In this paper we have explored how technology can support new forms of pedagogical storytelling, where the narrated experiences of practitioners across diverse skill domains can be found and delivered to learners within the contexts of immersive training environments. First, we have shown that stories related to specific activities and skill sets can be harvested from public Internet weblogs, and efficiently retrieved using text retrieval technologies that incorporate relevance feedback. Second, we have shown that analogies can be drawn between activities narrated in public weblogs and a broad range of skillsets that are the focus of training in a large organization, the U.S. Army. Third, we provide a mechanism by which authors of intelligent tutoring systems for immersive training environments can automate the delivery of stories to trainees, linking specific stories to specific items in an explicit task model. Fourth, we conducted an evaluation of the feasibility of this authoring process, finding promising initial results, but also finding that stories from analogous domains may pose additional difficulties for instructional designers.

Ultimately, the technologies and methodologies described in this paper will only be useful if they facilitate learning. What is needed is additional research to evaluate the training effectiveness of pedagogical storytelling in immersive training environments, both for within-domain and analogous-domain narratives.

References

- Day, S., and Goldstone, R. (2012) The Import of Knowledge Export: Connecting Findings and Theories of Transfer of Learning. *Educational Psychologist* 47(3):153-176.
- Dixon, N., Allen, N., Burgess, T., and Kilner, P. (2005) *Company Command*. West Point, NY: Center for the Advancement of Leader Development & Organizational Learning.
- Gick, M. and Holyoak, K. (1983) Schema induction and analogical transfer. *Cognitive Psychology* 15:1-38.
- Gordon, A., Wienberg, C., and Sood, S. (2012) Different Strokes of Different Folks: Searching for Health Narratives in Weblogs. ASE/IEEE International Conference on Social Computing, Amsterdam.
- Kim, J., Hill, R.W., Durlach, P.J., Lane, H.C., Forbell, E., Core, M., Marsella, S., Pynadath, D., and Hart, J. (2009). BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context. *International Journal of Artificial Intelligence in Education*, 19(3), 289-308.
- Lane, H.C., Hays, M.J., Core, M.G., & Auerbach, D. (2013). Learning intercultural communication skills with virtual humans: Feedback and fidelity. *Journal of Educational Psychology* 105(4):1026-1035.
- Swanson, R. (2011) Enabling open domain interactive storytelling using a data-drive case-based approach. Dissertation. University of Southern California.
- United States (2012) *The Army Universal Task List: Field manual 7-15*, Washington, DC: Headquarters, Dept. of the Army.
- Wienberg, C. and Gordon, A. (2012) PhotoFall: Discovering Weblog Stories Through Photographs. ACM Conference on Information and Knowledge Management (CIKM-2012), Maui, HI.

Acknowledgments

The projects or efforts depicted were or are sponsored by the U. S. Army. The content or information presented does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

From Playing a Game to Solving an Equation

Shulamit Kapon, Faculty of Education, University of Haifa, 199 Aba Khoushy Ave., Mount Carmel, Haifa 3498838, ISRAEL, shulamit.kapon@edu.haifa.ac.il

Abstract: Although the potential of digital games as instructional resources is well known, facilitating deeper learning from games is acknowledged to be a considerable challenge. This paper argues that examining transfer from the realm of a digital game to the realm of a discipline as a case of analogical reasoning can be productive when trying to understand and later design scaffolds that connect the knowledge representations embedded in games to formal disciplinary knowledge representations used at school. This claim is supported through (1) an analysis of the design of a multiplatform commercial educational game Dragon Box 5+, aimed at teaching children to solve algebraic equations; and (2) tracking the experiences of several players. The transfer from procedures in the game to algebraic procedures is explored as a case of analogical reasoning, specifically addressing central sub-processes such as retrieval, mapping, evaluation of inferences, and the role of prior knowledge in these processes.

Introduction

Digital games are considered to be a powerful medium for learning (Aldrich, 2004; Gee, 2003). However, some studies of game-based learning as an instructional resource in the formal disciplines suggest that the links between the game world and formal disciplinary knowledge required in a school-based context are not straightforward (Holbert & Wilensky, 2012; Kebritchi, Hirumi, & Bai, 2010).

Clark et al. (2011) and Habgood and Ainsworth (2011) showed that digital games that integrate the subject matter with a game idea in which the external representation of the learning content is explicitly explored through the core mechanics of the game are more efficient in creating this link. Clark et al. termed this "conceptually-integrated games" whereas Habgood and Ainsworth called it "intrinsically integrated games", and both acknowledge the challenge of transfer from the game world to formal disciplinary-based reasoning. Holbert and Wilensky (2012) attributed the transfer problem to the incongruence of the epistemological framings (Hammer, Eleby, Scherr, & Redish, 2005). More specifically, the set of expectations that a student has about a real-world or school-based situation that requires disciplinary-based reasoning is very different from the set of expectations in the game world. Based on this explanation they suggested that when the appropriate formal representations are integrated into the player interface, the likelihood for transfer increases since the players are implicitly encouraged to connect different epistemological framings.

The present paper suggests a complementary explanation. As in Holbert and Wilensky's paper, the theoretical approach employed is the Knowledge in Pieces epistemological perspective (e.g., diSessa, 1993; diSessa & Sherin, 1998; Wagner, 2006); however, following Kapon and diSessa (2012), the transfer from the realm of the digital game to the realm of the discipline is theoretically explored as a case of analogical reasoning. Central sub-processes of analogical reasoning such as retrieval, mapping, evaluation of inferences, and the role of prior knowledge are examined in the context of playing an educational digital game. The theoretical claims will be illustrated through an analysis of an award-winning commercial educational digital game entitled Dragon Box Algebra 5+ (WeWantToKnow), and examples from players' experiences (7 children, 5th, 6th and 7th grades). The 7th graders who played the game played after learning to solve linear equations in school. The 5th and 6th graders played before a formal presentation of algebra.

Theoretical Framework

Analogies and analogical reasoning have been studied from a wide range of perspectives. A core focus of attention has been the ways in which analogies foster understanding in some new situation or domain (the target) by comparing it to a more familiar one (the source/base) (Gentner, 1983; Holyoak & Thagard, 1989). Several models suggest that spontaneous analogical transfer is more likely to occur when the target shares multiple features with the source analogue (Falkenhainer, Forbus, & Gentner, 1986; Holyoak & Koh, 1987) and that this process is mediated by pragmatic goals (Holyoak & Thagard, 1989; Keane, 1996). These models cohere with the design principles of conceptually integrated games, in which the degree of similarity of representations and actions between the realm of the game and realm of the discipline is by definition increased.

The ability to apply relational knowledge across different contexts is a powerful learning mechanism. In analogical reasoning this transfer takes place when relations from a familiar source domain are mapped onto relations and objects in the target domain, allowing the individual to analogically infer new relations or objects in the target domain (Dunbar, 1995; Gentner, 2010; Holyoak & Thagard, 1989). However, spontaneous retrieval of a relevant, appropriate, and productive source for an analogy is not trivial, and often superficial similarity

plays a major role in analogical retrieval rather than deep relational structures (Gentner, Rattermann, & Forbus, 1993; Gick & Holyoak, 1980). The issue of retrieval seems highly germane to the development of links between the game world and formal disciplinary knowledge required in a school-based context. After all, why should a student even consider an activity in a game setting as similar to an activity in a formal disciplinary context?

Studies of students who reason through instructional analogies show that the evaluation of the plausibility and applicability of an analogical inference is not just a question of the degree of similarity and pragmatic considerations. Rather, these studies suggest that the learner's knowledge in the target domain directly affects the acceptance or rejection of this explanation (Brown & Clement, 1989; Kapon & diSessa, 2012). Kapon and diSessa (2012) argued that learning through analogy always involves a bootstrapping process *within* the target domain. Drawing and elaborating on the model of p-prims (diSessa, 1993), they suggested that when reasoning through instructional analogies, students activate particular knowledge elements, termed explanatory primitives, or e-prims, in light of which they assess the plausibility of the analogical inference. Differences between individuals accepting or rejecting the suggested analogical inferences were explained by differences in the repertoire of cued knowledge elements and a dynamic assessment of their applicability to the target domain. Thus, it is reasonable to expect that prior knowledge in general and in the target disciplines in particular will affect the analogical transfer from the realm of the game to the realm of the discipline, not only in terms of retrieval but also in terms of the evaluation of inferences.

The Game

This paper analyzes a commercial educational digital game that aims to teach children to solve algebraic equations. Dragon Box 5+ (WeWantToKnow) is an award-winning multiplatform commercial game. Players start with a two sided game board, a box and some cards. The goal is to get the box alone on one side of the game board by cancelling out negative cards, adding cards to both sides, placing cards below or next to a group of cards. Gradually the cards are replaced by numbers and variables, and actions in the game start to *look* like addition, division and multiplication operations (see Figure 1). The game has 5 chapters with 20 problems ("levels") each. As the game progresses, instead of getting the box alone the player solves (at least in terms of representations) algebra equations for x . During the game the player gets constant feedback.

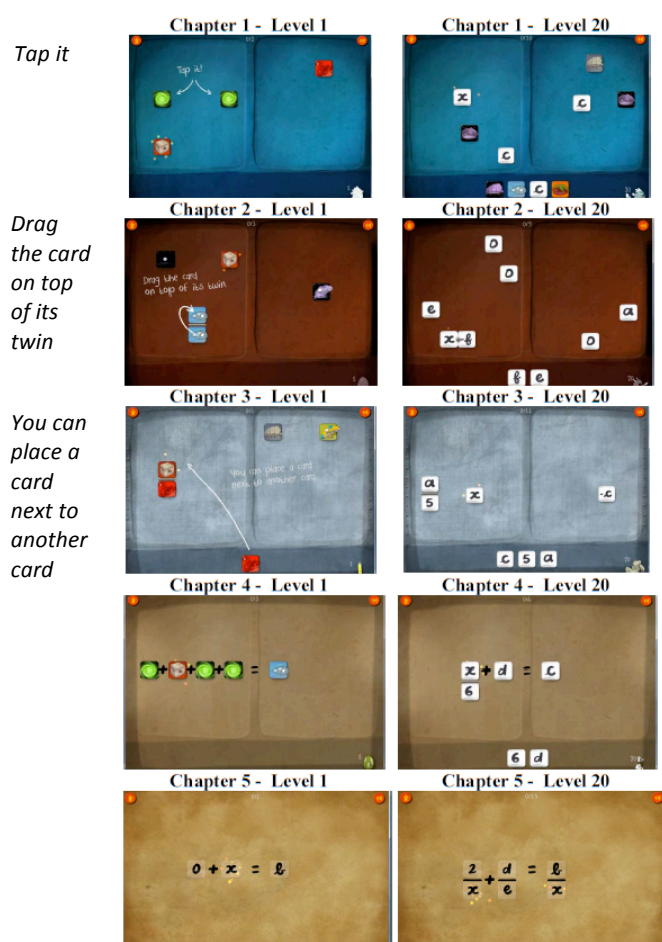


Figure 1. The game board as the game progresses. On the left are the instruction that appear on the screen.

Tracking Players' Experience

Data on players' experience were collected by graduate students enrolled in a seminar taught by the author. A considerable part of the seminar dealt with the notion of transfer of learning from multiple perspectives. Seven graduate students (5 math teachers, 1 science teacher, and 1 psychologist) participated in the seminar. Each graduate student worked with one child (5th-7th grade) who volunteered to participate in the study. Each child (N=7) met with the graduate student five times (25-50 minutes each). During each meeting the child played and completed one chapter and the grad student observed. The computer screen (game board and in the corner, a video webcam of the child) and all the conversations were recorded. The students were instructed that if the child asked for assistance during the game they should only use "game language" unless the child explicitly used mathematical terms. A short semi-structured interview took place, after each chapter. The interviewer asked the child to return to one or two levels, solve them again and explain how he or she solved them. No questions were asked during the game to avoid interrupting the flow (Czikszenmihalyi, 1990). Additional questions were asked to probe whether the child made connections to the mathematical meaning. The interviewers also took field notes regarding difficulties on specific levels, requests for assistance, etc. All the sessions (play and interview) were transcribed including a time stamp and a description of the child's actions on each level. Particularly difficult levels were identified based on the relative time spent on the level, the number of repetitions needed to get it right, and explicit requests for assistance.

Analogy from the Realm of the Game to the Realm of Algebra

The next subsections discuss central terms and sub-processes in analogical reasoning, such as mapping, surface and relational similarities, retrieval, the source of the analogy, and the role of prior knowledge in the evaluation of the analogical inference, as they come to bear on the design and player experience of Dragon Box 5+.

Relational Similarities

At the design level, Dragon Box presents an analogy with rich relational similarities. The player's actions in the game as well as the rules that are enforced with the constant feedback given to the player are isomorphic to the actions and rules used when solving linear algebraic equations. (1) When adding a card (equivalent to arithmetic addition) to one side of the game board, the player is graphically reminded to add the same card to the other side. In algebraic equations the rule is to add variables and numbers to both sides of the equation. (2) When placing a card next to (equivalent to arithmetic multiplication) or below (equivalent to division) another card, the player is graphically reminded to do so to all groups of cards on both sides of the game board. Again, when we multiply or divide one side of an equation by a number or variable, we should do the same to the other side of the equation, and this means multiplying or dividing every component on each side of the equation with the same number or variable. (3) The player receives feedback after completing each level: (a) the box is alone - equivalent to success in isolating the variable; (b) right number of moves - equivalent to success in efficiently solving the equation with a minimal number of steps; (c) right number of cards - equivalent to canceling redundant numbers and variables (e.g., cancelling $-2+2$). In fact one might even be tempted to think that the actions in the game and the algebraic actions are isomorphic production rules (Anderson, 1987).

Analogical Transfer

Despite the rich relational similarities in the structure of the actions in the game to algebraic procedures, very few students spontaneously ascribed mathematical meaning to their activities, and those who did so ascribed only partial meaning (e.g., treating the equivalent to addition in the game as addition, but not seeing the equivalent to multiplication in the game as multiplication). The literature on analogical reasoning suggests several explanations for the spontaneous failure of analogical transfer: (1) The students did not understand the source of the analogy (as in Clement, 1993). (2) The students were unable to retrieve the relevant information (as in Gentner, Loewenstein, Thompson, & Forbus, 2009). (3) Based on their prior knowledge about games and algebra the students failed to see the game situation and the algebraic situation as similar, and thus rejected the entailed analogical inferences (as in Brown & Clement, 1989; Kapon & diSessa, 2012). In the following sections I discuss each explanation in the context of the game and its pedagogical implications. Note that the children were fully engaged in the game and enjoyed it. They often wanted to play an extra chapter after each session, they were happy to meet with the graduate student for the next session, and said that the game was fun. Thus the difficulty to ascribe mathematical meaning to the actions in the game cannot be attributed to affect.

An Accessible Source for the Analogy

The game has five chapters, and the rules of the game are gradually presented in the first three. As the students move through the chapters, the symbolic representations gradually transform into formal symbolic algebraic representations (see Figure 1). The fact that the participants completed chapters 4 & 5 relatively quickly compared to the time spent on chapters 1, 2 and 3 suggests that they had mastered the procedures of the game, and that they were almost unaware of the transformation in the symbolic representations. Hence the students

fully understood the source of the analogy. Moreover, the time the children spent playing the entire game and successfully completing all 5 chapters (e.g., 100 exercises) ranged from a total of 1 to a total of 1.6 hours, which is much faster than the time it generally takes students to master the equivalent math rules in a classroom context. This suggests that the procedures of the game can be considered as a productive source of the analogy.

Retrieval

Superficial similarity plays a major role in analogical retrieval; i.e., the retrieval of a base for analogy is easier if it shares similar objects, similar properties, and a similar general theme with the target, although the structure of relations are what makes the analogy a powerful cognitive tool. As the students move through the chapters, the symbolic representations gradually transform into formal symbolic algebraic representations (see Figure 1). The observations indicated that the children were almost unaware of the transformation in the symbolic representations; namely they treated them implicitly without thinking about their mathematical meaning. For instance, in the game each "night card" has a "day card" with opposite coloring. Dragging a night card over a day card cancels it (one gets a green whirlwind card, which when clicked over disappears). During chapter 1, level 12, suddenly instead of a drawing on a card, the player sees the letter "a" on a card and "-a" on another card. All the children treated the "-a" card as a "night card" and dragged it on the "a" card, but the 5th and 6th graders did not use the words "positive" or "negative". For instance one of the children said in the interview afterwards that "it just looked like the most appropriate thing to do", but she did not regard the green whirlwind as "zero". The three 7th graders did better than the younger students in this respect, but the ascribed mathematical meaning was nevertheless limited for them as well.

Scholarly work on learning from games suggests that external scaffolds are required to facilitate deeper learning from games (e.g., Barzilai & Blau, 2014; Honey & Hilton, 2010). The literature on analogical reasoning suggests that explicit alignment may facilitate analogical encoding, and subsequent retrieval of the relevant schema in other relevant situations (Gentner, Loewenstein, & Thompson, 2003). Thus explicitly aligning the actions and rules of the game with the corresponding algebraic actions and rules might prove to be a productive scaffold.

Evaluation of Analogical Inferences

Let us examine two excerpts from the interviews: (1) A 5th grader, who had not been introduced yet to formal algebraic representations, was interviewed after completing chapter 1. He was asked how he knew that if he dragged the "-a" over "a" he would cancel both cards and get the green whirlwind. He pointed to the "-" sign and said: "it symbolizes opposites". After a minute he stopped, looked at the interviewer, smiled and said "well, it doesn't really make sense to put a minus before a letter, why would they do that?" (2) A seventh grader was interviewed after successfully completing all 5 chapters. She was asked by the interviewer who pointed to the fraction line on the game board, "what is the meaning of the line?" Her answer was: "It separates both parts..... at first I thought it was a fraction line." When the interviewer asked why she decided it was not, she answered: "because a fraction line actually contains things, this divided by that, and if you divide them....., it's equal to something, and I cannot see anything here that is a division."

The two examples illustrate how the students' knowledge in arithmetic and algebra affected their evaluation of the analogical inference, a finding that is consistent with prior work on students' reasoning and the evaluation of instructional analogies. This work (Brown & Clement, 1989; Kapon & diSessa, 2012) suggests that a carefully planned discussion of meaning should be integrated into the explicit alignments suggested above.

A Concluding Remark

This paper argues that examining transfer from the realm of a digital game to the realm of the discipline as a case of analogical reasoning can be productive when trying to understand and later design scaffolds that connect the knowledge representations embedded in games to formal disciplinary knowledge representations used at school. The claim was illustrated through an analysis of the design of a commercial computer game (Dragon Box 5+) and the experiences of a few players.

One may question the generalizability of the claim made here. It is possible that the analogical mapping between the procedures in the game and the disciplinary procedures are applicable only to Dragon Box 5+. The strong analogy in the game was indeed what attracted my attention to it in the first place. However, even simulation games do not reproduce an activity identical to what they simulate, and the construction of correspondence relations between the real world and the game requires some degree of analogical reasoning.

References

Aldrich, C. (2004). *Simulations and the future of learning: An innovative (and perhaps revolutionary) approach to e-learning*. San Francisco, CA: Wiley.

- Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem situations. *Psychological review*, 94(2), 192.
- Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, 70, 65-79.
- Brown, D. E., & Clement, J. (1989). Overcoming misconceptions via analogical reasoning: Abstract transfer versus explanatory model construction. *Instructional Science*, 18(4), 237-261.
- Clark, D. B., Nelson, B. C., Chang, H.-Y., Martinez-Garza, M., Slack, K., & D'Angelo, C. M. (2011). Exploring Newtonian mechanics in a conceptually-integrated digital game: Comparison of learning and affective outcomes for students in Taiwan and the United States. *Computers & Education*, 57(3), 2178-2195.
- Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching*, 30(10), 1241-1257.
- Czikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: HarperCollins Publishers.
- 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.
- Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Steinberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 365-395). Cambridge MA: MIT Press.
- Falkenhainer, B., Forbus, K. D., & Gentner, D. (1986). The structure-mapping engine. Urbana-Champaign, IL: University of Illinois at Urbana-Champaign, Department of Computer Science.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York, NY: Palgrave Macmillan.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155-170.
- Gentner, D. (2010). Bootstrapping the Mind: Analogical Processes and Symbol Systems. *Cognitive Science*, 34(5), 752-775.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95(2), 393-405.
- Gentner, D., Loewenstein, J., Thompson, L., & Forbus, K. D. (2009). Reviving Inert Knowledge: Analogical Abstraction Supports Relational Retrieval of Past Events. *Cognitive Science*, 33(8), 1343-1382.
- Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive psychology*, 25(4), 524-575.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive psychology*, 12(3), 306-355.
- Habgood, M. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *The Journal of the Learning Sciences*, 20(2), 169-206.
- Hammer, D., Eleby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective*. Greenwich, CT: Information: Age Publishing.
- Holbert, N. R., & Wilensky, U. (2012). *Designing Video Games that Encourage Players to Integrate Formal Representations with Informal Play*. Paper presented at the The Future of Learning: The 10th International Conference of the Learning Sciences (ICLS 2012), Sidney, Australia.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, 15(4), 332-340.
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13(3), 295-355.
- Honey, M. A., & Hilton, M. (2010). *Learning science through computer games and simulations*. Committee on Science Learning: Computer Games, Simulations, and Education, National Research Council, USA: National Academies Press.
- Kapon, S., & diSessa, A. A. (2012). Reasoning Through Instructional Analogies. *Cognition and Instruction*, 30(3), 261-310.
- Keane, M. T. (1996). On adaptation in analogy: Tests of pragmatic importance and adaptability in analogical problem solving. *The Quarterly Journal of Experimental Psychology: Section A*, 49(4), 1062-1085.
- Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education*, 55(2), 427-443.
- Wagner, J. F. (2006). Transfer in pieces. *Cognition and Instruction*, 24(1), 1-71.
- WeWantToKnow. DragonBox Algebra 5+. from <http://www.dragonboxapp.com/>

Acknowledgments

I am grateful to the graduate students who helped track the players' experience: Hana Almog, Sokaina Awawde, Inbal Ben Arie, Angela Halloun, Lilach Hershkowitz, Areefa Kasim Ali, Hanan Tarabieh.

Supporting Conceptual Understandings Outdoors: Findings from the Tree Investigators Mobile Project

Heather Toomey Zimmerman, Susan M. Land, Brian J. Seely, Michael R. Mohney, Gi Woong Choi, and Lucy R. McClain, Penn State University, University Park, PA 16802 USA
heather@psu.edu, sland@psu.edu, bjs5142@psu.edu, mrm126@psu.edu, gxc207@psu.edu, lucy@psu.edu

Abstract: This design-based research project examines two iterations of *Tree Investigators*, which supports science learning with mobile devices in an arboretum. Researchers coded videorecords of children and parents (n=40 people) to understand how digital augmentations influenced observation and understandings about trees. In Iteration 1, learners focused on tree identification; Iteration 2, learners focused on tree life cycles. We focus here on Iteration 2, where children completed a pre- and post-test assessment and participated in a photographic collage task to document a tree's life cycle. Findings suggested that a touch-screen conceptual organizer that provided a model of the life cycle, along with text and contrastive images, supported people's observations. The learners also collected photographic evidence of life cycle stages in a knowledge generative task. Increases in factual and conceptual knowledge of the life cycle were observed pre- and post-test; however, learners did not show consistent appropriation of new scientific vocabulary.

Our *Tree Investigators* research and design intention is to engage people in science learning during their out-of-school time with the use of digital materials deployed on mobile computers. Our goals are to support people so that they become (a) adept *observers* who can coordinate scientific knowledge with their sensory experiences in the outdoors and (b) proficient *explainers* of scientific phenomena related to ecological systems based on their interactions with plants and animals. We adopted mobile computers given their increasing ubiquity in everyday life (Warschauer & Matuchniak, 2010) and increasing reach of mobile computers into families of modest socioeconomic means (Yardi & Bruckman, 2012).

Theoretical Framework

To support learners in outdoor environments with mobile computers, our theoretical framework brings together research findings about technological supports for science learning (Chen, Kao, & Sheu, 2005; Liu et al., 2009; Rogers et al., 2004; Squire & Jan, 2007; Squire & Klopfer, 2007; Tan et al., 2007) and research findings from place-based education (Gruenewald, 2003; Lim & Calabrese Barton, 2006; Semken, 2005; Smith, 2002).

Supporting Learners with Mobile Computers

Researchers have studied learner engagement, content knowledge acquisition, and enjoyment in outdoor settings including historical locations (Tan & So, 2011), woodlands (Rogers et al., 2004), wetlands (Liu et al., 2009), parks (Tan et al., 2007), and gardens (Chen, Kao, & Sheu, 2005). We also consider findings related to learner engagement that come from those mobile learning projects that augment real-world locations with virtual data and gaming scenarios (e.g., Dunleavy & Dede, 2014; Klopfer, 2008; Squire & Jan, 2007). Researchers have reported design elements for mobile devices that encourage data collection (Squire & Klopfer, 2007) and engagement in discourse (Hsi, 2003; Rogers, et al, 2004; Tan et al., 2011) that support science learning.

Place-Based Education

Place-based education is a pedagogical perspective that advocates for designing learning within and about local communities (Gruenewald, 2003; Smith, 2002; Sobel, 2004). Researchers (e.g., Lim & Calabrese Barton, 2006) use place-based education to understand the connections between abstracted, disciplinary knowledge and people's local knowledge and practices (Gruenewald, 2003). Place-based education in school (e.g., Sobel, 2004) advocates designing curriculum to make school-based learning pertinent to local issues and considerations. Place-based education in out-of-school settings (e.g., Tzou, Scalone, & Bell, 2010) highlights the problems that arise when the focus on disciplinary concepts disregards the manner in which learners' lives are embedded within local systems, histories, and interactions.

Semken (2005) offers a framework for science-related place-based teaching with five aspects. Semken suggests place-based teaching: (1) focuses on a setting's natural history, (2) considers the varied meanings that a place has for people who use it, (3) incorporates explorations that use authentic artifacts and representations, (4) includes ecologically and culturally appropriate pedagogy and (5) acknowledges and fosters a "sense of place" of learners, educators, and others. We adopt Semken's perspectives on place-based education to connect out-of-school learners using mobile computers to the outdoor settings in their communities. Specifically, we used

mobile devices to connect learners to scientific practices and concepts embedded within a natural garden setting of historical and ecological importance—especially in a historically important old growth stand of trees.

Methodology

We conducted two iterations of a design-based research project at an arboretum (n=40) where we collected 7.5 total hours of video. Video data were analyzed to elucidate the role of digital media deployed by mobile devices to support people in scientific observations and explanation-building talk. Our goal was to advance theory related to learning outside of school and to distill design principles related to the development of mobile computer apps and websites that can enhance families' and youths' experiences as they explore the outdoors.

Research Question

Our research investigates the following questions: How do young people and their families talk together about trees and life cycles while using the Tree Investigators mobile computer app? How does a knowledge-building photographic collage task support the development of conceptual understandings of the stages of trees' life cycles for children visiting an arboretum?

Design for Mobile Computers: Tree Investigators

In Iterations 1 and 2 of Tree Investigators, an onsite naturalist directed the families to observe trees and to coordinate their place-based observations with disciplinary information delivered by a mobile computer. Both iterations included child-friendly text, consisting of short sentences; the text's reading level was determined by the Flesch-Kincaid rating system to be at a 3rd grade level. A Ph.D.-level botanist reviewed the tree content for scientific accuracy. The mobile computers augmented seasonally or developmentally unavailable aspects of trees—mostly via digital photographs and descriptive text.

We designed Tree Investigators in Iteration 1 as a mobile website (see Figure 1) that used augmented reality (images and text layered onto the physical space) to support families to develop observations and explanations related to tree biodiversity. Iteration 1 was organized by tree species with each of eight species having its own online materials accessed by a QR code. We re-designed Tree Investigators in Iteration 2 as a mobile app (see Figures 3 and 4) that did not rely on the Internet. Iteration 2 was organized conceptually by aspects of a tree's life cycle in contrast to Iteration 1's species-centered presentation of content. Learners began Iteration 2 with observing evergreen and deciduous tree in a botanical garden and coordinated this sensory information with the conceptual model of a tree's life cycle on a mobile app (Figure 3 and 4). The final Iteration 2 activity included using a photographic collage app where learners collected photographic examples of the stages of a tree's life cycle (i.e., seed, seedling, sapling, mature tree, and dead/snag tree) in a forested area.

Participants and Setting

Across both Iterations, 40 people participated. The participants in Iteration 1 were 25 people from 11 families and the participants in Iteration 2 were 15 people from 6 families. The children were 6 to 12 years old. Given that we designed Tree Investigators for users of informal sites, we strategically recruited families who were current users of nature centers and arboreta for intergenerational education and recreation.

The research setting was the Arboretum at Penn State, which includes various groomed and curated gardens as well as a stand of old-growth hardwood trees with a network of trails. The oldest trees in this 42-acre stand pre-date the construction of the University campus in 1859. Given the logging in this area for development and for the iron industry throughout the 1800s, this old-growth stand of trees holds a protected status due to the historical, scientific, and cultural value to the area. Iteration 1 used the trees in the groomed gardens while Iteration 2 used both the groomed gardens and the old-growth forested area. Inclusion of the old-growth forest allowed for a clearer realization of our study's place-based education aims and for the learners to see actual tree specimens in all stages of the life cycle (e.g., seed, seedling, sapling, mature, dead).

Data Collection and Analysis

Families were videorecorded during a 1-hour guided tour using augmented photographic and textual elements of Tree Investigators on iPad tablet computers and iPod Internet-enabled mp3-players. Given our interest to support science learning in informal spaces, for Iterations 1 and 2 we employed an analytical framework of conversation elaboration (Leinhardt & Crowley, 1998), with talk as both a product and the process of learning. We used a theoretical-driven approach to code transcripts for evidence of observational and explanation-building talk that was derived from Allen (2002): *perceptual talk* (identification, naming, and describing species); *conceptual talk* (inference, interpretation, and prediction); *connecting talk* (life, knowledge, and inter-species connections); and *affect talk* (emotional expressions of positive or negative feelings). We conducted a detailed line-by-line analysis of families' talk using the Allen (2002) coding analysis framework.

In Iteration 2, we conducted two additional analyses. First, given our interest in digital photographs to foster the connection of local, place-based knowledge to domain knowledge in science (Land, Smith &

Zimmerman, 2013), the children in Iteration 2 took pictures of elements of the forest at the Arboretum and made a tree life cycle photo collage. The photo collages had empty slots for photographs, and the families were tasked to find an example of each stages of the life cycle in their collage, in effect requiring them to apply what they had learned to observations in the forest. We analyzed the debrief interview and the actual collage artifacts. Second, we conducted a brief (7-10 minute) pre- and post-test that was implemented as a short interview.

Findings

Iteration 1: Photographic Images Supporting Observations and Understandings

In Iteration 1, children and their families used the Tree Investigator's mobile media to connect their observations of trees to new understandings of related biological concepts. Given our goals to enhance place-based understandings, families used the mobile devices in the Arboretum to coordinate their on-site observations with abstracted scientific knowledge. Images and prompts (see Figure 1) that were part of a mobile website were used to support family observational practice and to develop explanations about the differences in trees and their characteristics related to biodiversity. As reported in Zimmerman, et al. (2013), Iteration 1 supported learners to: (a) identify relevant aspects of the trees on-site; (b) articulate an understanding of scientifically-relevant characteristics of trees' natural history (Figure 2); and (c) identify salient differences between evergreen and deciduous trees using both mobile images and specimens at the Arboretum. Our findings suggested the importance of augmented photographic elements of trees that were not seasonally or developmentally visible as contrastive cases to the onsite tree specimens.

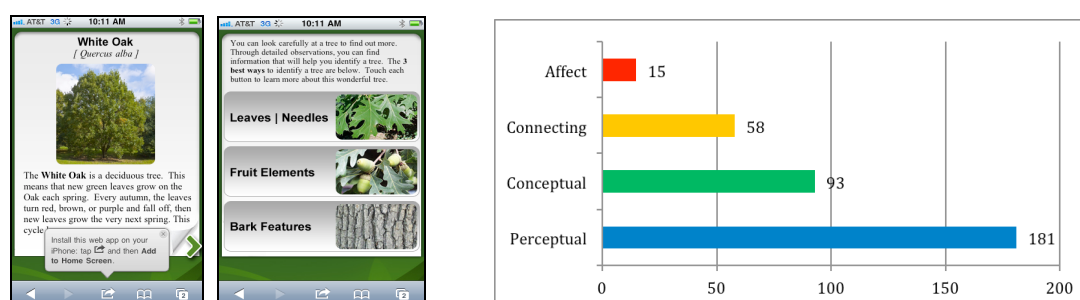
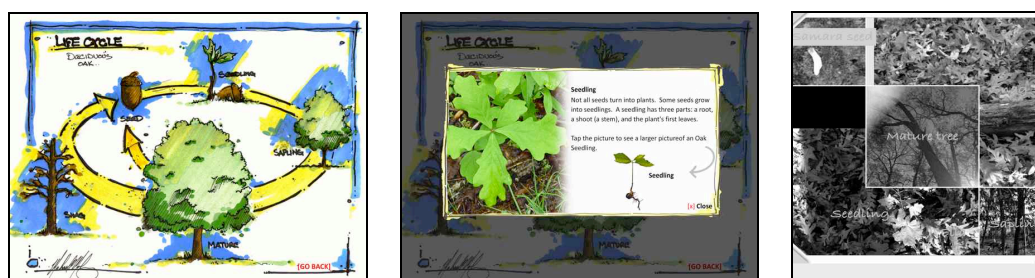


Figure 1. The Iteration 1 interface for the white oak. Figure 2. Chart of types of talk from Iteration 1 families.

Iteration 2: Conceptual Organizer Supporting Observations and Understandings

Iteration 2 was designed to address limitations identified from Iteration 1 and to expand the focus to life cycle elements due to families' observed interests. In Iteration 1, observations of characteristics of trees in the Arboretum (e.g., a pine cone that was open with its seeds dropped vs. a closed pine cone on a tree) often led to discussions of life cycle concepts. Thus, Iteration 2 focused on: (a) providing a graphic organizer (Quintana et al., 2004) of trees' life cycle processes (see Figure 3); (b) indexing science content to local, indigenous trees (Semken, 2005); and (c) including a generative task (Land, Smith, & Zimmerman, 2013) whereby participants used the photographic capabilities of the iPad to document the various parts of the life cycle processes in an old-growth forest (see Figure 5 for a participant's life cycle collage created at the Arboretum).



Figures 3, 4 and 5. Conceptual organizer screen (left); Conceptual organizer seedling (middle); Participant's life cycle collage created onsite with photographs taken at the Arboretum using an iPad (right).

We report Iteration 2 findings from three data sources: (a) an assessment of declarative and conceptual knowledge, (b) photo collage artifacts developed by each participant, and (c) videos of participants interacting at the Arboretum and a final interview. The participating children were given an 8-item assessment of life cycle facts and concepts, provided in an open-ended response format, both before and immediately following the learning activities. Learners received 1 point for each correct answer, and 0 points for an incorrect answer or no response, for a total possible score of 23 points. The pre-test mean score was 4.5 points (standard deviation 2.6)

while the post-test mean was 14.3 (standard deviation 2.7) showing significant improvements ($t = -8.647$, $p < 0.001$) after the exploration (see Table 1).

Table 1: Paired-samples t -test of the pre- and post-test scores

	Mean	N	S.D.	T
Pre-test scores	4.500	10	2.6352	-8.647***
Post-test scores	14.300	10	2.7101	

*** $p < .001$

For the photo collages, we analyzed artifacts and videorecords including: (a) the processes learners used to create the photo collages, (b) elicitation interviews with the children about the photo collages, and (c) the actual photo collages created by the learners. An analysis of the artifacts showed that participating families accurately applied what they had learned during the Tree Investigator activities to identify exemplars of the various life cycle stages in the old growth hardwood forest. In the video analysis, we documented that families shared their life cycle observations aloud with each other and with other families. For example, a mother and daughter who found an oak seedling offered the seedling to other families to include on their photo collages. Overall, we found that the process of creating the photo collage artifacts was a collaborative endeavor between children and adults, child and child, and or child and naturalist. Some children consulted the naturalist to ensure that they were photographing a sapling versus a seedling, as they observed tree species in the woods that they had not encountered during learning with mobile app (which focused on the oak tree). This sapling and seedling distinction was a difficult conception for children in the debriefing interviews, with many children simply stating that the sapling was a “bigger” or “older” tree than the seedling. (The seedling/sapling distinction was also a challenge on the knowledge assessment.) While a few science vocabulary words proved to be difficult for some children to appropriate, the children consistently explained the life cycle conceptually during their debriefing interview. For example, during their interviews, often children used general terms to describe the tree life cycle—such as the seed grew in steps to become a grown tree capable of growing flowers and seed and then trees eventually died. All participating children were able to take photographs and create a life cycle collage; in fact, two learners took over 100 additional photographs while on-site during their visit. These two children did engage in learning tree life cycle concepts even while taking these extra photographs.

Discussion

As indicated by our Tree Investigators Iterations 1 and 2 preliminary findings, mobile devices can be used to deliver science content, support families’ scientific talk, provide access to related knowledge not always visible in a place via augmentations, and create artifacts on-the-go through mobile apps. Across the Iterations of Tree Investigators research, we found that mobile devices enabled the engagement with actual specimens at the Arboretum. Researchers have expressed concern about “heads-down” technologies (e.g., Hsi, 2003), where learners spend time looking at the computer, rather than engaged with the scientific phenomenon. Here, we found that the images and text, when supported by a naturalist, encouraged visitors to engage with the trees and other plants around them. Place-based educational aims (e.g., Semken, 2005) that guided our design were successful in using the specific examples at the Arboretum as exemplars, as evidenced by the fact that learners looked for and found seeds, seedlings, saplings, mature trees, and snags (as well as fallen dead trees).

Conclusion

The contribution of our design-based research study is in informing technologically-enhanced designs for learning outside of school. While small in scope, this study suggests that place-based pedagogical efforts that utilize mobile devices to support informal science education can enhance families’ learning experiences in the outdoors. We advocate for additional research, based on the results from our exploratory study, on how mobile technologies can be used by families and other learners in out-of-school venues in relation to the creation of photographic artifacts, the need for scaffolding (Yoon, et al, 2012), and the role of different app structures.

References

- Allen, S. (2002). Looking for learning in visitor talk: A methodological exploration. In G. Leinhardt, K. Crowley, & K. Knutson (Eds.), *Learning conversations in museums* (pp. 259-304). Mahwah, NJ: LEA.
- Chen, Y. S., Kao, T. C., & Sheu, J. P. (2005). Realizing outdoor independent learning with a butterfly-watching mobile learning system. *Journal of Educational Computing Research*, 33(4), 395-417.
- Dunleavy, M., & Dede, C. (2014). Augmented reality teaching and learning. In M.J. Bishop & J. Elen (Eds.), *Handbook of Research on Educational Communications and Technology* (4th ed., Volume 2), pp. 735-745. New York: Macmillan.

- Gruenewald, D. A. (2003). The best of both worlds: A critical pedagogy of place. *Educational researcher*, 32(4), 3-12.
- Klopfer, E. (2008). *Augmented Learning: Research and Design of Mobile Educational Games*. Cambridge, MA: MIT Press.
- 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.
- Land, S. M., Smith, B. K., & Zimmerman, H. T. (2013). Mobile technologies as tools for augmenting observations and reflections in everyday informal environments. In (eds.) J. M. Spector, B.B. Lockee, S.E. Smaldino, & M. Herring. *Learning, problem solving and mind tools: Essays in honor of David H. Jonassen*. (pp. 214-228). New York: Routledge.
- Leinhardt, G., & Crowley, K. (1998). Museum learning as conversational elaboration: A proposal to capture, code, and analyze talk in museums. Report available at <http://mlc.lrdc.pitt.edu/mlc>
- Lim, M., & Calabrese Barton, A. (2005). Science learning and a sense of place in a urban middle school. *Cultural Studies of Science Education*, 1(1), 107-142. doi:10.1007/s11422-005-9002-9
- Liu, T.-C., Peng, H., Wu, W.-H., & Lin, M.-S. (2009). The effects of mobile natural-science learning based on the 5E learning cycle: A case study. *Educational Technology & Society*, 12(4), 344-358.
- Quintana, C., Reiser, B., Davis, E., Krajcik, J., Fretz, E., Duncan, R., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, 13(3), 337-386.
- Rogers, Y., Price, S., Fitzpatrick, G., Fleck, R., Harris, E., Smith, H., Randell, C., Muller, H., O'Malley, C., Stanton, D., Thompson, M., & Weal, M. (2004). Ambient Wood: Designing new forms of digital augmentation for learning outdoors. *Proceedings of the 2004 Conference on IDC*. (p. 3-10).
- Semken, S. (2005). Sense of place and place-based introductory geoscience teaching for American Indian and Alaska Native undergraduates. *Journal of Geoscience Education*, 53(2), 149-157.
- Sobel, D. (2004). Place-based Education: Connecting Classroom and Community. *Nature and Listening*, 4. <http://www.antiochne.edu/wp-content/uploads/2012/08/pbexcerpt.pdf>
- Smith, G. (2002). Place-based education: Learning to be where we are. *Phi Delta Kappan*, 83, 584-594.
- Squire, K., & Klopfer, E. (2007). Augmented reality simulations on handheld computers. *Journal of the Learning Sciences*, 16(3), 371 - 413.
- Squire, K. D., & Jan, M. (2007). Mad City Mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of Science Education and Technology*, 16(1), 5-29.
- Tan, T. H., Liu, T. Y., & Chang, C. C. (2007). Development and evaluation of an RFID-based ubiquitous learning environment for outdoor learning. *Interactive Learning Environments*, 15(3), 253-269.
- Tan, E., & So, H. J. (2011). Location-based collaborative learning at a geography trail: Examining the relationship among task design, facilitation, and discourse type. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Proceedings of the 2011 CSCL Conference* (p. 41-48).
- Tzou, C., Scalone, G., & Bell, P. (2010). The role of environmental narratives and social positioning in how place gets constructed for and by youth. *Equity & Excellence in Education*, 43(1), 105-119.
- 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
- Yardi, S., & Bruckman, A. (2012). Income, race, and class: exploring socioeconomic differences in family technology use. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (pp. 3041-3050). ACM. Doi: 10.1145/2207676.2208716.
- Yoon, S. A., Elinich, K., Wang, J., Steinmeier, C., & Tucker, S. (2012). Using augmented reality and knowledge-building scaffolds to improve learning in a science museum. *International Journal of Computer-Supported Collaborative Learning*. doi:10.1007/s11412-012-9156-x
- Zimmerman, H. T., Land, S. M., McClain, L.R., Mohny, 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.

Acknowledgement

We appreciate our partners that made this learning research possible: the Arboretum at Penn State, Shaver's Creek Environmental Center, 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 colleagues in the Augmented and Mobile Learning Research Group (sites.psu.edu/augmentedlearning/). We express our gratitude to the family participants for joining our project.

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

Iolanda Garcia, Universitat Oberta de Catalunya, C/Roc Boronat, 117, Barcelona 08018, Spain,
igarcia@uoc.edu

Abstract: In this paper we report an ongoing research project that aims to develop, apply, assess and accurately document a method for co-designing learning scenarios in higher education with the participation of teachers and students. The theoretical and empirical background is described and a model for the study and analysis of students' participation in co-design processes is proposed. Finally, the main implications, challenges and potential contributions of the application of this model are discussed.

Introduction

The study of student participation and consideration of their perspective in planning and curriculum design has been approached from different perspectives. There is, first, a line of work on "student engagement", dedicated to the analysis of proposals based on encouraging active involvement and commitment on the part of students. It is assumed that offering opportunities for students to take control of their learning activities, reflect and become aware of what they do to learn, may favor the adoption of deeper learning approaches (Bain & Zimmerman, 2009). Another line of work focuses on the study of "student voice". This is both a theoretical approach and a set of practices that place students as active agents in the analysis and review of teaching and learning proposals. From this perspective, students' perception and experience about their learning is recognized as unique and essential to the improvement of education (Bovill, Cook-Sather & Felten, 2011). A third line of study, more recent and more focused on the field of higher education, is framed under the expression "students as producers" also referred to as "students as researchers". Neary (2012) refers to this approach as way of turning students into key agents of a radical transformation of higher education institutions. Although the approach may have different concretions, one of those most widely taken is the use of academic research as a core undergraduate curriculum, involving teachers and students as partners in academia's function of knowledge production.

The literature developed in the last decades around the issue shows that, although there is a clear stream of research that supports the need to address learning in universities from the shared responsibility of teachers and students, examples of transformation of concrete contexts of practice are almost anecdotal and confined to specific projects. Although the literature puts forward a number of reasons for student involvement in curriculum design, there is still little systematic evaluation of their real impact and specific dynamics (Bovill, Cook-Sather & Felten, 2011).

Moreover, several authors highlight the risks of participatory approaches, and even maintain a critical stance regarding the way "student voice" has been interpreted and represented, particularly in the field of higher education. According to Seale (2010), in the discourse of these approaches there is the implicit assumption that the students' feedback will have a transformative impact on teaching practices and curriculum development. However, these expectations of transformation are not articulated or explicit, nor do they sufficiently describe how this type of participation can be realized to achieve the mentioned purposes. Finally, the same author asserts that the lack of clarity about the real commitment towards students' participation and empowerment is reflected in the way "student voice" is defined and the roles assigned to students in some projects. Based on her review, Seale (2010) argues for the need to create a framework to connect theory and practice from which to articulate the implementation of concrete actions so that the "student voice" can really have an impact on universities. According to this author, participatory research methods have this potential.

In this paper we report an ongoing research project that aims to develop, apply, assess and accurately document a method for co-designing learning scenarios with the participation of teachers and students, including its articulation, roles, components, phases, conflicts and turning points. In the following pages the theoretical and empirical background is described and a model for the analysis of co-design processes is proposed. Finally, the main implications, challenges and potential contributions of the application of this model are discussed, as well as its relevance within the framework of the 2014 ICSL Conference.

Contextualization and Rationale of the Project

Although the idea of student participation in curriculum design is not new and has been developed mainly in primary and secondary education especially in the USA, it continues to have a very modest and partial scope in the field of higher education. Existing literature is limited and tends to focus on the description of varied experiences rather than on the abstraction and evaluation of specific methodological parameters, so that progress in knowledge on the subject is still not evident (Bovill, Morss & Bulley, 2008; Seale, 2009). However, all the latest pedagogical literature provides arguments in favor of the idea of actively involving students in curriculum

design, either from the perspective of enhancing the development of their critical judgment, increasing their commitment and responsibility for their own learning, enabling more authentic and meaningful learning experiences, improving the options to personalize learning, or understanding the very idea of curriculum as a co-creation task between teachers and students (Bovill, Morss & Bulley, 2009). Below we provide a summary of some of the most relevant results from the study of university students' participation in curriculum design.

Evidence on Students' Participation in Curriculum Design

Seale (2009) carries out a comprehensive review and accounts for the heterogeneous ways of approaching and materializing the study of "student voice" in higher education. According to this author, there are two purposes that are mainly linked to the student voice projects in higher education: quality enhancement and assurance and staff or professional development, which, in turn tend to be aligned respectively with higher education policies on assessment (evaluation and feedback) and teachers' reflective practice.

The most widespread form of university students' participation is through giving feedback on courses through means such as staff-student liaison committees, questionnaires and surveys, focus groups and electronic voting systems (Bovill, Morss & Bulley, 2009), although the real impact of this kind of procedures on the design of courses is not clear. Some authors have proposed methodologies in which students and teachers participate in "co-generative" discussions, aimed at sharing their views and reflecting together on how to improve the practice of teaching and learning (Roth, Lawless & Tobin, 2000). There is less evidence of the procedures used in the exploration of the direct involvement of students as "co-designers" in the field of higher education. Cameron and Tanti (2011) report the methods utilized in a project where students were asked to take an active role in planning and creating their own learning tasks through a problem-based learning approach supported by social media tools. Bovill, Cook-Sather and Felten (2011) describe a multiple case study of students' participation in course re-design conducted in different universities in the USA and UK. Those experiences involved different strategies such as the creation of small "course design teams" (CDT) formed of teachers and students that co-create or re-create a course syllabus; the commissioning of a small group of students to design a new virtual learning environment based on case studies, also including the production of written, audio and video resources; and engaging a whole group of students in decision making over the actual content of the course according to what they considered they needed to learn. Finally, Bovill, Morss and Bulley (2009) propose as an alternative way of participation, enabling students to participate in the design process at a later stage in order to capitalize on their experience of a course.

Regarding the conditions and factors that may determine or affect a co-design process, it seems clear that time is a key factor. A sustained dialogue over time is required to generate the context of trust needed to deconstruct mutual prejudices between students and teachers and allow them to express their ideas clearly. It also takes time and external facilitation for students to come to understand and use certain pedagogical concepts. There seems to be a turning point from which the co-design process becomes more fluid and rich, when students understand that their views will be taken into serious consideration (Bovill, Morss & Bulley, 2009).

Cameron and Tanti (2011) highlight the role that technology and specifically social media can play in facilitating these processes, to enable students to create, discuss and publish new knowledge objects easily and quickly (Chang, Kennedy & Petrovic, 2008). In this sense, the OER movement offers plenty of opportunities to put into practice the concept of "students as producers" by involving them in activities to inform and enhance learning designs and to produce and improve learning materials (Greaves, 2012).

But the reviewed studies also found points of tension that may hinder the co-design process. According to Cameron and Tanti (2011), for students to assume the role of co-designers, teachers must relinquish some control. They need to be more self-aware, flexible and knowledgeable to respond to student learning needs and to accept their opinions and demands. There are also institutional factors that can constrain student participation, such as professional requirements or regulatory frameworks (Bovill, Morss & Bulley, 2009).

Moreover, participatory approaches have a number of drawbacks if they are not genuine or not proposed in an appropriate manner. Sometimes student participation is treated as an aesthetic and superficial issue, limited to spaces too specific and timely to have a real impact on the learning design process. There is a risk of falling into the objectification and generalization of the perspective of students, ignoring their diversity of profiles and underlying motivations. Students used to maintaining a relationship under the authority of teachers may show resistance, fears and even feel manipulated when facing such proposals (Bovill, Cook-Sather & Felten, 2011). According to Seale (2011), some student voice projects are teacher-centric, in the sense that they put students in the role of teacher, rather than empower them from their position and enable equal and constructive dialogue.

Theoretical and Methodological Approach

The project discussed here, based on the use of co-design methods, takes as a reference the revised work on "student voice", "student engagement" and "students as producers", to propose the direct and active participation of students in the process of co-designing learning scenarios they will take part in. In addition, it combines the

methodologies and tools of the domain of learning design with those of the participatory design field. Roschelle, Penuel & Shechman (2006:606) describe co-design as "*a highly-facilitated, team-based process in which teachers, researchers and developers work together in defined roles to design an educational innovation, realize the design in one or more prototypes, and evaluate each prototype's significance for addressing a concrete educational need*". According to these authors, the co-design process is not completely democratic –as it is in participatory design– because researchers often hold the ultimate responsibility for the quality decision-making. Many of the co-design practices in education are associated with the design and implementation of technological tools to support learning processes (Mor & Winters, 2006; Penuel, Roschelle & Schechtman, 2007). In general, it is about funneling educational innovations that currently involve the use of technological resources in the school context. These experiences have typically involved teachers, researchers and often software developers, as partners in the process of co-designing educational innovation. This procedure ensures the connection and the orchestration of the theory, practice models, tools and participants' insights.

The project aims to study the development of learning scenarios that are more authentic, contextualized and focused on learners, through a co-design process involving students and teachers in the negotiation of the design principles of such scenarios to assess the potential of this approach as a catalyst for change and innovation in higher education. The study applies the methodology of design-based research. The design of the investigation is iterative, situated and intervention-led but underpinned by theory. The object of study is therefore the very process of co-designing, taking as key agents both the teachers and the students to whom those practices are addressed. A mixed approach (quantitative and qualitative) is used for data collection, analysis and interpretation. The research design involves several cycles of iterations with the aim of refining the process of co-design. The research questions that address the issue of students' participation are as follows:

- a. How are participants' roles and levels of intervention negotiated, assigned and managed?
- b. What are the stages and critical issues to consider in the process of co-design?
- c. What are the effects of the co-design process on teachers and students' perception of learning, teaching and learning design?
- d. How does context (university model, area of knowledge, profiles) influence the co-design process?

The participants are a group of teachers and students from two universities with different models, one of them blended (University of Barcelona) and the other virtual (Open University of Catalonia). Of the four mentioned contexts reported in this study, two are from the blended university and two are from the virtual one. The four design contexts correspond to different disciplinary areas, such as medical informatics, economics, tourism and communication. The co-design of a learning scenario for each context of practice aims to ensure that results are accessible, acceptable and useful to all participants, and can be effectively used to evaluate, report and improve practice in these and other contexts (Anderson & Shattuck, 2012).

The co-design process takes place cyclically, through the following four main stages: exploration, envisionment, operationalization and assessment and reflection. The first phase of the co-design process consists of a series of participatory workshops with six teachers from the four selected practice settings, along with members of the research team. These sessions have the function of introducing participants to the dynamics of co-design and gaining a deeper understanding of the contexts of practice. They are also oriented to designing learning scenarios based on inquiry-based learning and technology-enhanced learning principles, involving the prototyping of the designed learning scenarios. In the second phase of the co-design process, two students selected from each context join the design work groups (a total of eight students) with the aim of critiquing the prototypes of the learning scenarios designed by teachers, bringing their own ideas and perspectives and thus validating the final designs. This is done through the same dynamic of participatory workshops used in the first phase but adding the implementation of the learning scenarios in the corresponding real contexts of practice. In this way the prototypes of the learning scenarios are tested and feedback is collected from the entire student group. Implementation occurs in iterative cycles that enable the designed scenarios to be refined based on collected feedback, reflection and discussion by the design teams. Thus, it addresses two levels of student participation: a) students who are part of the co-design team of each context, b) students who make up each of the four student groups where learning scenarios are implemented.

Throughout this entire process exhaustive data collection is performed using several research instruments: initial interviews with teachers and students, participant observation and audio recordings of joint work sessions, short questionnaires addressed to both teachers and students after each work session, classroom observation and field notes, a questionnaire to the whole student group, final interviews with teachers and discussion groups with students. The purpose is to thoroughly keep track of the co-design process, enabling the research questions that were posed to be answered. To do this, an analytical model has been developed that integrates all the meaningful dimensions for studying the co-design according to the literature review and research questions. Specifically, the relevant dimensions for the analysis of students' participation reported in this paper are: *methodological aspects of co-design, roles, dialogue and negotiation, phases, conflicts and key issues, student and teacher profiles, students and teachers' change, institutional and curricular factors*.

These dimensions are discussed from four perspectives or different viewpoints that reinforce each other, as shown in Figure 1: joint work sessions, classroom implementation, perceptions and design product. It is a holistic model of analysis of the co-design process, which in turn allows the triangulation of data and methods, in order to preserve the research trustworthiness. For the analysis of qualitative data (interviews, observations, field notes and post-session questionnaire) the constant comparative method and grounded theory (Glasser & Strauss, 1967) is used, while for the quantifiable data collected with the post-session questionnaire and the general questionnaire applied to the whole students group, a descriptive statistical analysis is performed.

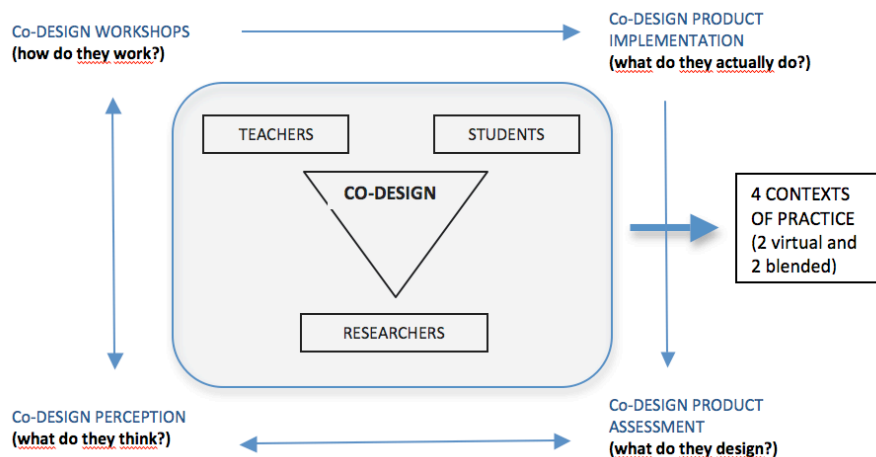


Figure 1: Co-design process analysis framework

Design-based research is coupled with the use of participatory design methods. The principle of performing various iterations allows students to participate in all stages of the process even though they are incorporated into a second phase. Teachers and students work in close contact and are placed at the same level through the entire process of co-design, although the former will have the ultimate responsibility for the scenarios designed. Similarly, teachers and students assume the role of co-researchers to engage in the production and analysis of collective research knowledge and enhancement of the group to solve the problems identified (Fielding, 2004). However, the research team is ultimately responsible for the co-design process.

Discussion on Opportunities for Productive and Innovative Inquiry

The research presented is intended to advance the study of participatory and collaborative learning design methods that involve teachers and students in the field of higher education. With this goal, inquiry-based and technology-enhanced learning scenarios will be co-designed. It is considered that the design of learning scenarios –including the sociocultural context in which they are framed, the chosen pedagogical approach and the objects that make up the learning situation– in which learning activities are inserted, can elicit the processes intended to be facilitated and promoted among students. Moreover, the co-design process itself is proposed as a learning experience. In the case of teachers it is viewed as an opportunity for professional development and in the case of students it is expected to cause a change in their conceptions, understanding and attitude towards learning from the perspective of learning design. Therefore, this research is based the following assumptions:

- students' participation in the co-design process can integrate their perspective, interests and needs more effectively into the designed learning scenarios, ultimately promoting deeper learning by different student profiles and in different learning contexts;
- participation in the co-design process can make students more aware of the processes involved in learning and learning design, and become more autonomous in the direction of their learning activities;
- co-design processes involving students and teachers in collaborative negotiation of the design principles applied to learning scenarios can facilitate the adoption of an inquiry-based learning model mediated by a more mature and autonomous use of technology by students;
- students' participation in the co-design process is expected to encourage teachers to reflect on their practice. Similarly it is expected that the supporting co-design methods and instruments will help participant teachers to develop awareness and skills as learning designers.

Conclusions and Relevance of the Topic

Involving students in learning design processes is a complex task because it implies profound changes in how to approach teaching and learning in higher education. However we consider it indispensable to walk this path in order to offer learning scenarios that are more authentic, contextualized and meaningful to current students.

We have shown the positive implications of the adoption of this approach, both in terms of improving the quality of learning, changing roles and relationships between teachers and students, and as a catalyst for teachers' professional development processes. The absence of clear and detailed references on the implementation of such participatory processes hinders its application and its generalization. For this reason, we believe that the adoption of a design-based research approach could be useful to faithfully capture the co-design experience in a real context and simultaneously feed the knowledge in this field.

If the goal of universities is to train autonomous, critical, creative students who are able to develop themselves throughout their lives they must begin to adopt the vision of students as producers. To do this, it seems essential to empower students and to provide them with the agency to truly influence and bring about change in the university. Models of inquiry-based learning combined with the use of social media offer many opportunities in this regard, which we intend to explore in this research. In turn, teachers have a crucial role to play in facilitating this process, as well as in the transformation of learning design proposals in higher education.

From our point of view the subject of the reported research is closely related to the theme chosen for the 2014 ICLS Conference and would be an appropriate and relevant contribution to it. The focus of the inquiry is in fact the learning experience of teachers and students based on the co-design process. We believe that co-design processes are of special richness because they are located at the intersection between different types of practices: sociocultural practices of collaboration between participants and the actual practices of teaching and learning in institutional and epistemic contexts.

References

- Anderson, T. & Shattuck, J. (2012). Design-Based Research: A Decade of Progress in Education Research? *Educational Researcher*, 41, pp. 16-25.
- Bain, K., & Zimmerman, J. (2009) Understanding great teaching. *Peer Review*, 11 (2), pp. 9-12.
- Bovill, C., Cook-Sather, A. & Felten, P. (2011) Students as co-creators of teaching approaches, course design and curricula: implications for academic developers. *International Journal for Academic Development*, 16 (2). pp. 133-145.
- Bovill, C., Morss, K. & Bulley, C.J. (2009) Should students participate in curriculum design? Discussion arising from a first year curriculum design project and a literature review. *Pedagogic Research in Maximising Education*, 3 (2). pp. 17-26
- Cameron, L. & Tanti, M. (2011) Students as learning designers: using social media to scaffold the experience. *eLearning Papers*, 27.
- Chang, R., Kennedy, G. & Petrovic, T. (2008). Web 2.0 and user-created content: Students negotiating shifts in academic authority. *Proceedings ascilite Melbourne 2008*.
- Fielding, M. (2004) Transformative approaches to student voice: theoretical underpinnings, recalcitrant realities, *British Educational Research Journal*, 30(2), pp. 295–311
- Glasser, B. & Strauss, A. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Chicago: Aldine Publishing Company.
- Greaves, L. (2012) Feed-forward for Informed Learning (FfIL): learner-generated materials for personalised learning. *Enhancing Learning in the Social Sciences* 4(3).
- Mor, Y., & Winters, N. (2007), Design approaches in technology enhanced learning, *Interactive Learning Environments*, 15 (1), 61-75.
- Neary, M. (2012) Student as producer: an institution of the common? [or how to recover communist/revolutionary science] *Enhancing Learning in the Social Sciences* 4(3).
- Penuel, W.R., Roschelle, J. & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process. *Research and Practice in Technology Enhanced Learning*, 2(1), pp. 51-74.
- Roth, W.M., Lawless, D.V. & Tobin, K. (2000) {Coteaching | Cogenerative Dialoguing} as Praxis of Dialectic Method. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(3), Art. 37.
- Roschelle, J., Penuel, W. R. & Schechtman, N. (2006). Co-design of innovations with teachers: Definition and dynamics. Paper presented at the *International Conference of the Learning Sciences*, Bloomington, IN.
- Seale, J. (2009) Doing student voice work in higher education: an exploration of the value of participatory methods, *British Educational Research Journal*, 36 (6), 995-1015.

Acknowledgments

The project is funded within the R+D National Plan, by the Ministry of Economy and Competitiveness of Spain. We also would like to thank the professors and students participating in the research for their involvement and collaboration.

Authenticity Matters: Youth and Science Participation in Design-Based Learning Environments

Gavin Tierney and Déana Scipio, University of Washington, 1100 NE 45th Street, #200, Seattle, WA 98105
gtierney@uw.edu, descipio@uw.edu

Abstract: This paper reinterprets the notion of the “competent outsider” (Feinstein, 2010) through the lens of authenticity. We present a three-pronged conceptual framework about authenticity: to the domain, to the world, and to students. Data from Project COOL and the Knowledge In Action Project raise questions about power and ownership of “authentic” science. Findings suggest considering these multiple aspects of authenticity can help in design and instruction of formal and informal science learning environments.

Major Issues Addressed

When educators and designers of learning environments ask kids to take on the task of solving problems in Science Technology Engineering and Mathematics (STEM) disciplines, we are also asking them to choose to participate in particular ways. The ways that we frame problems for youth, the tools we offer them, the authority we grant, and the ways we hold them accountable are all influenced by our ideas of what it means for students to be or become a version of themselves that can participate in the real world. These ways of participating are housed within the domain and linked to a variety of disciplinary practices, thus when we think about authentic disciplinary practices we need to consider the types of authentic problems, practices, and identities that we are asking students to take on/ become. Learning in practice raises questions about the kinds of practice-linked identities that youth are encouraged to take on (Eisenhart, 1996; Nasir, 2002; 2013).

Youth might understand themselves as doing science practices but still not feel they are being scientists (Barab & Hay, 2001). van Eijck & Roth (2009) traced the trajectory of a young man from the Salish tribe to explore the impact of “authentic science experiences” on his scientific literacy development as well as his science career choices. Even though he participated in experiences that would be considered authentic based upon disciplinary and scientific measures, he did not persist in the field. In a related study (van Eijck, Hsu & Roth, 2009) the researchers while exploring the relationship between youth’s experiences of science practices and how they translated them into images of science practice, discovered that engagement in science practices can serve to calcify youth’s images of science. A longitudinal study conducted in the UK around youth identity and STEM participation highlighted a “mismatch between popular representations of science, the manner in which it is taught, and the aspirations, ideals, and developing identities of young adolescents” (Archer et al, 2010). This paper approaches this mismatch between science practices and students lives through the lens of authenticity, looking at authenticity in science, authenticity in the real world, and in students’ worlds. This paper leverages data from two empirical studies to answer the research question: How do students’ perceptions of authenticity within project-based curricula influence their disciplinary participation?

Potential Significance

Understanding the ways that youth experience authenticity in their lives is an important consideration for designers of learning environments. Our data suggests strong links between youth engagement, sense of authenticity, and choices to increase or sustain participation. These findings raise new questions for educators and designers of learning environments that seek to involve youth in STEM practices. This raises many questions: To what end? To whose end? What is the goal of STEM education? Is it transfer into everyday practices? Are we trying to create scientists, or competent outsiders? Feinstein (2010) argues for a re-imagination of the utility of science, based in the everyday lives of students. Feinstein calls scientifically literate citizens “competent outsiders” and advocates for student voice and connections to relevant everyday settings. We leverage Feinstein’s ideas and reinterpret them through the lens of authenticity, considering authenticity of problems, practices, roles, and settings. We offer a three-pronged framework for thinking about authenticity: authenticity to the domain, authenticity to the world, and authenticity to students. We suggest ways, through examples from data, in which these three types of authenticity can be used in educational design and instruction.

Theoretical Framework

To explore the ways youth participate in authentic project-based curricula we have used Productive Disciplinary Engagement (PDE) (Engle & Conant, 2002) as our theoretical lens. Engle and Conant define PDE as making intellectual progress or getting somewhere (productive), a connection between what the students are doing and the practices and discourse in the discipline (disciplinary), and making substantial coordinated contributions that include emotional displays and spontaneous re-engagement over time (engagement) (Engle & Conant, 2002, p. 402). PDE is linked to authenticity in that when students are productively engaged in the discipline, they are

pursuing the same authentic practices and tasks as disciplinary experts. In our analysis we specifically look at youth conceptions of what constitutes productivity. The discipline of science has defined and evolving definitions of what is productive in the discipline. We were interested in exploring students' goals, and thus their concepts of productivity, as they engaged in the discipline.

By exploring students' concepts of productivity, we can also see what is valuable, relevant, and authentic to students. In this paper we suggest that authenticity is not a property of a learning space but rather an ever-changing practice around which components of social learning can be organized. This is true for all three prongs of authenticity we consider in this paper. Just as the discipline is in constant change, so to is the world at large and students' worlds, meaning that what is authentic to any of these domains is also in flux. We conceptualize authenticity to students as in what is considered authentic and relevant to them. Though a nuanced point, we believe that authenticity to students requires projects or activities to leverage students' out-of-school lives, but not necessarily integrate their out-of-school lives into the projects. Previous research on contextualized scenario-based learning has demonstrated ways in which students can feel an activity or simulation is authentic because of the authentic roles they are given or the authentic performances required at the end (Cognition and Technology Group at Vanderbilt, 1992; Rivet & Krajcik, 2008). Put another way, we consider authenticity to students as what 'feels real' to students, looking at the ways this form of authenticity connects to what disciplinary experts believe is authentic to the discipline and what teachers, mentors, and designers believe is real to the world at large.

Project Descriptions

The two projects in this analysis are both design-based research projects which leverage project based learning. The Knowledge In Action Project is a project-based Advanced Placement (AP) Environmental Science curriculum. Chemical Oceanography Outside the Lab (COOL) is a self-choice, afterschool chemical oceanography program for middle school youth. Designed environments focused on project-based learning have particular affordances for being and becoming in practice. The two projects for this analysis were chosen because both offer students opportunities to participate in projects and activities that are authentic to the Science discipline and the real world. The two projects were also chosen to represent a school setting (Knowledge In Action) and an out-of-school setting (COOL). Finally, the projects were chosen so we could explore authenticity across a larger sample of qualitative data (Knowledge In Action) and a single case study (COOL).

AP courses are typically known for the vast amount of information that students are asked to learn for the AP test taken at the end of the year. The Knowledge In Action Project proposed that by putting students in roles and asking them to tackle challenging, authentic projects that they would do as well or better on the AP test and, compared to conventional AP courses, would connect what they learned to their out-of-school lives.

COOL is a design-based research study focused on approaches to broadening participation in the geosciences for youth from groups traditionally underrepresented in STEM disciplines. COOL proposed that hybrid learning environments designed to have youth and STEM mentors work collaboratively would provide youth with models of participation and create opportunities for explicit conversations between youth and mentors about ways of navigating into new discipline-linked identities. In this learning environment, mentors' interactions with youth occupy a hybrid space between academic and interpersonal relationships. COOL sought to leverage these hybrid relationships to bring youth and mentors together to engage in authentic STEM practices. Here authenticity means projects that simultaneously hold personal, community, and disciplinary relevance for all participants.

Data and Analysis

Data for this paper were pulled from larger Design-Based Implementation Research projects on student learning and engagement. Specific data for this analysis were chosen based on previous work in each project. We chose data that highlighted the link between authentic contexts or activities and student participation. The Knowledge In Action Project used data from youth individual interviews, youth group ("fishbowl") interviews, and teacher interviews across three years. Project COOL used data from youth and mentor interviews and video observations. All data Video and audio records were transcribed verbatim and then collected in a hermeneutic unit in computer programs for qualitative analysis. There were two distinct phases of analysis. In the first phase, research teams on each project utilized a grounded theory approach, focusing on student engagement and identification with the discipline as foundations for beginning codes. Then each research team included an initial set of code categories based on each research group's research questions, including codes around *engagement*, *identity*, *transfer*, *disciplinary thinking*, *identification with the domain*, *positioning*, *stance*, and *citizenship*. Coding was iterative and collaborative, with research group members proposing new codes and code categories, negotiating codes and their definitions, and co-producing analytic comments and memos. Analysis proceeded until no new codes were needed to characterize the data. Patterns for The Knowledge In Action Project included student reports of transfer, proposed solutions to environmental science problems, and identification with certain disciplinary practices. Patterns for Project

COOL included deepening participation in STEM practices, identification with the domain of science, and sharing chemical oceanography information with new communities i.e. home and school.

In the second phase of analysis, the authors reanalyzed sections of data across the two projects through the lens of productive disciplinary engagement and the three prongs of authenticity. We analyzed the projects proposed goals and curriculum for opportunities for authentic practice. We then identified moments in the data where the participants talked about their perceptions of the authentic opportunities provided them. These moments were coded using open and theoretical codes based on Productive Disciplinary Engagement (Engle & Conant, 2002). Patterns in the data were explored using discourse analytical techniques and, when applicable, connecting perceived patterns with video of youth participation in the Project COOL. Discourse analysis included exploring language around agency and identity and the contexts and scenarios referenced during those instances. Patterns across data sets included students' perception of authenticity of tasks extending beyond tasks only being authentic to the discipline. Tasks that were authentic to the real world and students' worlds increased perceived relevance and identification with certain disciplinary practices.

Findings

We present data from the two projects, exploring the ways in which authenticity to the domain, to the real world, and to students' worlds were present and how these forms of authenticity impacted the students' STEM engagement and identification. The Knowledge In Action Project highlights student responses to a redesign of an AP Environmental Science unit that asked students to use authentic disciplinary practices and thinking to address real world problems within the context of their home lives. Project COOL was chosen as a comparison because it presents a close analysis of the mechanisms of sense of authenticity as it related to the student's identification with the domain of chemical oceanography. In both projects students use practices and thinking that designers saw as authentic to the scientific disciplines. Both projects also asked students to tackle authentic real world problems around human's impact on the environment. Finally, students from both projects saw the project activities as authentic to their worlds as students. In The Knowledge In Action Project students used data from their own lives to address the real world problem of sustainability. In Project COOL the student saw the activity as personally authentic because it was a performance that mattered. These two cases, taken together offer a broad picture (APES) and a close analysis (COOL) of the impacts of sense of authenticity on student's participation.

Authenticity in the Real World — Authenticity in Students' Worlds

In The Knowledge In Action Project AP Environmental Science (APES) course the projects begin locally, with students analyzing environmental science issues in their own lives. The projects then grow in scope, until, at the end of the year, the students are looking at environmental science on a global scale. In contrast to more hard sciences, Environmental Science focuses far more on citizen science or how to be a scientifically informed sustainable citizen. This is particularly the case in the project-based APES curriculum, which contains a course master question of "How can we live more sustainably?" With a focus on sciences and sustainability in the world, authenticity within the course typically refers to authentic scientific practices and authenticity to the real world. Authenticity to the real world encompasses a number of project qualities: authentic real world roles, authentic challenges, authentic overarching questions (master course question), and authentic real world problems that the students are asked to tackle in the projects.

In previous analysis however (Nolen et al., 2013), the authors have seen that the part of the APES course that is the most memorable to students and has the strongest impact on their daily lives is the one in which they do not take on a role, but analyze their own practices. In the first project of the AP Environmental Science course, titled "Ecological Footprint", the students are asked to analyze the daily habits of themselves and their families. This starts with the students calculating the number of earths that would be needed if everyone on the planet lived the way they do. Then, analyzing different areas of consumption such as transportation or electricity use, the students propose a change to their family.

In our analysis of the first year of the Ecological Footprint cycle we saw the ways in which the project impacted the students practices outside of school. One female student from the second year of the course implementation talked about the relevance to her life:

I feel as if the things that we're learning really relate to me, especially like the mobility analysis, the water usage, and the electricity usage. I can totally see how that relates to me, and how that impacts my life because I can see, 'Oh! This is how much gas I use. This is how much I don't walk (November 14, 2011).

The authenticity real world problem of sustainability connected directly to the student's life and practices. This is contrasted with the first year of the curriculum, where, for the Footprint cycle, students were put in an authentic real world simulation (planning an eco-friendly wedding) that did not relate to their own authentic

lives. One student commented on that version of the cycle: “I didn't see the connection to that and anything that I could take away from that and to the future. It was very irritating” (May 10, 2011). The first year version of the cycle seemingly contained authenticity to the domain, incorporating various scientific thinking and practices and it was authentic to the real world, with the authentic problem of sustainability and the authentic real world context. However, it was not until the second year, when the cycle was redesigned to also include authenticity to the students’ worlds, that the impact of the cycle was felt. A male student in the second year of the project discussed the connections between science practices, the real world problem of sustainability, and his own life.

[In class] we had to write down... a whole list of electrical appliances that we might use, and so we put in how long each day they're used or plugged in. And then we put it in this Excel sheet and got the math of how many kilowatt-hours are being used from that, and then we printed it out and made a pie graph. I immediately showed that to my mom and she was amazed and right off the bat we just started changing the way that we use stuff, and started to unplug stuff when we weren't using them (October 31, 2011).

By connecting authentic citizen-scientist practices with the students’ lives the project became authentic to not just the real world, but also the students’ real world. The Ecological Footprint cycle introduced numerous authentic scientific concepts and practices, such as collecting, analyzing, and interpreting their own data (NAP, 2011), however, the authenticity of the project rested on the reality of issue of sustainability in the world and the students’ day-to-day practices. In thinking about the ways students adopt authentic scientific practices into their daily lives and into their developing ideas of what it means to be a citizen and a scientist, it may be useful to think about the ways authentic scientific practices connect to authentic experiences in students’ lives.

Authenticity in Performance Spaces — Authenticity in Students’ Worlds

In the second year of Project COOL, one of the young women in our study- Kelly (pseudonym) stood out as a case of how sense of authenticity can lead to deepening participation in STEM contexts. As a context COOL gave her opportunities to engage with “real” data, participate in scientific communication, discover new things, and to be trusted in places where her actions could have real consequences. Kelly, commented on the sense of newness she felt within COOL and went on to explain, “the school doesn't teach us this types of stuff and just because a lot of the experiments that we did, they were stuff that you guys didn't know either. Whereas at school we're learning stuff but it's experiments and the results that the teacher already knows” (2011). Kelly mentioned a field trip we took to the ocean sciences department of a local university in her follow up interview demonstrating that place was also an important component of her sense of authenticity: “Especially the sterile area it shows like, like how, how little you can do to mess up to like, the littlest thing can destroy a whole science project over something that people have been working on for months” (Kelly, 2011). For Kelly, visiting the sterile area and understanding that “the littlest thing can destroy a whole science project” was crucial. This was an important place she had been trusted to enter. These experiences helped Kelly develop a sense of authenticity, she felt empowered to step into a leadership role when faced with obstacles to participation. She did so in the face of a challenge by other members of her group, pushing everyone to step up and get their work done for the group’s final project.

I remember I did take charge and say something about what we should do and why everybody was actually there. Like the reason they were there wasn't just to be there for after school. Because if it was just after school, they could have joined another something else or were they there for science and actually learning something (July, 2012).

Kelly described herself as someone who was “there for the science.” She saw her work in COOL as meaningful because her experiences had allowed her to create a storyline of someone who was “there for the science.” Her sense of the actions and discourse as connected to authentic science shaped her perception of what she was learning. It shaped her desire to make sure that the final project was completed. This case study illuminates the ways that a sense of authenticity developed during a young woman’s interactions with a set of designed sociomaterial arrangements (Bell et al., 2012). Kelly’s sense of authenticity within the COOL Program led to deepening participation in STEM practices. The challenge to educators, designers, and theorists is to figure out how to create and sustain sense of authenticity within all learning environments.

Relevance to the Conference Theme

The processes of being and becoming in practice are directly related to the ways that youth experience the connection between the things they are doing and the practices of disciplinary experts in the real world. The ways that youth perceive these connections are sometimes even more important than the observable links that teachers or curriculum designers intend. Our data suggests the possibility that youth’s perceived connections

between their lived experiences and disciplinary practices impact their identification with the discipline, and whether they choose to deepen participation in the discipline. The three prongs of authenticity intersect in important ways as students engage productively with the science domain, developing identities as scientists and citizen scientists. In year one of The Knowledge In Action Project, the course was authentic to the science domain and authentic to roles and problems existing in the real world, however it wasn't until the addition of authenticity to students' worlds that students began reporting identification with the domain and the transfer of practices out of school. Similarly, Project COOL was constructed to be authentic to the domain of science and to the real world, but also to include authentic performances and roles. These three prongs of authenticity worked in concert to help the case study youth participate in the practices of contemporary science. Future research could explore ways the three prongs of authenticity intersect, authenticity as a changing practice, and ways to integrate authenticity into learning environments, materials, and teacher pedagogical practice.

References

- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2010). "Doing" science versus "being" a scientist: Examining 10/11-year-old schoolchildren's constructions of science through the lens of identity. *Science Education*, 94(4), 617–639.
- Barab, S. A., & Hay, K. E. (2001). Doing science at the elbows of experts: Issues related to the science apprenticeship camp. *Journal of Research in Science Teaching*, 38(1), 70–102.
- Barron, B. (2004). Learning Ecologies for Technological Fluency: Gender and Experience Differences. *Journal of Educational Computing Research*, 31(1), 1–36.
- Bell, P., Tzou, C., Bricker, L., & Baines, A. M. D. (January 01, 2012). Learning in Diversities of Structures of Social Practice: Accounting for How, Why and Where People Learn Science. *Human Development*, 55, 269–284.
- Cognition and Technology Group at Vanderbilt. (1992). The Jasper series as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist*, 27, 291–315.
- Heath, S. B. (2012). *Words at work and play: Three decades in family and community life*. Cambridge: Cambridge University Press.
- Eisenhart, M. (1996). The production of biologists at school and work: Making scientists, conservationists, or flowery bone-heads? In B. Levinson, D. Foley, & D. Holland (Eds.), *The cultural production of the educated person* (pp. 169–185). Albany: State University of New York 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.
- Feinstein, N. (2010). Salvaging science literacy. *Science Education*, 95(1), 168–185.
- Markus, H., & Nurius, P. (1986). Possible Selves. *American Psychologist*, 41(9), 954–969.
- Nasir, N.S. (2002). Identity, Goals, and Learning : Mathematics in Cultural Practice. *Mathematical Thinking and Learning*, 4(2), 213– 247.
- Nasir, N. S. (2011). *Racialized identities: Race and achievement among African American youth*. Stanford, California: Stanford University Press.
- National Research Council. (2011). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, D.C.: The National Academies Press.
- Nolen, S. B., Tierney, G. Goodell, A., Eng, S., Whitfield, L., Nguyen, D., Cooper, S., Becherer, K., Pak, S., Abbott, R. (2013). Teaching and Learning Practice-Linked Environmental Identities: A Design-Based Study. Paper presented at the biennial meeting of the European Association of Research on Learning and Instruction, Munich, September.
- Rivet, A. E., & Krajcik, J. S. (January 01, 2008). Contextualizing instruction: Leveraging students' prior knowledge and experiences to foster understanding of middle school science. *Journal of Research in Science Teaching*, 45(1), 79–100.
- van Eijck, M., Hsu, P.L., & Roth, W.M. (2009). Translations of scientific practice to "students' images of science." *Science Education*, 93(4), 611–634.
- van Eijck, M., & Roth, W.M. (2009). Authentic science experiences as a vehicle to change students' orientations toward science and scientific career choices: Learning from the path followed by Brad. *Cultural Studies of Science Education*, 4(3), 611–638.

Acknowledgements

This work was supported by grants 1261892 and 1251595 (Susan Nolen, PI) from the National Science Foundation, and grants from the George Lucas Educational Foundation (John Bransford, PI). This work is also supported by a grant from the National Science Foundation (NSF#9014483) and is partly supported by the Learning in Informal and Formal Environments (LIFE) Science of Learning Center (<http://life-slc.org/>) (NSF# 0354453). All opinions, findings, and conclusions expressed here are those of the authors.

Personas as a Powerful Methodology to Design Targeted Professional Development Resources

Adrian Madsen and Sarah B. McKagan, American Association of Physics Teachers, College Park, MD 20740

Eleanor C. Sayre, Kansas State University, Manhattan, KS 66506

Matthew Martinuk and Alexander Bell, Cognition Technology, Vancouver, BC V5N3Z2

Email: {amadsen, smckagan}@aapt.org, esayre@ksu.edu, {sandy, alex}@cognitiontec.com

Abstract: Scaling and sustaining educational innovations is a common problem in the learning sciences. Professional development resources around educational innovations that are personalized to appeal to the needs and motivations of different types of faculty offer a possible solution. The method of developing “personas” to represent key types of users is commonly employed in user-interface design and can be used to produce personalized resources. Personas are fictional named archetypes of users encompassing generalizations of their key characteristics and goals that emerge from interviews. This method is especially powerful because personas succinctly package information into the form of a person, who is easily understood and reasoned about. Herein we describe the creation of a set of personas focusing on the needs and motivations of physics faculty around assessment and teaching. We present the personas, a discussion of how they inform our design and how the method can be used more broadly.

Introduction

A common issue in the learning sciences is the scaling and sustaining of educational innovations. Educators and researchers continually focus on creating new curricula and teaching methods, but many promising innovations are not as effective when implemented by different teachers in new settings (Penuel, Fishman, Haugan Cheng, & Sabelli, 2011) so faculty stop using them (Henderson, Dancy, & Niewiadomska-Bugaj, 2012). The primary strategy used in science professional development at the university level is to disseminate educational innovations as finished products (Henderson, Finkelstein, & Beach, 2010). This strategy assumes that successful implementation does not depend on the details of the particular faculty members, students and institutions. It does not leave room for the faculty member to play a meaningful role in the improvements or take into account their unique environment (Henderson et al., 2010). This is problematic as faculty value the uniqueness of their personal styles, skills and preferences when choosing a teaching method (Henderson & Dancy, 2008). It may be that there is too much focus on getting faculty to use the educational innovation “as is” and not enough focus on the individuality of the faculty and their environments.

One way to address scale and sustainability is through personalized professional development designed to take into account the diverse needs, motivations and environments of teachers and faculty who are potential or current users of educational innovations. We believe that personalized professional development resources will increase implementation and continuing use of these innovations because they target the things that faculty members actually care about. We are specifically designing a professional development website, the Physics Education Researcher (PER) User’s Guide (www.perusersguide.org), to support faculty use of curricular and assessment innovations in physics at the university level. Our aim is to understand the needs, motivations and environments of different types of faculty, prioritize which types of faculty our site will be designed for and design our website in a way that speaks to each and meets real needs.

We propose that the method of creating “personas” is a powerful way to understand different types of faculty members who are potential or continuing implementers of educational innovations. This method is primarily used in the user-interface design process where designers create new technology products to match the needs of potential users (Cooper, 2004; Jacobs, Dreessen, & Pierson, 2008). The creation and use of personas is a well-established research technique where rich sets of qualitative data about users’ goals and experiences are synthesized into user archetypes called “personas.” These personas are generalizations of key characteristics and goals of potential users. Personas are described as “fictional, detailed archetypical characters that represent grouping of behaviors, goals and motivations observed and identified during the research phase” (Calde, Goodwin, & Reimann, 2002). They are usually named and assigned a picture to help them feel like real people. Personas as a method has roots in ethnography (Blomberg & Burrell, 2009) as well as human computer interaction (Sears & Jacko, 2007).

A common problem in the design process (we use the term “design” in the engineering sense) is having too vague an idea of who the users are. Without clarity, it is impossible to communicate about the specific needs and goals of the users and design a resource or experience to meet these. This is referred to as the problem of the “elastic user” who stretches and changes as the needs and constraints of the project change (Cooper, 2004). To address this problem, one could design for a specific user, but this produces a product that is too narrowly

focused (Jacobs et al., 2008). Instead one could compile information on many users in the form of a report on key demographics and characteristics of users. The detailed nature of these reports makes them cumbersome to use (Pruitt & Adlin, 2010) and they are not naturally reasoned about by the human mind (Grudin, 2006).

In order to help designers understand, communicate about and effectively design for a variety of users, personas are used. These clearly defined representations of people package a large amount of information into a succinct format that is easy to understand and gives the design team something concrete to refer to and discuss (Pruitt & Adlin, 2010; Yström, Peterson, Von Sydow, & Malmqvist, 2010). Personas help the designers become user-centered and avoid self-centered designs (Pruitt & Adlin, 2010). Personas are an especially powerful communication tool because we as humans are innately equipped to generate and engage with representations of people, as opposed to statistical summaries, which are more difficult for us to engage with (Grudin, 2006). The personas do not represent individual users but contain key differentiating characteristics from an amalgamation of users. This allows designers to focus on important characteristics and not get sidetracked by superfluous details. The abstraction of personas away from the interviewees upon which they are based also ensures the anonymity of participants, which is especially important when focusing on sensitive issues that participants would not discuss unless anonymity were guaranteed. Further, there are cases where even though participants are not identified outright, their uniqueness creates an anonymity set too small to protect their identity. Creating personas for these kinds of participants ensures their protection. Finally, personas are referred back to throughout the entire project to ensure that the product does not depart from the needs of the actual users.

There are several limitations of the method of personas. Because personas blend together characteristics of many users, it is hard to tell which and how many users each persona actually represents (Chapman & Milham, 2006). There is a propensity for designers to base their understanding of users on stereotypes and not the actual users (Turner & Turner, 2011). This can be mitigated by closely coupling data to the persona creation. Personas are difficult to validate, as they emerge from an interaction of the specific set of users interviewed, the questions discussed and decisions by the design team (Chapman & Milham, 2006). The designers make conscious decisions about what aspects of the users to include in the personas based on which aspects they understand are important to the final product. It is likely that given a similar topic but different set of users and designers, different personas would emerge. That being said, validity can be established with expert review (Yström, Peterson, Sydow, & Malmqvist, 2009). The effectiveness of the final product for the intended user base is an additional measure of validity.

In this paper we detail our methodology, describe the personas that emerged from our interviews and discuss their usefulness in designing professional development resources around educational innovations.

Methodology: Creation of Personas

We used semi-structured interviews to learn about faculty members' and department heads' current practices and needs around assessment in physics. We interviewed 13 physics faculty and 11 physics department heads in online interviews. Video was used in most. We recruited faculty members from a list of participants from a professional development workshop for new faculty held several years earlier. We sent faculty an email invitation to participate. We subsequently received recommendations from current participants and the project staff for other faculty who might be willing to participate and invited them. We recruited department heads from a list of participants from a different professional development workshop held the previous summer. A sufficient number of these department heads agreed to participate, so no additional invitations were needed. Our overall sample included eight participants from undergraduate serving public institutions, six from undergraduate serving private institutions and 10 from research institutions.

As part of our interview protocol, we asked participants about their background, school and department, current teaching practices and use of assessment, needs around assessment and how our online resources might meet their needs. Our interview team included two user interface designers (one is a former physics education researcher) and a physics education researcher. During most of the interviews, one team member engaged with the interviewee while the others listened and took notes. The interviews were recorded and after each interview, the team members individually wrote down the key points they noticed, primarily attending to the user's motivations and goals, tasks that they commonly completed around assessment, attitudes and beliefs, needs, pain points and constraints. These items were often discussed amongst the team members before the next interview. After several interviews, the team identified a few characteristics that varied among the interviewees and strongly influenced their teaching and assessment practices. For example, the interviewees' level of "buy-in" to educational innovations and knowledge of these resources varied significantly and had a strong impact on their use of research-based materials. We plotted the participants on these two axes and found groupings of participants, which helped in the initial development of the personas. As the interviews progressed, we continued to identify key differentiators and used a constant comparison protocol to determine if they were meaningful. As we reached saturation, we developed an initial profile for each persona and tested these to see if they fit the actual users we had interviewed.

After we had completed all the interviews, one team member re-listened to most of the interviews and took additional notes on goals, motivations and needs, writing down key illustrative quotes. These quotes were used to create a document synthesizing the tasks, behaviors and attitudes of each persona. These quotes and synthesis were used to once again modify the personas. Finally, we discussed the current version of the personas as a team and made additional changes based on evidence presented in the discussion.

Using the final versions of the personas, we ranked each on the variety of differentiating characteristics identified during the persona construction process, e.g., the degree to which faculty valued evidence or intuition to guide their teaching and assessment practices. These rankings helped fill out the important characteristics of the personas. With a rich understanding of the personas, we created task flows and outlined associated pain points. Task flows are detailed lists of steps describing common tasks users complete around the aim of the project, in our case, common tasks around teaching and assessment. Examples include finding a new research-based assessment or teaching method or comparing the results of an assessment. Within each task flow, we identified “pain points” discussed by users. Pain points are any inconvenience, confusion or frustration that users encounter when performing a specific task, for example, uncertainty on how to interpret the results of an assessment or difficulty finding appropriate data to compare one’s results to.

Finally, the design team met with the larger group of project stakeholders to prioritize the personas with respect to the design of the website. They discussed their own goals for the website and the likelihood that a particular persona would benefit from the site’s offerings in order come to a consensus decision on which personas to design the site for.

Results: The Personas

Five personas emerged from our interviews (Table 1). In order to discuss these personas most naturally, they were given names and assigned pictures. We also chose key quotes from the interview transcripts to help us grasp who each persona represented and outlined their goals and motivations as related to assessment and teaching. Table 2 contains rankings of the personas on key characteristics that differentiate them from one another. Tables 1 and 2 are used together to get a full picture of what constitutes each persona. The demographics of the personas are not meant to be a meaningful characteristic. We recognize that these do not reflect the demographic distribution of US physics faculty.

After creating the personas, we outlined the key task flows and pain points that each would have when interacting with our online resources. For example, a common task flow for Diane, the pragmatic satisficer, is to search for a research-based assessment, run the assessment, analyze and interpret the results and compare the results with others. Her pain points include a concern that students are not taking it seriously, too much time is required to analyze the results and it is difficult to find data from other institutions for comparison. Several other task flows and associated pain points emerged for Diane during the interviews. This information gives the team a rich sense of who Diane is and her wants and difficulties, which allows us to design our online resource to meet her stated needs as well as infer solutions to other needs that users like her may not have voiced.

After review and discussion of the personas and our own goals for the site, the stakeholder team prioritized who the site will be designed for. We believe that we can have the most impact on scaling and sustaining educational innovations if we provide resources to faculty who are already motivated to improve student learning. We feel our impact would be lessened if we focused on convincing skeptical faculty to use these innovations. Further, faculty who are already successful at using innovations don’t need our help as much as those with little experience. So we prioritized in order from most to least important, Raphael the motivated novice, Tim the seeker, Diane the pragmatic satisficer, Marge the proto-researcher and Paula the skeptic.

Analysis and Discussion

Understanding who we should design for, their task flows and pain points, and prioritizing these personas tells us a lot about how to design our online resources so they are targeted and personal. For example, different personas thrive with different levels and kinds of information about educational innovations. Raphael is brand new to using educational innovations and doesn’t have a lot of time, so he needs simple practical information and ready-to-go teaching and assessment materials. He also needs resources on how to face obstacles he will encounter when using these new methods which will support him in persisting in his use of the innovations. On the other hand, Tim has experience using educational innovations and wants lots of details about the research-validation and theoretical underpinnings of the innovation. He would also like information on how this innovation compares to other innovations he is familiar with. This level of detail will help him intelligently choose the most appropriate innovation for his own environment. Paula has a low buy-in to the value of educational innovations but cares about student learning, so she needs resources to help her understand how educational innovations can improve learning before being interested in the details of a new innovation.

There is rarely a professional development situation where only one persona is represented. Further, an individual faculty member may have characteristics from several personas. To address this, different aspects of the professional development resources should be designed to appeal to the personas based on the prioritization

decided on by the stakeholder team. This is in contrast to the common practice of providing one type of professional development resource that might appeal to only one persona or contain so little or so much information that it benefits no one at all. We briefly discuss how we personalized our website based on the personas. The needs of Raphael were our top priority, so we designed a simple home page with very basic but enticing information and a jump bar with links to articles answering his commonly asked questions. From the home page, we provided links to more detailed pages so that Diane can, for example, compare her students to others using a database of assessment results to get the number she cares about. We also provided links from the home page for Tim leading him to detailed information about the research validation for each innovation and recommendations for new innovations to try. We believe design choices like these will help faculty who identify with a given persona easily find what they need to implement or continue using educational innovations.

We believe the method of personas is useful whenever researchers seek to design a professional development resource or experience aimed at scaling or sustaining educational innovations. Personas encourage the design team to focus on the users and transform a large amount of detailed information about individuals into digestible outlines of fictional people who can be easily be understood and discussed. This method also provides a framework for how to design professional development targeted at different types of users and fills in the details of who the users are and what needs or motivations they have that the innovation can satisfy. With an increased focus on the users and not just the innovations themselves, we believe more educators will take up new innovations and continue to use them.

Table 1: Descriptions of five personas emerged from interviews with physics department heads and faculty. These personas represent user archetypes, but no persona exactly represents any one user. (Names and pictures are included to help communicate about the personas as if they were real people.)






	Role	Key Quote	Motivations and Goals
Paula the Skeptic 	Adjunct New Faculty Exp. Faculty Dept. Head	"The entire assessment movement seems to ignore that we are already assessing our students."	<ul style="list-style-type: none"> • Unconvinced that research-based techniques offer anything more than traditional/current methods • Assessment to reaffirm validity of current teaching methods • External pressures for assessment
Raphael the Motivated Novice 	Adjunct New Faculty Exp. Faculty Dept. Head	"... make it useful for the students. I mean at the end of the day, that's all I really care about."	<ul style="list-style-type: none"> • Thoughtful teacher who cares about student learning • Traditional background or new to teaching • Acknowledges challenges with traditional methods • Time-pressured
Diane the Pragmatic Satisficer 	Adjunct New Faculty Exp. Faculty Dept. Head	"The ability for faculty to compare what they are doing with other faculty would be really useful and spur them on to improve their teaching."	<ul style="list-style-type: none"> • Interested in the result of assessment: the number • Gives assessment to compare results • Wants to know what teaching methods "work" • Prefers summary information
Tim the Seeker 	Adjunct New Faculty Exp. Faculty Dept. Head	"We're driven by data & cause & effect. If the data shows that the students are getting more out of the course, then that gives you reason to change."	<ul style="list-style-type: none"> • Generally curious and interested in student learning improvement • Asks questions which lead to change/improvement • Wants to find suitable existing tools • Internally motivated
Marge the Proto-researcher 	Adjunct New Faculty Exp. Faculty Dept. Head	"It's hard to argue that you don't want data about what your students are learning in your class."	<ul style="list-style-type: none"> • Has questions or needs beyond those already answered by educational research • Wants to assess new aspects of students understanding / development • Needs to do more analysis of results

Table 2: Ranking of personas on characteristics that emerged in user interviews. P = Paula the Skeptic, R = Raphael the Motivated Novice, D = Diane the Pragmatic Satisficer, T = Tim the Seeker, M = Marge the Proto-researcher, PER = physics education research.

High PER knowledge	M	T	D	R	P	Low PER knowledge
High PER buy-in	T M	D	R		P	Low PER buy-in
Reformed teacher	M	T	D	R	P	Traditional teacher
Evidence-based	M	T	D	R	P	Intuition-based
Teacher responsibility for learning		T M	D	R	P	Student responsibility for learning
Cares about meaning of results	M	T	R P		D	Cares about numerical result
Seeks expert opinion	D M	T			P	Confident in own opinion/skills
High influence over teachers	M	T		D P	R	Low influence over teachers

References

- Blomberg, J., & Burrell, M. (2009). An ethnographic approach to design. In A. Sears & J. A. Jacko (Eds.), *Human-Computer Interaction: Development Process* (pp. 71–94). Boca Raton: CRC Press.
- Calde, S., Goodwin, K., & Reimann, R. (2002). SHS Orcas : The first integrated information system for long-term healthcare facility management. In *CHI '02 Case Studies of the CHI2002 | AIGA Experience Design Forum* (pp. 2–16). New York: ACM.
- Chapman, C. N., & Milham, R. P. (2006). The personas' new clothes: methodological and practical arguments against a popular method. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, 50(5), 634–636.
- Cooper, A. (2004). *The inmates are running the asylum: Why high-tech products drive us crazy and how to restore the sanity*. Indianapolis: Sams - Pearson Education.
- Grudin, J. (2006). Why personas work: the psychological evidence. In J. Pruitt & T. Adlin (Eds.), *The persona lifecycle: Keeping people in mind throughout the product design* (pp. 643–663). Burlington: Morgan Kaufmann.
- Henderson, C., & Dancy, M. H. (2008). Physics faculty and educational researchers: Divergent expectations as barriers to the diffusion of innovations. *American Journal of Physics*, 76(1), 79–91.
- Henderson, C., Dancy, M., & Niewiadomska-Bugaj, M. (2012). Use of research-based instructional strategies in introductory physics: Where do faculty leave the innovation-decision process? *Physical Review Special Topics - Physics Education Research*, 8(2), 020104. doi:10.1103/PhysRevSTPER.8.020104
- Henderson, C., Finkelstein, N., & Beach, A. (2010). Beyond dissemination in college science teaching: An introduction to four core change strategies. *Journal of College Science Teaching*, 39(5), 18–25.
- Jacobs, A., Dreessen, K., & Pierson, J. (2008). “Thick” Personas – Using ethnographic methods for persona development as a tool for conveying the social science view in technological design. *Observatorio Journal*, 5, 79–97.
- 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.
- Pruitt, John, & Adlin, T. (2010). *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. Burlington: Morgan Kaufmann.
- Sears, A., & Jacko, J. A. (Eds.). (2007). *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications* (2nd ed.). New York: Taylor and Francis.
- Turner, P., & Turner, S. (2011). Is stereotyping inevitable when designing with personas? *Design Studies*, 32(1), 30–44.
- Yström, A., Peterson, L., von Sydow, B., & Malmqvist, J. (2009). Using personas as a mediating tool in the development of engineering programs. In *Proceedings 2: a Utvecklingskonferensen för Sveriges ingenjörsutbildningar*.
- Yström, A., Peterson, L., Von Sydow, B., & Malmqvist, J. (2010). Using Personas to Guide Education Needs Analysis and Program Design. In *Proceedings of the 6th International CDIO Conference*. Montréal.

Engaging in Educational Design Processes for Sustainable Learning: Learning and Becoming in Practice

Mona Holmqvist Olander and Clas Olander
University of Gothenburg, IDPP, Box 300, SE-405 30 Gothenburg, Sweden
mona.holmqvist@gu.se, clas.olander@gu.se

Abstract: The aim of this study is to describe in what way the students' learning outcomes and theoretical assumptions are used by teachers to enhance the design of lessons with a focus on development of sustainable learning over time. Seven teachers and 71 students in an upper secondary school participated, along with one researcher. Three lessons were designed and implemented in three different student groups ($N=24, 23, 24$), with one intervention for each group. The interventions were designed in an iterative process: the result of the first lesson formed the base for the development of the next, and so on. A multi-case analysis was conducted within and between each intervention. The results of the study describe correlations between teachers' deepened theoretical knowledge, changes in lesson design, and impact on students' learning outcomes in a long-term perspective. We found students' learning outcomes increased during the three interventions in three different groups of students (A +0.17, B +0.87 and C +1.54). At the delayed post-test, the learning gain in the third group sustained to be higher than in the previous two groups of students (A -0.08, B +0.26, C +0.75).

Survey of the Field

This study addresses a general issue in educational research, developing sustainable learning. Transfer is focusing how the things we learn in one situation affect our ability to handle other and new situations, related to what we have previously learned. Marton (2006) questions the assumption that transfer always builds on similarities (instead of differences) between different situations, that the similarity is defined by the researcher (not the participants), and that the relation is between two specific situations (rather than between "fields" of situations) should be questioned. Engle (2006) describes generative learning as the ability to appropriate the application of something in a new situation which differs from the first even when the first and second situations as such in some ways are related to each other. Our definition differs as we do not think it is enough to examine only two situations or to evaluate the learning in only one post-unit assessment activity. The design of assessing generative learning (learning that generates further learning) in our definition, means that we assess students' knowledge immediately before and after the intervention, which may be seen as one occasion (pre-test, intervention, post-test) because it results in a single analysis of the effect of the initial intervention. The weeks after the intervention, students participate in several different contexts, and they have the opportunity to discern the object of learning in other situations, depending on the degree to which the intervention has developed their potential generative learning. This can be seen as a second activity, which means we imply that the context the students meet after the intervention contributes to their learning. If their ability to learn further is enhanced by the initial lesson, their learning increases or sustains during the weeks and will be possible to assess in a new test several weeks after the initial intervention. As Carraher and Schliemann (2009) suggest, we have taken into consideration how prior knowledge plays a role in transfer; seeing transfer as a theory of learning there is a need to develop a hypothesis about what it takes to learn. This means that the instruction in the initial learning situation have an impact on the learning beyond the immediate learning situations if it is designed based on theoretical assumptions about learning. As the lesson forms students' prior knowledge in a new situation and opens up their ways of thinking about the content, different ways to design the lessons might have different impact on the learning outcome beyond the lesson, and that is what is studied in this project.

Purpose

The aim of this study is to describe in what way theoretical assumptions about learning and the students' learning outcomes are used by teachers to enhance the design of lessons with a focus on development of sustainable learning over time. Teachers' developed skills to design lessons in and by practice to enhance student learning is in focus. The design of the lesson plans are analysed together with the results of the students' learning outcomes in three different groups. By the use of an iterative process called learning study (Marton & Tsui, 2004; Holmqvist, 2011; Lo & Marton, 2011), three teachers teach three different lesson designs in three groups of students about the same content at three different lessons (one lesson in each group of students). The analysis of the learning outcomes and the design of the lessons are made collaborative in the whole team of teachers.

Theoretical Background

Variation theory (Marton & Booth, 1997; Runesson, 1999; Holmqvist, Gustavsson & Wernberg, 2008) has been used as the theoretical framework to design and analyse how students' generative learning is developed. In this framework, learning is seen as a qualitatively changed understanding of a phenomenon in the environment. This means that learning is an ongoing process where the learner gradually experiences instances in a more and more developed way. Learning is thus not seen as either correct or incorrect, for in all misconceptions there are some correct assumptions that can be developed. Variation theory is based on three corner-concepts: discernment, simultaneity, and variation. They all require each other in a learning situation and are intertwined. To learn something new, you have to discern this new part from a previously discerned background. This means you compare and see the new thing in contrast with other phenomena you have previously experienced. This variation makes it possible to suddenly experience something new in a former, familiar situation. In a very simplified way – you might not experience one particular house in a street that you travel every day until one day the owners repaint the house in a new colour. The variation of colour (once red and now yellow) makes you discern the house in a new way as you simultaneously remember what it looked like before and what it looks like right now. This theory has been used to develop learning in several different countries (Marton & Tsui, 2004), a number of different school subjects (Lo & Marton, 2011; Holmqvist Olander & Sandberg, 2013; Ljung-Djärf & Holmqvist Olander, 2013; Holmqvist Olander & Bergentoft, 2014), and beyond the bounds of the ordinary school (Holmqvist, 2009).

Multiple-Case study

The development of lessons designed for generative learning was studied in a multiple-case study (Stake, 2006), with three cases, called “learning study cycles” consisting of five steps each: design of one lesson, pre-test, intervention, post-test, and post-test session. The teachers' aim in the learning study was to develop learning designs in science education dealing with a general issue concerning students' learning of the “nature of science” (Lederman, 2007) with a specific object of learning, “What constitutes a scientific theory?” When the teachers discussed what topic to choose before the learning study, they agreed that the nature of science was a topic that was both challenging for them to teach and for students to learn; furthermore, it is a vital part of science subjects in Sweden (biology, chemistry, and physics). The learning study was enacted as part of a course in biology, where the idea of the nature of science is emphasised in the syllabuses, especially under the heading of “Character and working methods of biology,” that is, “Models and theories as simplifications of the world – How they change over time” and “The importance of the experimental working process in testing, re-evaluating, and revising hypotheses, theories, and models.”

Methodology, Participants, and Implementation

A learning study is a kind of action research; the learning study in this case was part of a school development project in an entire upper secondary school, where different subject-based teams of teachers chose different objects of learning. The participants were seven teachers, one researcher, and 71 students (approximately 17 years old) in three groups ($N=24$, 23, 24) in upper secondary school. Two group interviews about the content were conducted with equivalent students ($N=8$) who should not participate in the multiple-case study. The interview focused on probing students' views of how science could make claims and how you could judge whether a theory is scientific or not. It turned out that the students could rather easily spot whether a theory was scientific or not, but they were very vague when it came to reasons for their conclusions. The result of the screening was discussed at a meeting between the researcher and the teachers; the object of learning was formulated as “understanding what constitutes a scientific theory,” instead of the previous focus on what a scientific theory is.

Data consisted of three videotaped research lessons in three different groups of students (taught by teacher A, B and C), pre-, post-, and delayed post-tests seven weeks after each intervention. There were also meetings (in total, six) before and after each lesson where teachers and researcher evaluated the previous lesson and suggested ways of communicating identified critical aspects in the next lesson. The data source for the teachers' experiences consisted of notes and minutes from six 2-hour meetings: 11/9, 27/9, 4/10, 9/10, 16/10 and 11/12 at the school. Between the meetings, the whole team of teachers assessed students' tests and, if possible, watched videos taken during the lessons. At least two teachers were also present during the research lessons, in that one was teaching and another was videotaping. The video-recordings represent an important data source since they point out in what way the teachers could become more skilled in designing the lessons based on students' previous knowledge.

Tests and Assessments

The teachers decided to teach about the nature of science. The idea of including the nature of science (NOS) in the science curricula is a worldwide phenomenon; however, there seems little agreement on what NOS is; Abd-El Khalick (2012) concludes that we currently “have no well-confirmed general picture of how science works,

no theory of science worthy of general assent” (p. 354). However, some components seem inevitable, e.g. definitions and relationships between concepts, laws, models, and theories. Suggestions of constituents of a scientific theory include such aspects as being testable, replicable, and/or leading to predictions (p.737); they are fallible human constructions but have substantive evidence that supports them (Dagher, Brickhouse, Shipman & Letts, 2004, p. 739). This definition was the point of departure for teachers in this study, and the lesson designs are used as data to describe in what way the teachers develop their lesson design based on theoretical knowledge and the students’ learning outcomes.

The pre-test, post-test, and delayed post-tests were identically articulated, although some students answered with paper and pencil while others logged on to a Web platform to enter and submit their answers. The analysis of results shows no difference in the way students answered regardless of response method. To keep the answers anonymous, students were given a three-digit code to be used on all three tests. The question was presented to students as shown in Figure 1.

<p>What constitutes a scientific theory?</p> <p>I claim that the following examples are scientific theories.</p> <p>(a) gravitation (deals with the attraction of bodies to each other)</p> <p>(b) telepathy (deals with the transmission of thoughts)</p> <p>(c) astrology (deals with the influence of astronomical bodies on the life of humans)</p> <p>(d) astronomy (deals with the structure and motion of astronomical bodies)</p> <p>Argue, i.e. give reasons, as to whether they are or are not scientific theories.</p>	Code:
--	-------

Figure 1. Pre-, post- and delayed post-test.

When assessing students’ answers, we searched for the following criteria: A theory ...

- 1) may be investigated/tested *and/or* be falsifiable.
- 2) is supported by previous findings.
- 3) may be used to make predictions.
- 4) may be modified and developed *and/or* the formulation is a human construct

Taking into consideration all four aspects of what a theory is would count as 4 points on the test, one point for each aspect. However, the assessment was not primarily based on a quantitative scoring; rather, the responses were analysed qualitatively to see in what way the students had developed understanding or not. Table 1 shows excerpts from one student’s answers at pre- and post-test. The student has answered correctly as to whether the examples are scientific theories or not; however, the aspects she used to make these decisions resulted in different scores and showed (on the post-test) a deeper understanding that was beyond merely answering the direct question correctly.

Table 1: Answers from student #330

	Pre-test	Post-test
(a) gravitation	Yes, because the scientific arguments support the theory. (It is most likely, there is “evidence.”)	Gravity is consistent with experiments and so on. You can try to contradict the theory but with no success. That is correct!
(b) telepathy	No, because there are no scientific arguments that support the “theory.”	Experiment is not consistent with this hypothesis. There is no evidence that it is correct/there is evidence that it is wrong because the experiments are not consistent with the hypothesis. It’s wrong!

Results

Differences in Teachers’ Handling of the Content

Each lesson started with a brief introduction by the teacher and a framework of some aspects defining a scientific theory (hypothesis – testing – theory). After the introduction came a group activity where students in groups of three or four were given a “black box” (a large sealed box filled with unknown objects). The students were to figure out what was inside and give reasons for their assumptions (hypothesis). After that, the teacher made a PowerPoint presentation on the work of Ignaz Semmelweis and his hypothetic deductive method when trying to discover the cause of puerperal (childbed) fever. This part of the lesson resulted in a change of focus, as students became occupied in finding out why mothers died of puerperal fever instead of develop an understanding of what a scientific theory is. The teacher aimed to build a structure of relevance, but instead the students in lesson A and B became very focused on finding out why doctors in the 19th century could not stop

the incidence of puerperal fever. The theory was kept in the background and the disease in the foreground. After this presentation, the framework was further developed on the whiteboard and related to the Semmelweis case. This was presented sequential in all three lessons, one at a time, instead of simultaneously or integrated. However, in lesson C the teacher made simultaneous references to the Semmelweis case when she drew the schedule on the black board.

At the end of the each lesson, the students were supposed to contrast what a scientific theory is with what it is not. In the first lesson (A), the example used was whether the theory of a flat earth could be considered as scientific; in lessons B and C, the contrast presented was between Semmelweis's theory and numerology. The flat earth theory was, however, all too clearly inaccurate to be a good example to find out if it is a scientific theory using the model drawn at the black board. This resulted in the teachers changing the example, in that way offering students a contrast between two phenomena from different areas. The strategy was to sharpen the contrast (Semmelweis versus numerology) even further in lesson C by also letting students do an online numerology test (<http://www.twice.se/astro-numerologisk-kalkyl.html#UhfRDD9YW2k>).

Except for the difference in the examples of contrasts (flat earth and numerology); the main difference between the lessons lay in which aspects the teacher offered students when explaining what constitutes a scientific theory. This only took a few minutes of each lesson (five minutes in lesson B and eight in lesson C), but was the crucial part of the learning opportunity to understand what a scientific theory is, as the rest of the lesson was mainly about building a structure of relevance or background. In lesson A, there were few aspects of the structure devoted to the teacher explaining what a scientific theory is; the teacher mainly pointed out that observations lead to a hypothesis that could be tested and falsified or supported. During the post-lesson session after lesson A, the teachers decided to offer the students more aspects. This resulted in new dimensions of a variation of ways to reinforce a conclusion (experiments, earlier models, and predictions) which would lead to a generalisation and, finally, a theory. In lesson C, the pattern of variation was changed again, keeping the dimension of variation regarding the reinforcement but also opening up a dimension of variation regarding what a theory is (possible to develop, testable, and predictive). Thus, the analysis of teachers' intentions gradually includes all four aspects that constitute a scientific theory: (1) may be investigated/tested *and/or* be falsifiable, (2) is supported by previous findings, (3) may be used to make predictions, and (4) may be modified and developed *and/or* the formulation is a human construct. In lesson A, the aspects offered were about aspect (1), while the rest were not offered and therefore could not be discerned. The dimensions of variation opened up in lesson B, also concerning aspect (2) and (3); while in lesson C students had the opportunity to explicitly discern all four aspects and the teachers' intended or planned object of learning was thus closer to the enacted object of learning offered the students in lesson C than in lessons A and B.

Results of Students' Learning

Students' responses to pre-tests and post-tests (Table 2) show that students' learning gains in lesson A were low (and not significant, measured with paired t-test, 5%-level). However, their responses in lessons B and C were more in line with the scientific view after the lesson than before. This change is significant in both lesson B and C, but in lesson B the only category that the students seem to have discerned was "Scientific theories may be investigated/tested" (20 out of 23). After lesson C, the results show that the students in this group were able to give a more varied view; for example, 7 out of 24 argued with the help of three or four criteria, and another 10 students used two criteria (six students used only one criterion, and one failed to use any). Lesson C was also the lesson where students were able to retrieve developed knowledge after seven weeks; these students retained the knowledge over time and thus evidenced a more sustainable learning.

Table 2: Mean scores, differences, and range (0–4) in tests

	Research lesson A			Research lesson B			Research lesson C		
	Mean	Range	+/-	Mean	Range	+/-	Mean	Range	+/-
Pre-test	0.25	0–1		0.00	0–0		0.54	0–2	
Post-test	0.42	0–1	+ 0.17	0.87	0–1	+ 0.87	2.08	0–4	+ 1.54
Delayed post-test	0.17	0–1	- 0.25	0.26	0–1	- 0.61	1.29	0–4	- 0.79
Pre-test / Delayed post-test			- 0.08			+ 0.26			+ 0.75

In the results of the delayed post-test, we found a recession in all groups (from post-test to delayed post-test), but in lessons B and C there is an increased mean score in number of criteria mentioned when comparing pre-test and delayed post-test – specifically 0.26 in lesson B and 0.75 in lesson C (both significant). The students in lessons A and B did not retrieve more than one of the assessment criteria, however not necessary the same after as initially (as the student in Table 1). This was the same at the pre-test in lesson C, however in the post-test 10 students mentioned two criteria, four students mentioned three and three students mentioned all four criteria. At the delayed post-test, nine students mentioned two criteria and one student

mentioned three criteria. The students in this group did develop a more differentiated knowledge than the students in the two other lessons.

Discussion

In this study, teachers' use of students' learning outcomes and gradually more developed theoretical knowledge about what it takes to learn have been studied. The design of the lessons became, over time, more and more explicitly focused on the aspects critical to discern for the students to enhance knowledge. The iterative process; plan, implement, assess, analyze and new planning based on the previous findings, enables a cumulative learning process for the teachers. When enacting the third lesson and making the contrast between Semmelweis and numerology; teachers opened up a dimension of variation where the two contrasting theories were offered to discern simultaneously. The increase of the students' learning outcomes was highest in the third lesson, supporting the hypothesis that teachers' developed understanding of what it takes to learn the defined content, and by that meeting the students' needs in a more powerful way, enhances student learning. When participating in a learning study, the first step is to articulate what you intend to teach and how (i.e. the intended object of learning), then put your knowledge into action in the classroom together with the students (i.e. enacted object of learning), and finally analyse the classroom practice through students' learning outcomes (i.e. lived object of learning). We argue that these steps are most effective if undertaken in authentic practice (Brown, 1992) and in a collegial context where in-depth discussions of teaching strategies are grounded in observation or video-recordings. In that respect, learning study as an in-service training model has potential for developing teachers' learning and becoming in practice regarding how to stimulate students to develop sustainable learning.

References

- Abd-El-Khalick, F. (2012) Examining the Sources for our Understandings about Science, *International Journal of Science Education*, 34(3), 353-374
- Brown, A. L. (1992). Design experiments': Theoretical wild duck methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, 2 (2), 141-178.
- Carraher, D., & Schliemann, A. (2009). The transfer Dilemma. *Journal of the Learning sciences*, 11(1), 1-24.
- Dagher, Z., Brickhouse, N., Shipman, H., & Letts, W. (2004). How some college students represent their understandings of the nature of scientific theories, *International Journal of Science Education*, 26(6), 735-755.
- Engle, R. (2006). Framing Interactions to Foster Generative Learning: A Situative Explanation of Transfer in a Community of Learners Classroom. *The Journal of the Learning Sciences*, 15(4), 451-498
- Holmqvist, M., Gustavsson, Laila & Wernberg, Anna (2007): Generative learning. Learning beyond the learning situation. *Educational Action Research*, 15(2), 181-208
- Holmqvist, M., Gustavsson, L. & Wernberg, A. (2008) Variation Theory – An Organizing Principle to Guide Design Research in Education. In Kelly, A.E., Lesh, R., & Baek J. (eds) *Handbook of design research methods in education*, 111-130. New York: Routledge.
- Holmqvist, M. (2011). Teachers' learning in a learning study. *Instructional Science*, 39(4), 497-511.
- Holmqvist Olander, M. & Sandberg, B (2013). Democratic literacy - a complex object of learning. *The International Journal for Lesson and Learning Studies*, 2(3), 251-289
- Holmqvist Olander, M. & Bergentoft, H. (2014). Regulation of tenseness. *International Journal for Lesson and Learning Studies*, 3(1), 24-45
- Lederman, N. (2007). Nature of science: Past, present and future. In Abell, S & Lederman, N. (Eds) *Handbook of Research on Science Education* (pp. 831-880). Lawrence Erlbaum Associates, New Jersey/London.
- Ljung-Djårf, A. & Holmqvist Olander, M. (2013). Using Learning Study to Understand Preschoolers' Learning: Challenges and Possibilities. *International Journal of Early Childhood*, 1(45), 77-100.
- Lo, M.L., Chik P.M., Ko, P.Y., Leung, Y.L., Pong, W.Y., Lo-Fu, Y.W. & Ng, F.P. (2005) Learning Studies: Development and Impact. In Lo, M.L., Pong, W.Y. & Chik, P.M. (Eds.) *For Each and Everyone: Catering for Individual Differences through Learning Studies* (41- 73). Hong Kong: Hong Kong University Press.
- Lo Mun Ling & Marton, F. (2011). Towards a science of the art of teaching: Using variation theory as a guiding principle of pedagogical design. *International Journal for Lesson and Learning Studies*, 1(1), 7 – 22
- Marton, F. & Booth, S. (1997). *Learning and awareness*. Mahwah: Lawrence Erlbaum Associates.
- Marton, F. & Tsui, A.B. (2004). *Classroom discourse and the Space of Learning*. London: Lawrence Erlbaum Associates.
- Marton, F. (2006). Sameness and Difference in Transfer. *The Journal of the Learning Science*, 15, 499-535.
- Runesson, U. (1999). *Variationens pedagogik. Skilda sätt att behandla ett matematiskt innehåll*. [The Pedagogy of Variation] Göteborg: Acta Universitatis Gothoburgensis.
- Stake, R. (2006). *Multiple Case Study Analysis*. New York & London: The Guilford Press.

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

Jakita O. Thomas, Spelman College, jthoma41@spelman.edu

Abstract: Computational algorithmic thinking (CAT) is the ability to design, implement, and assess the implementation of algorithms to solve a range of problems. It involves identifying and understanding a problem, articulating an algorithm or set of algorithms in the form of a solution to the problem, implementing that solution in such a way that it solves the problem, and evaluating the solution based on some set of criteria. This report introduces and describes CAT as explored through the Supporting Computational Algorithmic Thinking (SCAT) project, an on-going longitudinal between-subjects research project and enrichment program that guides African-American middle school girls through the iterative game design cycle resulting in a set of complex games around broad themes.

Introduction

Jeanette Wing (2006) defines computational thinking as “*a way humans solve problems...*”. This research makes explicit a critical aspect of computational thinking through its focus: the design, development, and implementation of algorithms to solve problems. An algorithm is defined as a well-ordered collection of unambiguous and effectively computable operations that, when executed, produces a result and halts in a finite amount of time (Schnieder & Gersting, 2010).

Computational algorithmic thinking (CAT) is the ability to design, implement, and assess the implementation of algorithms to solve a range of problems. It involves identifying and understanding a problem, articulating an algorithm or set of algorithms in the form of a solution to the problem, implementing that solution in such a way that it solves the problem, and evaluating the solution based on some set of criteria. CAT has roots in Mathematics (Polya, 1973), through problem solving and algorithmic thinking (Kramer, 2002). CAT lies at the heart of Computer Science, which is defined as the study of algorithms (Schneider & Gersting, 2010). CAT embodies the ability to think critically and creatively to solve problems and has applicability in a range of areas from Computer Science to cooking to music (ISTE NETS.S, #4, 2007; Polya, 1973; Wing, 2010).

Supporting Computational Algorithmic Thinking (SCAT) is a longitudinal between-subjects research project exploring how African-American middle-school girls develop CAT capabilities over time in the context of game design. SCAT is also a free enrichment program designed to expose middle school girls to game design. The goals are: 1) to explore the development of computational algorithmic thinking over three years in African-American middle-school girls as they engage in iterative game design, and 2) to increase the awareness of participants to the broad applicability of computational algorithmic thinking across a number of industries and career paths. Spanning three years, participants, called SCAT Scholars, develop CAT capabilities as they design more and more complex games. SCAT Scholars begin the program the summer prior to their 6th grade year and continue through their 8th grade year. They engage in 3 types of activities each year: 1) a two-week intensive game design summer camp; 2) Two (2) six-week technical workshops where Scholars implement the games they have designed using visual and programming languages in preparation for submission to national game design competitions; and 3) field trips where Scholars learn about applications of CAT in different industries and careers. This work aims to explore the following research questions:

1. How do individual and small-group computational algorithmic thinking capabilities of African-American middle school girls develop over time?
2. What difficulties do learners face as they engage in computational algorithmic thinking?
3. What do those difficulties suggest about supporting learners as they engage in computational algorithmic thinking?
4. How does participating in SCAT impact participants' perspectives of computational algorithmic thinking as well as their perceptions of themselves as problem solvers and game designers?

Game design has been chosen as the domain for a number of reasons. First, game design is a domain with which middle-schoolers have a great deal of familiarity as consumers (Kafai, 2006; Irvine, 2010). The Pew Internet & American Life Project's survey revealed that among young people, ages 12 – 17, 97% of respondents play video games (Lenhart, Kahne, Middaugh, Macgill, Evans & Vitak, 2008). As such, this domain can provide motivation as learners “look under the hood” of their favorite games to understand how they are designed and implemented. Second, game design is centered around the iterative design, representation, and implementation of algorithms, which makes it an ideal domain to understand and describe the development of

CAT over time (Crawford, 2010). Third, based upon industry practices, game designers iteratively move from game conceptualization to production and release over time (Fullerton, et. al., 2004), making game design an ideal domain for conducting longitudinal studies. Lastly, game design is a domain in which African-American women are grossly under-represented (Brathwaite, 2009). Despite the fact that, of the 97% of young people who stated they played games in the Pew Institute's survey, over 94% of girls play video games with little difference in the percentages by race, ethnic group, or socio-economic status (Lenhart, Kahne, Middaugh, Macgill, Evans & Vitak, 2008), women represent only about 10 – 12% of the game design workforce, and Latinos and African-Americans comprise less than 5% combined (Plutzik, 2010).

While there is a great deal of research that examines how to engage students in computational thinking and learning in Computer Science (CS) or that focuses on how game design improves IT fluency, algorithmic thinking, collaboration, programming capability, and broader participation from under-represented groups, there is a scarcity of research that focuses on understanding and describing how the development of CAT happens over time as a complex cognitive capability (Repenning & Ioannidou, 2008; Owensby, 2006; Thomas, 2008; Werner, Campe & Denner, 2005; Maloney, Burd, Kafai, Rusk, Silverman & Resnick, 2004; DiSalvo, Guzdial, Mcklin, Meadows, Perry, Steward & Bruckman, 2009; Barnes, Richter, Chaffin, Godwin, Powell & Ralph, 2007; Kafai, 2006; Koschmann, et. al., 1996; Papert, 1993). Furthermore, there is less research that focuses on understanding how the development of these kinds of complex cognitive capabilities can impact not only how we leverage game design to teach and support students as they develop these capabilities, but also how we define and measure the learning that happens during that development.

This report describes the SCAT project, which is in the first of three years of data collection. The next section of this report will provide the background context that grounds the research. Then, the SCAT learning environment, including the scaffolds that support Scholars as they engage in game design, is described. As this work is in its first year, this report does not currently have any findings or results to share. Instead, this report is designed to introduce computational algorithmic thinking as a complex cognitive capability and SCAT as a project and program designed to develop CAT within this population, similar to Maloney, et. al (2004).

Background

The National Research Council (NRC, 2011), in their report entitled *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, outlines eight practices as being “essential elements of the K-12 science and engineering curriculum”. Among them are: defining problems, developing and using models (physical or mathematical models and prototypes), planning and carrying out investigations, analyzing and interpreting data, using mathematics, information & computer technology, and computational thinking, designing solutions, engaging in argument from evidence, and obtaining, evaluating, and communicating information. While the major competencies that students should have by the 12th grade and sketches regarding how that competence should progress are described, the NRC identifies that those sketches are based on The Committee on a Conceptual Framework for New Science Education Standards' judgment as “there is very little research evidence as yet on the developmental trajectory of each of these practices” (p. 3-6).

As a domain, engaging in game design aligns with the eight practices outlined by the NRC (2011). The iterative game design lifecycle involves several phases, which are also iterative (Fullerton, et. al., 2004). During brainstorming, game designers generate many ideas for games and present those ideas. Once an idea is selected, paper-and-pencil drawings are created, called storyboards, that include demo artwork. Playtesting is next, which involves bringing actual players from the target user group in and observing them as they play the game (or engage with the storyboard) in real time, getting feedback about the game experience to inform the design of the game (Fullerton, et. al., 2010; DiSalvo, Guzdial, Mcklin, Meadows, Perry, Steward & Bruckman, 2009). Next, game designers create a playable physical prototype using paper-and-pencil and/or craft materials. Then, a rough software prototype is created which models some aspect(s) of core gameplay. Then follows more playtesting. Next comes creating the design document, which outlines every aspect of the game and how it will function. This is followed by implementing the game with playtesting throughout implementation. Finally, quality assurance testing is done with continued playtesting.

The acquisition and development of skills, capabilities, and practices involves the changing of declarative knowledge, or independent pieces of factual knowledge, to procedural knowledge, or connected knowledge that forms a process for carrying out a skill (Anderson, Kline, Greeno & Neves, 1981; Anderson, 2000). Applied in context and/or among a community, a process evolves into a practice (Nersessian, 2008; Lave & Wenger, 1991). While skills, or abilities refer to what one can do in the present, capabilities refer to what one can learn to do with instruction and support, or scaffolding (Bandura, 1994; Bandura, <http://des.emory.edu/mfp/self-efficacy.html#bandura>, Vygotsky, 1978; Bransford, Brown & Cocking, 1999; Tabak, 2004). However, moving learners from capability to ability requires several things (Bransford, Brown & Cocking, 1999; Owensby, 2006; Thomas, 2008). First, learners need opportunities to make connections between their experiences and the knowledge or skills they are learning. Second, learners need enough time to learn and develop skills and capabilities so that they can use them flexibly in appropriate situations. Third,

learners should be supported as they attempt to represent problems at higher levels of abstraction. Finally, learners should be encouraged to monitor their learning and should be supported as they learn metacognitive strategies.

SCAT Learning Environment

The facilitator plays a major role in the development of Scholars' CAT capabilities in the SCAT learning environment, as she serves first as the primary modeler and then as a just-in-time coach (Collins, Brown & Newman, 1989). In addition, the facilitator leads and supports discussions that help Scholars as they think through their designs, helps them make connections across dyad experiences and problems, and models the kinds of questions Scholars should be asking themselves and their peers as they develop algorithms for their game designs, move through the iterative game design cycle, and reflect on their use of CAT (Koschmann, et. al., 1996). As dyads work on their game designs, she walks from group to group asking them questions about their designs, helping them identify problems and issues, illustrating for them how to use the Design Notebook and other tools and resources provided to them to help them design their games, and serves as a sounding board for dyads as they design. Although the facilitator is a critical component to the SCAT learning environment, she cannot be with every group or individual all the time. To help overcome that limitation and to help Scholars develop more expert CAT capabilities, the Design Notebook has been created to coach Scholars as they engage in CAT through game design. The Design Notebook has been integrated into SCAT activities, affording Scholars multiple opportunities to develop CAT capabilities while working individually and collaboratively in dyads.

The Design Notebook contains paper-and-pencil based tools that coach groups and individuals in the ways cognitive apprenticeship suggests (Collins, Brown & Newman, 1989; Puntembekar & Kolodner, 1998) by using a system of scaffolds (Owensby, 2006; Thomas, 2008). Each scaffold in the system supports groups and individuals in a particular way and addresses a particular difficulty that learners may face when engaging in complex cognitive skills, processes, and capabilities like designing an experiment, interpreting and applying the experiences of experts, or engaging in CAT. The system of scaffolds has 5 parts (Owensby, 2006; Thomas, 2008). First, tool sequences make process sequence visible. This scaffold addresses the structuring of tools to suggest a high-level process that learners are engaging in. Second, within each tool, structured questioning or statements make the task sequence clear. This scaffold addresses prompts, which are questions or statements used to focus learners' attention as they are carrying out or reflecting on a task. Third, for each prompt in the sequence, hints are provided. Hints are task-specific/domain-specific questions or statements used to refine a task. Fourth, for each prompt in the sequence, examples are provided. Examples are exemplars that can be used to model a process or a specific step of a process. Fifth, for some tasks in the sequencing, a template or chart to help with lining up one's reasoning is provided.

Given that Scholars will be able to move through the iterative game design cycle at their own pace, it is likely that those Scholars or dyads who are further along in the game design cycle will be able to scaffold dyads who are not as far along (Vygotsky, 1978; Roschelle, 1996; Owensby, 2006; Thomas, 2008; Palinscar, 1984). In addition, different Scholars will bring different perspectives to the dyad, which will contribute to greater understanding by the dyad. The literature shows that small group collaboration and discussion has many benefits (Feltovich, Spiro, Coulson & Feltovich, 1996; Koschmann, Kelson, Feltovich & Barrows, 1996; Roschelle, 1996; Bayer, 1990; Wells & Chang-Wells, 1992; Barron, et. al, 1998, Barron, 2003).

Current Project Status and Next Steps

We are currently working with 20 African-American sixth grade girls who attend various schools across the metro-Atlanta area. The SCAT project kicked off its first year of data collection this past summer (July 2013) beginning with a two-week summer camp. During the camp, Scholars were introduced to the game design cycle and engaged in brainstorming, storyboarding, and physical prototyping. They designed games for social change that address a range of topics from environmental sustainability to bullying. Scholars used the full range of scaffolds, including asking questions of the facilitator, using the Design Notebook, and each other as they not only engaged in whole class discussions, but also as they participated in gallery walks (Kolodner, et. al, 2003), where Scholars interacted one-on-one, playtested each other's games, and gave feedback to inform the next iteration of design. On the last day of camp, Scholars visited the Aware Home on the campus of Georgia Tech, which is an "intelligent" home as well as a research project aimed at exploring various technical, design, and social issues that people face in a home setting. Scholars toured the home, learned about the different algorithms that were designed and implemented to run the home, and interacted with undergraduate/graduate students who design, implement, integrate, and maintain those algorithms.

During the Fall 2013 semester, Scholars participated in monthly SCRATCH workshops, learning about the SCRATCH environment and programming common game functionality using SCRATCH (e.g., moving the player from one level to the next; creating the look and behavior of a sprite; and creating a scrolling scene). Beginning in February, Scholars will engage in two 6-week workshops where they will develop design

documents, implement their games using SCRATCH, and submit their games to two different national game design competitions. The 2014 STEM Video Game Challenge aims to motivate interest in STEM learning among America's youth by tapping into students' natural passions for playing and making video games. It was initiated in response to President Obama's initiative to promote a renewed focus on STEM education, the Educate to Innovate Campaign. Scholars can either submit written design documents or playable game prototypes. Scholars will submit both design documents and playable game prototypes to the Future Game Designers Competition.

As Scholars engage in SCAT activities, we have and will continue to collect video observations (both whole class and dyads as they work together), artifacts they create, survey data (pre- and post), and interview data (from both the facilitator as well as the Scholars) as data. Video observations allow us to have a visual record of the enactment, see Scholars struggle with CAT through game design, and reveal the discussions they are engaging in as they work, providing a link between what students talk about and what they write about in their Design Notebooks (Owensby, 2006; Thomas 2008). Observations also uncover the effects of the Design Notebook on both individual and dyad CAT development. Scholar artifacts (e.g., Design Notebook, physical prototypes, software prototypes, etc.) uncover how CAT capabilities develop over time. Pre-surveys assess Scholars' awareness of careers and opportunities related to CAT prior to SCAT. Post-surveys assess SCAT's impact on career and opportunity awareness related to CAT. Interview data will uncover both Scholars' understanding of what CAT is, how they have enacted it in the course of their activities, and their perceptions of themselves as problem solvers and game designers as well as facilitators' perceptions of Scholars' CAT capability development, their experience facilitating SCAT activities, and feedback to inform those activities.

References

- Anderson, J. R. (2000). *Cognitive Psychology and Its Implications: Fifth Edition*. New York: Worth Publishing.
- Anderson, J. R., Greeno, J. G., Kline, P.J. & Neves, D.M. (1981). Acquisition of problem-solving skills. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bandura, A. (1994). Self-efficacy. In R. J. Corsini (Ed.), *Encyclopedia of psychology* (2nd ed., Vol. 3, pp. 368-369). New York: Wiley.
- Bandura blog entry. *Ability and Capability*. Downloaded from Bandura's blog, <http://des.emory.edu/mfp/AbilityCapability.html>.
- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12, 307-359.
- Barnes, T., Richter, H., Chaffin, A., Godwin, A., Powell, E., Ralph, T., Matthews, P. & Jordan, H. (2007). *Game2Learn: A study of games as tools for learning introductory programming*. In Proceedings of SIGCSE2007.
- Barron, B., Schwartz, D.L., Vye, N.J., Moore, A., Petrosino, A., Zech, L., Bransford, J. D. & The Cognition and Technology Group at Vanderbilt (1998). Doing with understanding: Lessons from research on problem- and project-based learning. *Journal of the Learning Sciences*, 7(3&4), 271-311.
- Bayer, A. (1990). *Collaborative-apprenticeship learning: Language and thinking across the curriculum, K-12*. Mountain View, CA: Mayfield.
- Bransford, J. D., Brown, A.L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Brathwaite (2009). Interview on Women, Games and Design. Downloaded from Applied Game Design blog, <http://bbrathwaite.wordpress.com/2009/01/07/interview-on-womengames-and-design/>.
- 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: essays in honor of Robert Glaser*, 453-494. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Crawford, C. (2010) Glaser to Think: Algorithmic Thinking. In *the Journal of Computer Game Design* v.7.
- DiSalvo, B. J., Guzdial, M., Mcklin, T., Meadows, C., Perry, K., Steward, C. & Bruckman, A. (2009). Glitch Game Testers: African American Med Breaking Open the Console. In Proceedings of DiGRA 2009.
- Feltovich, P.J., Spiro, R.J., Coulson, R.L., & Feltovich, J. (1996). Collaboration within and among minds: Mastering complexity, individually and in groups. In T. Koschmann (Ed), *Computer systems for collaborative learning*, Hillsdale, NJ: Lawrence Erlbaum, 25-44.
- Fullerton, T., Swain, C., and Hoffman, S. (2004). *Game Design Workshop: designing, prototyping and playtesting games*. San Francisco, CA: CMP Books.
- International Society for Technology in Education – National Education Technology Standards (2007). *NETS for Students 2007*, downloaded from <http://www.iste.org/standards/netsfor-students/nets-student-standards-2007.aspx>.
- Irvine, M. (2008). "Survey: 97 Percent of Children Play Video Games". Downloaded from The Huffington Post, http://www.huffingtonpost.com/2008/09/16/survey-97-percent-ofchil_n_126948.html.
- Kafai, Y. B. (2006). Playing and Making Games for Learning: Instructionist and Constructionist Perspectives for Game Studies. In *Games and Culture*, 1(1), pp. 36-39.

- Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B. B., Gray, J., Holdbrook, J., & Ryan, M. (2003). Promoting Deep Science Learning Through Case-based Reasoning: Rituals and Practices in Learning By Design Classrooms. In Seel, N. M. (Ed), *Instructional Design: International Perspectives*, Lawrence Erlbaum Associates: Mahwah, NJ.
- Koschmann, T., Kelson, A.C., Feltovich, P.J., & Barrows, H.S. (1996). Computer-supported problem-based learning: A principled approach to the use of computers in collaborative learning. In T.D. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm* (pp. 83—124). Hillsdale, NJ: Lawrence Erlbaum.
- Kramer, Kramer, D. (2002). *Algorithms Should Mind Your Business*, downloaded from <http://www.outsourcing-russia.com/docs/?doc=680>.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Lenhart, A., Kahne, J., Macgill, A. R., Evans, C. & Vitak, J. (2008). Teens' gaming experiences are diverse and included significant social interaction and civic engagement. Report 202-415-4500 for the Pew Internet & American Life Project: Washington, D.C.
- Maloney, J., Burd, L., Kafai, Y., Rusk, N., Silverman, B., and Resnick, M. (2004). Scratch: A Sneak Preview. Second International Conference on Creating, Connecting, and Collaborating through Computing. Kyoto, Japan, pp. 104-109.
- Owensby, J.N. (2006). *Exploring the Development and Transfer of Case Use Skills in Middle-School Project-Based Inquiry Classrooms*. Completed Dissertation, Georgia Institute of Technology. Proquest (1115125971).
- National Research Council (2011). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press.
- Nersessian, N.J. (2008) *Creating Scientific Concepts*. Cambridge, MA: MIT Press.
- Polya, G. (1973). *How to Solve It: A New Aspect of Mathematical Method, 2nd Edition*. Princeton, NJ: Princeton University Press.
- Palincsar, A. & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117 – 175.
- Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer*. New York: Basic Books.
- Puntembekar, S., & Kolodner, J. L. (1998). The Design Diary: Development of a Tool to Support Students Learning Science By Design. *Proceedings of the Interational Conference of the Learning Sciences '98*, 230-236.
- Plutzik, N. (2010). "So, Only White Men Can Be Game Designers?" Downloaded from the NPR All Tech Considered blog, http://www.npr.org/blogs/alltechconsidered/2010/03/if_youre_not_white_and_male_yo.html.
- Repenning, A. and Ioannidou (2008). *Broadening Participation through Scalable Game Design*, ACM Special Interest Group on Computer Science Education Conference, (SIGCSE 2008), (Portland, Oregon USA), ACM Press.
- Repenning, A., Webb, D., & Ioannidou, A. (2010). *Scalable Game Design and the Development of a Checklist for Getting Computational Thinking into Public Schools*. Paper presented at the 2010 ACM Special Interest Group on Computer Science Education (SIGCSE) Conference, Milwaukee, WI.
- Roschelle, J. (1996). Learning by collaborating: Convergent conceptual change. In T. Koschmann (Ed.). *CSCL: Theory and practice of an emerging paradigm*, Mahwah, NJ: Lawrence Erlbaum, 209-248.
- Schneider, G. M. & Gersting, J. L. (2010). *Invitation to Computer Science, 5th Edition*. Boston, MA: Course Technology, Cengage Learning, 4-16.
- Thomas, J.O. (2008). *Scaffolding Complex Cognitive Skill Development: Exploring the Development and Transfer of Case Use Skills In Middle-School Project-Based Inquiry Classrooms*. VDM Publishing.
- Vygotsky, L. S. (1978) *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
- Wells, G. & Chang-Wells, G. L. (1992). *Constructing knowledge together*. Portsmouth, NH: Heinemann.
- Werner, L., Campe, S., & Denner, J. (2005). Middle school girls + games programming = Information technology fluency. *ACM special interest group in information technology education (SIGITE)*. Newark, NJ.
- Wing, J.M. (2006). *Computational Thinking*. In CACM Viewpoint, March 2006, pp. 33-35. Wing, J.M. (2010). "Computational Thinking". Presented at the Centre for Computational Systems and Biology, Trento, Italy, December 2010.

Acknowledgments

We are grateful for the generous support of the National Science Foundation (DRK-12 1150098).

Investigating Student Generated Computational Models of Science

Satabdi Basu, Anton Dukeman, John S. Kinnebrew, Gautam Biswas and Pratim Sengupta
Vanderbilt University, Nashville, TN, USA
{satabdi.basu, anton.dukeman, john.s.kinnebrew, gautam.biswas, pratim.sengupta}@vanderbilt.edu

Abstract: Computational Thinking (CT) is now considered a core competency in problem formulation and problem solving. In spite of the known synergies between CT and science education, integrating CT in K-12 science classrooms is challenging. This paper reports a teacher-led, multi-domain classroom study conducted with 6th graders using CTSiM – a learning environment for CT and middle school science. Pre-post comparisons show that students made significant gains, both in terms of computational thinking and the relevant science concepts. Furthermore, we developed measures for analyzing students’ computational models, and our results show that as challenges faced decreased, model accuracy not only increased in general, but also became a good predictor of individual learning gains.

Introduction

Computational Thinking (CT) refers to the concepts and representational practices involved in formulating and solving problems, designing systems, and understanding human behavior by drawing on computer science concepts like problem representation, abstraction, decomposition, recursion, simulation, and verification (Wing, 2010). CT parallels core practices of science education, and computational modeling is an effective vehicle for learning a variety of challenging science and math concepts (Harel, 1990; Guzdial, 1995; Vattam et al., 2011).

In spite of the known synergies between CT and science education, integrating computational programming in science classrooms is challenging. Furthermore, relatively little is known about students’ conceptual understanding and developmental processes in curricula that involve learning programming and computational modeling in conjunction with scientific concepts (Grover & Pea, 2013). Our study seeks to make contributions along both these dimensions. Over two years, we have developed CTSiM (Computational Thinking using Simulation and Modeling) (Sengupta et al., 2013), which provides students with an agent-based, visual programming platform where they can conceptualize given science phenomena, program agent-based simulations, and compare their simulations against expert simulations. We present a study in which a 6th-grade science teacher with no computer science background successfully led classroom instruction with CTSiM. The topics covered in the class included the study of kinematics (motion of objects) and fish tank ecosystems.

Specifically, we investigate the following research questions using students’ pre-post test scores and activity data from log files:

1. Does CTSiM help students learn both science content and CT skills synergistically? In particular, are the domain and computational measures correlated for (a) learning gains, and (b) correctness of primitives used in the student generated models as compared against the expert model?
2. Does the correctness of the student-generated models vary between different modeling activities and, if so, is this measure related to the number of challenges faced by students in each of the activities?
3. Can students’ computational models predict their learning gains from the pre- to post-test?

The CTSiM Learning Environment and Learning Activities

CTSiM adopts a learning-by-design pedagogical approach where students engage in scientific inquiry by participating in various phases of model building, simulation, and verification. Model building is done in two steps: conceptualization and computational modeling. First, students conceptualize the given science phenomenon by creating a concept map using a node-link interface. To construct this map, students select the appropriate entities of the system, their properties, behaviors, and interactions. Students then construct their computational models using a visual programming interface we call the ‘Construction world’ (see Figure 1). Students can select primitives from a library, which include domain-specific (e.g., “speed-up” in kinematics, “feed” in biology) and domain-general primitives (e.g., conditionals and loops). The set of domain-specific primitives are available to the students only if they have been correctly specified in the conceptualization interface. CTSiM provides scaffolds for algorithm visualization through simultaneous simulation and code step-through. Students can also verify their models by comparing a NetLogo (Wilensky, 1999) simulation of their models with an ‘expert’ NetLogo simulation for the phenomenon in the ‘Envisionment world’ (see Figure 2). However, they cannot inspect the expert computational model underlying the expert simulation. Identifying differences between the simulations helps students realize how to refine and correct their conceptual and computational models.

Currently, CTSiM comprises four modeling activities across two domains (Kinematics and Ecology):

Activity 1 – Shape drawing: Students generate algorithms to draw simple shapes to explore the relations between speed, acceleration and distance. They start by modeling shapes like squares and triangles where each segment of the shape is of the same length, implying constant speed. This also familiarizes them with programming primitives such as “forward”, “right turn”, “left turn”, and “repeat”. Then, students modify their algorithms to generate spirals of the same shapes, where each line segment is longer (or shorter) than the previous one, implying constant acceleration. Students also re-represent a speed-time graph by drawing shapes such that the length of segments in the shapes is proportional to the speed in the given graph.

Activity 2 – Modeling a roller coaster: Students progress from modeling simple shapes to modeling a real-world phenomenon – a roller coaster’s behavior as it moves along segments of a track: (1) up (pulled by a motor) at constant speed, (2) down (free fall), (3) flat (cruising), and (4) up again (without a motor). An expert simulation is provided and students try to build models to match the expert roller coaster behavior. This activity also involves use of more complex computational constructs like conditionals and nested conditionals.

Activity 3 – Modeling a fish tank (macro-level): Students progress from modeling a single agent with a single behavior in the previous activities to modeling multiple agents, each with multiple behaviors. They model part of a closed fish tank system – a macro-level semi-stable model with two agents: fish and duckweed. This involves constructing a model of the food chain, the respiration and reproductive processes of the fish and duckweed, and the macro-level elements of the waste cycle. The non-sustainability of the macro-model (the fish and the duckweed gradually die off), encourages students to reflect on what might be missing from the model, prompting the transition to Activity 4, after they identify the build-up of toxic fish waste as the culprit.

Activity 4 – Modeling a fish tank (micro-level): Students recognize the need for bacteria, and model two types of bacteria in the fish tank, which, through stages, help convert the fish waste to nutrients (nitrates) for the duckweed. They model the waste cycle where the Nitrosomonas bacteria convert the toxic ammonia in the fish waste to nitrites, and then the Nitrobacter bacteria convert the nitrites to nitrates, which help sustain the duckweed. The plots generated in the simulation provide students with an aggregate level understanding of the interdependence and balance among the different agents in the system.

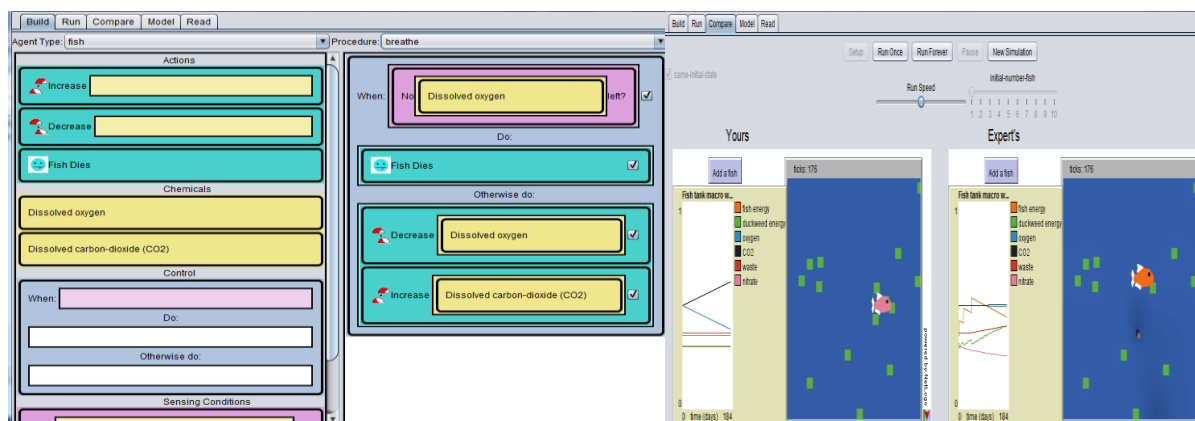


Figure 1. The Construction world with a ‘breathe’ procedure for ‘fish’ agents in a fish-tank unit

Figure 2. The Envisionment world for a fish-tank unit

Method

We conducted an in-class study using CTSiM with 25 6th-grade, middle Tennessee students (13 female, 12 male, average age=11.5). Each student worked individually on each of the four modeling activities described in Section 2. The study was run daily during the science period and the science teacher conducted all the instruction and led the classroom discussions. The teacher had no significant prior experience with programming and was introduced to CTSiM during two 90-minute training sessions before the study. He alternated between instructing (including facilitating class discussions) using CTSiM and having students work individually on the modeling tasks. One person from our research group was always present in the classroom to assist him with answering students’ questions, primarily those related to technical issues.

Students were administered pre- and post-tests in which the questions assessed Kinematics and Ecology knowledge, as well as Computational thinking skills. The Kinematics questions tested the concepts and relationships among speed, acceleration, and distance, including the generation and interpretation of speed-time graphs. The Ecology questions focused on the roles of species in the ecosystem, interdependence among the species, the waste cycle, and the respiration cycle. CT skills were assessed by asking students to use a provided set of primitives to model various scenarios that were not part of the CTSiM modeling activities.

Students worked on the Kinematics and Ecology units in 50-minute daily sessions for five days each. On Day 1, students took pre-tests for both the Kinematics and Ecology units. They worked on Modeling Activi-

ty 1 for days 2-4, and Activity 2 on days 5 and 6 before taking the Kinematics post-test on day 7. Students then moved on to the Ecology unit and worked on Activity 3 during days 8-10 and on Activity 4 on days 11 and 12. All students took the Ecology post-test on day 13. One student was absent for the Ecology post-test and is not included in the results presented for the Ecology unit. All student actions on the CTSiM system were logged for post-hoc analysis.

To evaluate the accuracy of the student-generated models in relation to an expert model, we calculated a vector-distance metric. Accuracy is measured as the distance between a student's model and the expert model, where a model distance of 0 implies that the model contained all the primitives in the expert model, and additionally contained no extraneous and incorrect primitives. Since our metric is based on a comparison with primitives in the expert model, and shapes in Activity 1 can be correctly modeled in a number of ways using different sets of primitives, vector distances were calculated only for models built in Activities 2-4.

Our distance measure for comparing models was derived from the Bag of Words metric (Piech et al., 2012). For each modeling activity, each procedure associated with each of the agents was represented as a bag of words, which was the collection of visual primitives used by a student in that procedure. Based on whether the primitives were computational or related to the domain being modeled, the bag of words was further labeled as computational or domain collections. To calculate a correctness measure, we normalized the size of the intersection of the student and expert models by the size of the expert model (summation of number of copies of each primitive). As a result, correctness scores are bounded between 0 and 1 inclusive. A score of 1 means the student had all the primitives necessary to build the model correctly. However, the bag of words does not analyze the structure further. For example, a student may have used all the right primitives, but the metric was not sensitive to incorrect assembly of the primitives, or if extraneous primitives were used in the procedure. Though a metric which neglects ordering of primitives might seem problematic, the specifics of our expert models and how our primitives were defined helped minimize the probability of such errors. For each modeling activity, we combined the correctness scores from its procedures using a weighted average based on the size of each procedure's expert model as shown in equation 1. Similarly, the computational and domain splits were used to get computational-specific and domain-specific correctness scores.

Since the correctness measure does not account for extraneous primitives, we also calculated an incorrectness measure for each procedure by counting the number of extra primitives in the procedure and normalizing it by the size of the expert model. A score of 0 meant no extra blocks were used in that procedure, while a score of 2 would mean that twice as many extraneous primitives as the total number of primitives required for a correct model had been used. We also combined the incorrectness scores from all procedures of a modeling activity using a weighted average as shown in equation 2. In order to assign an overall score to each model, we created a two-dimensional vector with the correctness and incorrectness measures and calculated its distance to the vector (1,0), as shown in equation 3.

Equation 1:
$$weightedAverageCorrectness = \frac{\sum_{each\ procedure} |user \cap expert|}{\sum_{each\ procedure} |expert|}$$

Equation 2:
$$weightedAverageIncorrectness = \frac{\sum_{each\ procedure} (|user| - |user \cap expert|)}{\sum_{each\ procedure} |expert|}$$

Equation 3:
$$distance = \sqrt{incorrectness^2 + (correctness - 1)^2}$$

Results

The intervention resulted in statistically significant learning gains for each of the two units, as seen in Table 1. Both units produced significant overall learning gains, as well as significant gains for domain content and CT skills, measured separately.

Table 1: Paired t-tests showing learning gains for the Kinematics and Ecology pre-post tests

Domain		Pre-test Mean(S.D.)	Post-test Mean(S.D.)	Max value	2-tailed <i>p</i> -value	Effect size
Kinematics (n=25)	Total	19.1(5.9)	25.9(7.1)	53	<0.0001	0.44
	Domain	15.3(4.3)	19.1(4.4)	36	<0.0001	0.34
	CT	3.8(2.5)	6.7(3.8)	17	<0.001	0.42
Ecology (n=24)	Total	15.3(5.9)	31.6(12.3)	62	<0.0001	0.65
	Domain	9.1(3.1)	19.7(7)	34.5	<0.0001	0.70
	CT	6.2(3.4)	12(6.3)	27.5	<0.0001	0.50

Correlations between Domain and Computational Measures

In order to investigate the potential synergy between learning CT and science in CTSiM, we first calculated the correlation between the domain and CT learning gains for both the science units. We used normalized learning gains to avoid potential ceiling effects for students with relatively high prior knowledge. The normalized gains were calculated by dividing the actual gains by the maximum amount that could have been gained based on pre-test scores. The correlations ($p < 0.01$) presented in Table 2 support our belief that CT and science concepts can be learned synergistically through computational modeling and simulation.

Table 2: Correlations between domain and CT learning gains

Normalized domain and CT gains	Pearson correlation coefficient (r)	2-tailed p-value
Kinematics (n=25)	0.52	<.01
Ecology (n=24)	0.56	<.01

To determine whether model accuracy scores for domain-related primitives and computational primitives were related in these activities, we calculated correlations between the vector distance scores for the domain-based and computational primitives for activities 2-4. Table 3 shows the correlations were high and significant ($p < .0001$). Since the domain and computational modeling scores are so strongly correlated, we only use the overall model accuracy score while reporting results in Sections 4.2 and 4.3.

Table 3: Correlations between domain and computational modeling scores (vector-distances)

Domain and computational vector distances	Pearson correlation coefficient (r)	2-tailed p-value
Roller coaster model (n=25)	0.72	<.0001
Fish macro model (n=24)	0.95	<.0001
Fish micro model (n=24)	0.93	<.0001

Variation of Student Model Correctness between Different Learning Activities

In order to answer Research question 2, we first compared the vector distance scores of students' roller coaster (RC) models (Activity 2) with that of students' fish tank macro models (Activity 3). A paired t -test comparing the vector distance scores for the RC models (mean vector distance = 0.23) versus the fish tank macro models (mean vector distance = 0.53) showed a significant decrease in accuracy ($t = 4.8$, $p < 0.0001$). However, a comparison of the vector distance scores between the fish tank macro models (mean = 0.53) and micro models (mean = 0.34) showed a significant increase in model accuracy ($t = 2.6$, $p < 0.05$). We hypothesized that the change in the model accuracy across units could be linked to the number of challenges students faced while developing the models for each of the units. For example, the RC modeling activity, though more challenging than Activity 1 (shape drawing), required modeling only one agent and a single procedure for that agent. On the other hand, when students switched to the fish-macro modeling activity, they had to deal with new domain content, multiple agents and multiple procedures for each agent. In a previous, one-on-one interview-based study, we coded student videos for the number of challenges faced by students in the different learning activities. This analysis confirmed that challenges increased from activities 2 to 3 and then decreased from activities 3 to 4 (Basu, et al., 2013). While no quantitative conclusions can be drawn from this pattern across two different studies, the basic modeling activities were the same in each study, so we believe that the change in the correctness metric observed between the units in this study can be linked to the corresponding change in the number of challenges students faced across units.

Student Models as Predictors of Learning Gains

In order to answer Research question 3, we calculated the correlations between students' pre-post learning gains and the correctness of their models as measured by the vector-distance scores. We used normalized learning gains and calculated correlations with overall gains, as well as gains for domain content and CT separately. Table 4 reports the correlations between the RC model and Kinematics learning gains, as well as correlations between Ecology learning gains and the fish macro and micro models, respectively. The 2-tailed p -values are listed only for the correlations that are statistically significant.

The low correlations and lack of significance for the RC and fish-macro models indicate that students' performance on these models is not predictive of the respective pre-post learning gains. However, students' fish-micro model's accuracy is strongly negatively correlated to their Ecology learning gains (note that a smaller distance value implies a better model). We believe that the lack of predictability of the RC and fish-macro models is linked to the initial complexity of those modeling activities, as reflected in the challenges faced by the stu-

dents. Learning during these activities involved learning about new computational abstractions and scientific concepts with students facing difficulties in both. However, we posit that the resolution of some of these challenges, specifically those related to domain learning, may not be reflected in the computational models. Though most students got better at using both domain and computational primitives as these activities progressed, their individual model accuracy is not predictive of learning gains. On the other hand, when students attempted the more familiar, and consequently less challenging, task in the fish-micro activity, their models were predictive of their learning gains with high confidence.

Table 4: Correlation between correctness of student generated models and pre-post learning gains

Vector distance	Kinematics normalized gains (n=25)	Kinematics Domain normalized gains (n=25)	Kinematics CT normalized gains (n=25)	Ecology normalized gains (n=24)	Ecology Domain normalized gains (n=24)	Ecology CT normalized gains (n=24)
RC model (n=25)	$r = -0.13$	$r = -0.01$	$r = -0.19$	-	-	-
Fish macro model (n=24)	-	-	-	$r = -0.06$	$r = -0.24$	$r = 0.17$
Fish micro model (n=24)	-	-	-	$r = -0.59$ ($p < .005$)	$r = -0.6$ ($p < .005$)	$r = -0.44$ ($p < .05$)

Discussion and Conclusions

As Grover & Pea (2013) point out, there are currently no accepted or standard assessments for measuring students' CT skills, and only a few studies have been conducted where students' progressive use of different programming artifacts has been studied over time. Our study makes a contribution along both these dimensions. Our CT assessment is based on computational modeling and students' abilities to generate a model or algorithm for a given scenario using computational and domain-specific constructs. Our vector-distance metric, as an extension to the well-known 'bag of words' measure, assesses the accuracy of students' models with respect to an expert model. Using this metric, we show that students' abilities to use domain-based and computational primitives are highly correlated, reflecting the synergy between CT and modeling science phenomena. Our results also suggest that when students find the challenges associated with a modeling activity manageable, their computational models generally improve and become better predictors of their pre-post learning gains.

References

- Basu, S., Dicks, A., Kinnebrew, J.S., Sengupta, P., & Biswas, G. (2013). CTSiM: A Computational Thinking Environment for Learning Science through Simulation and Modeling. In *Proceedings of the 5th International Conference on Computer Supported Education* (pp. 369-378). Aachen, Germany.
- Grover, S., & Pea, R. (2013). Computational Thinking in K-12 A Review of the State of the Field. *Educational Researcher*, 42(1), 38-43.
- Guzdial M. (1995) Software-realized scaffolding to facilitate programming for science learning. *Interactive Learning Environments*, 4(1), 1-44.
- Harel, I. (1990). Children as software designers: a constructionist approach for learning mathematics. *The Journal of Mathematical Behavior*, 9(1), 3-93.
- National Research Council. (2011). A framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC: The National Academies Press.
- Piech, C., Sahami, M., Koller, D., Cooper, S., & Blikstein, P. (2012, February). Modeling how students learn to program. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education* (pp. 153-160). ACM.
- Sengupta, P., Kinnebrew, J.S., Basu, S., Biswas, G., & Clark, D. (2013). Integrating Computational Thinking with K-12 Science Education Using Agent-based Computation: A Theoretical Framework. *Education and Information Technologies*, 18(2), 351-380.
- Vattam, S. S., Goel, A. K., Rugaber, S., Hmelo-Silver, C. E., Jordan, R., Gray, S., & Sinha, S. (2011). Understanding Complex Natural Systems by Articulating Structure-Behavior-Function Models. *Educational Technology & Society*, 14 (1), 66-81.
- Wilensky, U. 1999. NetLogo. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL.
- Wing, J. M. (2010). Computational Thinking: What and Why? *Link Magazine*

Acknowledgments

This work was supported by the NSF (NSF Cyber-learning grant #1237350).

Perspectival Computational Thinking for Learning Physics: A Case Study of Collaborative Agent-Based Modeling

Amy Voss Farris and Pratim Sengupta
Mind, Matter and Media Lab, Vanderbilt University – Peabody College
Email: amy.farris@vanderbilt.edu, pratim.sengupta@vanderbilt.edu

Abstract: We examine the process through which computational thinking develops in a perspectival fashion as two middle school students collaborate with each other in order to develop computational models of two graphs of motion. We present an interaction analysis of the students' discourse and computational modeling, and analyze how they came to a joint understanding of the goal of the modeling activity. We show that this process involves bringing about coherence between multiple perspectives: the object in motion, the computational agent, the other student, and graphs of motion.

Introduction and Background

Computational thinking, modeling and programming are now regarded as core epistemic and representational practices in K12 science and engineering (NGSS, 2013). Wing (2011) described computational thinking (CT) as the “thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent” (p. 1). The development of computational thinking is manifested in and deeply intertwined with representational practices such as programming and modeling, which in turn involve the design and development of computational abstractions such as defining patterns, generalizing from instances, and parameterization (Wing, 2006; Sengupta et al., 2013). In this paper, we present a theoretical framework for analyzing and understanding the role of perspectives, or points of view, in the development of collaborative computational thinking in the context of using agent-based programming and modeling for learning physics.

Prior research on learning physics using collaborative agent-based modeling has identified how shared understanding among student dyads develops through divergence and convergence of their conceptual understandings (Roschelle, 1992; Roschelle & Teasley, 1995). The students in these studies used a modeling environment called the Envisioning Machine, in which they directly manipulated a graphical simulation of the velocity and acceleration of a computational agent (similar to a Logo turtle) by altering the settings of the velocity and the acceleration vectors. Similar to Roschelle and Teasley, we also use an agent-based modeling environment, because previous studies have shown that students can indeed develop deep understandings of motion as a process of continuous change using such learning environments (Sengupta & Farris, 2012). However, in our study, the use of programming plays a central role in the students' interactions.

Theoretical Framework

At the broadest level, we seek to answer the following question: How does computational thinking begin to develop in a collaborative setting when students engage in agent-based programming in order to model the motion of an object as a process of change in distance and speed over time? To answer this question, we adopt the lens of *perspectival understanding* (Greeno & van de Sande, 2007; Greeno & MacWhinney, 2006). Greeno and van de Sande (2007) defined the construction of perspectival understanding as a process of constraint satisfaction. Following Thagard, they hypothesized that *coherence*—i.e., the compatibility or consistency of interacting representational elements (e.g., propositions, perspectives, meanings, etc.)—is the most general constraint (Thagard & Verbeurgt, 1998). Based on Greeno and his colleagues, we adopt the view that in a collaborative setting, manifold elements of the participating individuals' conceptions are taken up in a new joint understanding that is shaped by participants' points of view (POVs). For example, an interactant may operate from a point of view (POV) that is either enmeshed in the phenomenon or takes a depictive perspective, which views the phenomenon from a top-down or extrinsic perspective; however, the joint understanding that emerges during interaction can include expressed constituents of multiple POVs (Greeno & van de Sande, 2007).

Pedagogically, the ability of a user to take on the perspective of an actor within the system is an important affordance of agent-based models (Wilensky & Reisman, 2006), and virtual environments (Lindgren, 2012). In the context of understanding scientific phenomena that result from the aggregation of individual-level actions (e.g., change in motion over time, ecological interdependence, or formation of traffic jams) adopting an agent-perspective enables the learner to use their intuitive knowledge—which is often in the form of embodied knowledge (Papert, 1980)—in order to develop a deep understanding of how the collective phenomenon emerges from the aggregation of individual, agent-level actions. However, there is little understanding of the process through which learners *begin* to develop a perspectival stance. Our study offers a window into this process. We show how two collaborating learners shift across and negotiate multiple perspectives—the object in

motion, the other student, the programmable agent, and an aggregate, descriptive view of the graph—in order to interpret and model computationally the motion implied by the graph(s). In Thagard's terms, these perspectives represent the *elements* across which *coherence* is established through a process of shifts and negotiations between these perspectives (Thagard & Verbeurgt, 1998).

Data Collection and Analysis

The study took place in a classroom at a large private university in the mid-southern USA. The topic of the course was scientific modeling, and the course met on the mornings of six consecutive Saturdays during the regular academic year. Twenty students were recruited via web posting on the university website, ages 10 - 12 (Grades 5 and 6). They were enrolled on a first-come, first-served basis and had no prior programming experience. The second author taught the course, while the first author acted as a facilitator during the third day of the course, when the activity we report here took place. The children in the focal dyad are two 10-year-old males, Arnav and Liam (pseudonyms), who volunteered to do the task as a pair and agreed to talk about their thinking with each other while they worked.

We present in-depth analysis of 23 minutes of collaboration between the students. We use the case study method (Yin, 2009) and interaction analysis methods (Jordan & Henderson, 1995). The entire activity was video recorded by two cameras focused on the children and their screens. Additional data were collected as saved files and screen captures from the student computers. We created multimodal transcripts of the discourse, and coded them line-by-line for the perspective, or POV, of the speaker. Our analysis focuses primarily on the students' talk, and teaching moves and interaction between researchers and students are also described where relevant. We describe three significant episodes of interaction, and within each episode, we present an analysis of a salient, smaller segment of discourse. We also present a brief, summative analysis of the shifts and coherence in perspectives during the entire interaction.

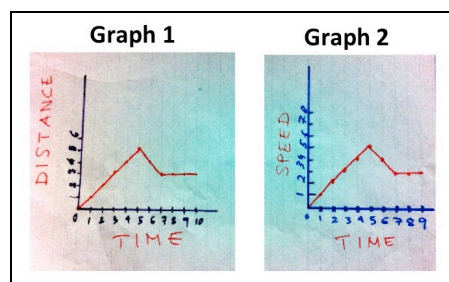


Figure 1. Distance-time (Graph 1) and speed-time (Graph 2) graphs

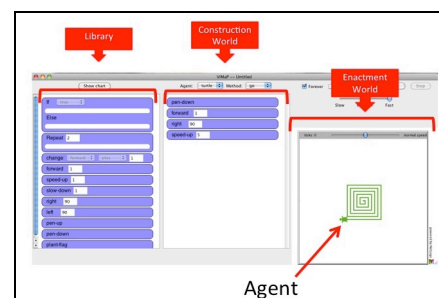


Figure 2. The ViMAP interface

The Activity and Instructional History

The goal of the focal activity was to generate an agent-based computational model for each of the two graphs, distance versus time (Graph 1, Figure 1), and speed versus time (Graph 2, Figure 1). In each model, students were asked to represent the motion of an agent in a manner that would match the motion represented in the respective graph. The graphs, as shown in Figure 1, were digitally projected on a large screen in the front of the classroom for the duration of the activity. The students were working with an agent-based modeling and visual programming environment called ViMAP (Sengupta, Farris, & Wright, 2012), where the user creates a computer program by dragging and spatially arranging programming commands from a library of command blocks (Figure 2). ViMAP commands were specifically designed to support domain-specific learning in kinematics (Sengupta et al., 2012). When the user runs the program, these commands are then acted out in representational space by a Logo turtle in a NetLogo microworld (Wilensky, 1999). In the activities leading up to the episodes reported here, students had developed ViMAP models of motion in which they generated turtle geometry shapes, in which step-size of the turtle represented the speed of motion. For example, the rectangular spiral in Figure 2 represents constant acceleration, as the step size of the turtle increases by the same amount every turn. In these activities, students developed models of several different kinematic phenomena, including motion on a roller coaster. All the students were familiar with line graphs from their regular science and math classes, although the activity we report here was the first time they encountered graphs of motion during the study.

Episode 1: Negotiating the Meaning of the Distance Graph

In the first episode Arnav and Liam engaged in negotiation of how to interpret Graph 1 (Figure 1). Arnav wanted to program the ViMAP turtle to reproduce the shape of the graph, which is a common novice approach to interpreting and representing graphs of motion (McDermott et al, 1987). Liam, however, argued that the graph meant that the distance was gradually going up. The segment below [2:29 – 2:48] illustrates Episode 1.

- Arnav: Yeah, but no. Distance 6 ((left arm used to approximate slope of first segment of graph)).
- Liam: It's GRADUALLY going up.
- Arnav: No, because the next one is speed.
- Liam: No, no, dude. Think about the first graph ((points to graph)). It's not you-something goes to 6, it's over time we go UP to 6. So then, (both?) is what we need to do.

Arnav's interpretation of Graph 1 ignores the role of time over which the change in distance takes place. We call this an *extrinsic graph* perspective, as he focuses on the shape of the line in Graph 1, and wants to program the ViMAP turtle to reproduce this line. In other words, even though he did consider the actions of the computational agent, these actions would simply reproduce the shape of the line in Graph 1—i.e., the graph extrinsic perspective was driving the actions of the agent. Liam countered with a more sophisticated interpretation: that the graph meant that something is “gradually going up” to 6. Liam interpreted the graph from an *intrinsic graph* perspective. That is, he focused on how the shape of the graph came to be, and to do so, he took on the perspective of actors (“we”), whose motion would gradually generate the shape of the graph (“...over time we go up to 6”). Also of interest here is that Arnav was unwilling to consider gradual change in distance as a representation of Graph 1, because Graph 2 shows speed—suggesting that he confused change in speed with change in distance over time.

Episode 2: How Far from Where You Are You're From

This episode began when Arnav and Liam asked for help from the first author (Amy), who in turn, asked them to explain to her what they inferred about the motion of the putative object from Graph 1 (Figure 1). The students described an object speeding up, slowing down, then staying the same. This verbal description matched their computational model, which used speed up and slow down commands. But the students were confusing speed with displacement; so, in order to encourage the students to think about displacement instead of speed, Amy then asked them if there are any two parts of the graph that are “the same.” Arnav said that times 3 and 7 (on the y-axis) indicated the same “distance from the starting point,” using the same words the instructor used with the whole class. After Amy left, the following exchange ensued [18:19 – 18:30], in which Arnav explains his interpretation of “distance from the starting point” to Liam.

- Arnav: It's not depending on how LONG it is ((uses pen as pointer to make an invisible line across the table surface)), it's depending on how far from where you are you're from, not how long the roller coaster actually is.
- Liam: Like he said, it how far FROM=
- Arnav: =No, so, if you're here ((right hand in front of chin))... and then you do a loop ((half-circle upward motion)), and you come back ((half-circle downward motion)), you'll be pretty much at the same distance as you started from.

In Arnav's final turn of talk, he uses the example of a loop, similar to a vertical roundabout segment during a roller coaster ride. His embodied definition and gestural enactment evidence his shift to thinking about the motion from the perspective of the agent in motion, which is at the same position in the beginning (time = 3) and end (time = 7) of the loop. Arnav and Liam's ViMAP model represented the changes in displacement over time, which evidences an agent perspective of getting further from, then closer to, a point in space. Arnav's construction fused inanimate physics entities with flexibly construed animate objects. The verb forms are enduring and simple present. Taking the perspective of a non-human agent allows the embodied action to transcend time and setting. Throughout this segment, Arnav's egocentric perspective was merged with the object perspective. Arnav and the (imagined) object in motion are conjoined in simultaneous, multiple constructed worlds: the here and now of the interaction, the visual representation, and the imagined physical processes. Ochs and colleagues have identified this kind of speech in the discourse of professional physicists (Ochs et al., 1996). In this utterance, the physics entities are distances and points, articulated with “how long it is” (where ‘it’ refers to the length of the roller coaster track), and “how far from where you are you're from,” in which “where you're from” is a point, and “how far you are from where you're from” is a distance. The only verb used in the utterance, “are,” is simple present and enduring (Ochs et al., 1996), as it allows the embodied action to transcend time and setting. Similar indeterminate grammatical constructions, along with gestural journeys through visual displays, as observed in physicists' discourse, constitute physicist and physical entity as co-experiencers of dynamic processes (Ochs et al., 1996).

Episode 3: Coordinating Relationships among Speed, Distance, and Time

In this episode, Arnav and Liam revised the loop to first make the turtle go down, and then up. Their thinking was that this motion could explain *both* Graph 1 and Graph 2, as the roller coaster car would get faster as it went down the downward half of the loop, and slow down as it went up the other half of the loop. The segment of

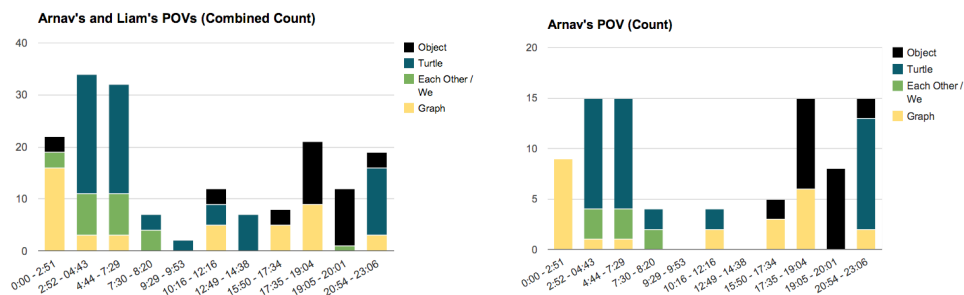
conservation we report here [22:19 – 22:46] results from their efforts to repair trouble: the period of acceleration in Arnav's proposed model did not correspond with the period of "getting further away" in Graph 1.

- Arnav: So we have to put it for 6 seconds ((adds repeating loop with a parameter of 6 six seconds))...forward 1, speed up 1 ((adds a forward command in the loop))
- Liam: Forward 1, speed up 1?
- Arnav: Yeah, because... Actually, we don't need the speed up, because, see, each point on that graph is one second, each point is one second, and all of them are the same (length?) we don't need a speed up.

Liam and Arnav referred back to the graphs, and then discussed the appropriateness of relevant ViMAP commands, "Forward <Step-size>" and "Speed-up <Change-in-Step-size>", for modeling the graphs. In doing so, they coordinated the change in distance as represented on the y-axis and time (x-axis) with the motion of an object, as Arnav's suggestion of a downward loop, gets taken up in joint action through *constructive listening* (Greeno & van de Sande, 2007) by Liam. Arnav then clarified that their ViMAP distance model did not require acceleration, because the length of the line (in the graph) during each second is the same, thereby implying that the object traveled the same distance during each second. This reflects coherence between an extrinsic graph perspective ("each point on that graph" and "all of them are the same (length?)"), an agent-perspective ("we don't need a speed up"), and an implied object perspective. For the first time in the activity, Arnav computationally parsed the difference between moving forward at the same speed and speeding up.

Shifts in Point-of-View During the Interaction: A Timeline

Figure 3A illustrates the counts of points of view, coded line-by-line throughout the interaction, for both Liam and Arnav. Intermittent periods of unrelated, off-task are not included in this analysis. Since a focus of this analysis is on Arnav's changing definition of distance from the starting point, his point-of-view counts are separated from Liam's and shown in Figure 3B. Four perspectives are represented: the object perspective, the agent perspective, the perspectives of one another, and perspectives based on the graph. The perspectives are color coded in black, teal, green, and yellow, respectively. At the beginning of the interaction (0:00 - 2:51), both students primarily took a graph perspective. The segment reported in Episode 1 illustrates Arnav's extrinsic graph perspective, while Liam adopted an intrinsic graph perspective. As the interaction continued, the students focused on generating a program for the turtle so that it would carry out specific actions (2:52 - 04:43), then edited the program and reran it (4:44 - 7:29), in order to generate the shape of the graph. Here, three perspectives are at work simultaneously: the graph (both extrinsic and intrinsic), turtle (agent), and egocentric (i.e., each other).



Figures 3A (left) and 3B (right). Arnav's & Liam's point-of-view, combined count (Figure 3A, left); and Arnav's points-of-view (Figure 3B, right).

Between 7:30 and 8:20, the instructor (second author) clarified the intended meaning of "distance" to the entire class, and Liam re-voiced this definition of distance to Arnav, who did not recognize any difference between that definition and the speed of the object. This was followed by an interaction between a facilitator and Liam, in which Liam explained the meaning of their (incorrect) model, but Arnav did not participate (09:29 - 09:53). In an extended interaction with the first author, (10:16 - 18:07), the students realized that their model showed changes in speed of the object in motion, but not its displacement. Episode 2, in which Arnav put forward a definition of "distance," began immediately after Amy left the students to change their model (18:09). In this segment, students integrated extrinsic and intrinsic graph perspectives with an object perspective (17:35 - 19:04). In Episode 3 (19:05 - 20:01 and 20:54 - 23:06), they begin enacting the motion (almost exclusively from an object perspective). However, when they began working on their program again, they coordinated the graph, agent perspective, and object perspectives (20:54 - 23:06). As these perspectives began to cohere, Arnav demonstrated his emerging understanding of the relationships among speed, distance traveled, and time, as shown in Episode 3.

Conclusion and Discussion

We have argued that when students engage in collaborative agent-based programming in order to model motion as a process of change over time, the development of computational thinking and learning physics co-occur through students' negotiations of multiple perspectives or POVs. *Coherence* between these perspectives serves as the *constraint* (Thagard & Verbeurgt, 1998), and the students' understanding of the relevant physics is propelled forward through a process of constraint satisfaction as these perspectives cohere. In our study, bringing about this coherence, in turn, was deeply tied to the children's computational *doing*, including agent-based programming and reflective discourse. In addition, the instructors' prompts also pushed the children toward particular points of view. Similar to previous studies using agent-based modeling (Wilensky & Reisman, 2006), our study also shows that the agent-perspective can indeed play a productive role in understanding the relevant scientific concepts; however, we also show that this perspective needs to be negotiated with others for conceptual growth. These other perspectives included the children's egocentric perspectives, perspectives based on the graphs, and that of the (imagined) physical entity in motion. Achieving coherence between all these perspectives enabled the learners to bridge what is happening *now* (i.e., the instantaneous position and speed of the object in motion) with what has happened *until now* (i.e., previous changes in the object's position and speed) – a feat that is challenging for even college-level physics learners (McDermott et al., 1987).

References

- Greeno, J. G., & MacWhinney, B. (2006). Learning as perspective taking: Conceptual alignment in the classroom. In *Proceedings of the 7th International Conference on Learning Sciences* (pp. 930-1).
- Greeno, J. G., & van de Sande, C. (2007). Perspectival understanding of conceptions and conceptual growth in interaction. *Educational Psychologist*, 42(1), 9-23.
- Lindgren, R. (2012). Generating a learning stance through perspective-taking in a virtual environment. *Computers in Human Behavior*, 28(4), 1130-1139.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1), 39-103.
- McDermott, L. C., Rosenquist, M. L., & Van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, 55(6), 503-513.
- Next Generation Science Standards: For States, By States*. (2013). The National Academies Press.
- Ochs, E., Gonzales, P., & Jacoby, S. (1996). When I come down I'm in the domain state: Grammar and graphic representation in the interpretive activity of physicists. *Studies in Interactional Sociolinguistics*, 13, 328-369.
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc.
- Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences*, 2(3), 235-276.
- Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning* (pp. 69-97). Springer.
- Sengupta, P., & Farris, A. V. (2012). Learning kinematics in elementary grades using agent-based computational modeling: A visual programming approach. *Proceedings of the 11th International Conference on Interaction Design & Children* (pp. 78-87).
- Sengupta, P., Farris, A. V., & Wright, M. (2012). From agents to continuous change via aesthetics: Learning mechanics with visual agent-based computational modeling. *Technology, Knowledge & Learning*, 17(1-2), 23 - 42.
- Sengupta, P., Kinnebrew, J. S., Basu, S., Biswas, G., & Clark, D. (2013). Integrating computational thinking with K-12 science education using agent-based computation: A theoretical framework. *Education and Information Technologies*, 18 (2), 351-380.
- Thagard, P., & Verbeurgt, K. (1998). Coherence as constraint satisfaction. *Cognitive Science*, 22 (1), 1-24.
- Wilensky, U. 1999. *NetLogo*. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, IL. URL: <http://ccl.northwestern.edu/netlogo>.
- Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep, or a firefly: Learning biology through constructing and testing computational theories—an embodied modeling approach. *Cognition and Instruction*, 24(2), 171-209.
- Wing, J. (2011). Research notebook: Computational thinking—What and why? *The Link Magazine*. Retrieved from <http://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why>
- Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). Sage.

Acknowledgements

Funding from the National Science Foundation (NSF Early CAREER # 1150230) is gratefully acknowledged. Thanks to Rogers Hall, Rich Lehrer & Kevin Leander for insightful comments on earlier drafts.

Children's Use of Inscriptions in Written Arguments About Socioscientific Issues

Sihan Xiao, William A. Sandoval, University of California, Los Angeles, Los Angeles, CA, USA
Email: shxiao@ucla.edu, sandoval@gseis.ucla.edu

Abstract: Engaging with science in everyday life typically occurs in the context of socioscientific issues, situations where science can play some role in helping people come to some judgment or decision. Such issues often require the interpretation of inscriptions (graphic representations of information) to evaluate claims. Research on socioscientific reasoning has shown that students often do not use relevant science or subordinate it to personal or social concerns. Research has not yet looked explicitly into how students interpret different types of inscriptions as potential evidence for use in socioscientific reasoning. This study compares upper elementary students' use of scientific and editorial inscriptions in their arguments about socioscientific issues. Using a pre/post design, we investigate students' use of inscriptions before and after a semester of science instruction focused on the coordination of claims and evidence.

Introduction

Engaging with science in everyday life often means being put in a position of having to evaluate claims. This includes not just deciding whether or not to believe a claim, but also judging its relevance to the issue at hand. Being scientifically literate, therefore, entails becoming a competent outsider to science that identifies personal relevance of particular scientific claims (Feinstein, 2011). Yet to complicate matters, in science-related issues people might encounter (e.g. What to do about climate change? Should I buy organic foods?), arguments for or against a particular position combine scientific and editorial claims. Educators have long wanted school science to prepare students to engage with science productively, but we have little evidence this happens. A good deal of research on science-related issues (i.e. so-called socioscientific issues) has explored questions around when and how students use relevant scientific knowledge to reason about an issue. Yet, one's knowledge on a topic is very often insufficient to contribute to a reasoned judgment. The problem becomes, then, how do students evaluate claims they encounter about which they know little? What do they think counts as evidence?

Socioscientific issues (SSI), those where science can inform the judgments or decisions we have to make in such fields as energy use, the environment, or health, are an arena in which students' thinking about evidence can be linked with their daily life. We hope students use scientific knowledge to help them come to reasoned arguments and decisions (e.g. Kolstø, 2001; Zeidler, Sadler, Simmons, & Howes, 2005), yet studies show students often rely more on their personal experience, value judgments, and moral concerns (Grace & Ratcliffe, 2002; Jiménez-Aleixandre & Pereiro-Munoz, 2002; Yang, 2004). Why is that the case? Previous research has proposed several explanations. First, a lack of understanding of scientific concepts makes them unlikely to be used (e.g. Hogan, 2002; Sadler & Fowler, 2006). Alternatively, students may understand applicable science knowledge but not consider it relevant to the situation at hand (e.g. Feinstein, 2011; Nielsen, 2012c). Further, it is only natural that in everyday situations science serves merely to frame the focus of seeing while values play a more significant role in reasoning (Nielsen, 2012a, 2012b). This view opens the possibility that students make judgments about the appropriate role of scientific knowledge in SSI.

This study explores the bases on which students might make such judgments, by focusing on a neglected yet critical aspect of SSI reasoning: the use of inscriptions as evidence. Inscriptions are graphical representations of information, such as photos, diagrams, graphs, tables, etc., mobilized in support of arguments (Latour, 1990). In science, inscriptions are a crucial means for representing the world and supporting or refuting explanations about how it works. In the classroom, as in the science lab, they mediate efficient communication, foster evaluation of scientific knowledge, and support conceptual understanding (Cobb, 2002; Roth & McGinn, 1998). Yet little is known about how students use inscriptions in SSI, even though when we face science in our daily life, we encounter inscriptions often simultaneously. For instance, a news article about hybrid cars may contain a diagram of CO₂ emissions by gasoline-powered cars and a photo of horrible traffic on a California freeway. A picture of a peaceful river farm may be attached to the grocery shelf as advertisement for organic foods. How do students perceive the role of these different inscriptions in reasoning about energy use or organic foods? Do they interpret such images as evidentiary, or as evocative of particular feelings or ideas?

To approach these problems, three research questions guide this study. First, what inscriptions do elementary-aged students use to argue about a SSI? Second, how do they use these inscriptions? Third, how appropriately do they use these inscriptions? Using a pre-/post-test design, we find patterns in which students use different types of inscriptions, in varying ways, to support their decision-making at the beginning and the end of a semester of instruction geared towards coordinating evidence and claims. We discuss potential

implications for science teaching that fosters scientific argumentation and productive engagement with socioscientific issues.

Methods

Participants

This study took place in a public university laboratory school located in a large, ethnically and economically diverse city. The student population was 36% Caucasian, 28% Multiethnic, 20% Latino, 9% Asian, and 7% African-American. One intact 5th-grade classroom was recruited for this study, involving 2 science teachers and 56 students aged 10-12 (30 boys, 26 girls). Students of this age have basic knowledge of science yet declining interest (Osborne, Simon, & Collins, 2003), and SSI addresses this very feature by approaching everyday situations, thus makes the sampled classroom analytically informative.

Setting

The two teachers of the sampled students were participating, with the other teachers at their grade level, in the second year of an ongoing “work circle” (Shrader, Williams, Walker, & Gomez, 1999) with researchers to develop practices to help students coordinate claims and evidence across the curriculum. In science, this included learning routines of productive science talk (Michaels & O’Connor, 2012). Teachers organized projects and short-term group work to encourage students to share their ideas and work toward consensus explanations for the science topics under study.

Measures

We used two written tasks as pre-/post-test design to measure participants’ performance at the beginning and end of the fall semester. In each task, participants individually read a single-paragraph scenario of a socioscientific issue (alternative energy use in one task and genetically modified food in another). The scenario presents a dilemma about which participants have to make a personal decision. We oriented participants to personal decision making because it represents how people commonly encounter and use science in everyday life. Participants were then asked to write down their decision and explanation, using inscriptions provided.

Each scenario includes 8 inscriptions participants may use to support their written response. Four are scientific inscriptions: 2 tables and 2 diagrams. Tables refer to the arrangement of information, mostly numbers, in rows and columns. Diagrams refer to graphs, models, or maps. Four editorial inscriptions include 2 photos and 2 drawings. Inscriptions were selected by reviewing a variety of news reports related to the topic of each scenario. Each inscription includes its original caption from the source article to help students comprehend the information it conveys. We divided inscriptions into scientific and editorial because students use both scientific and editorial resources to reason about SSI (Albe, 2008; Wu & Tsai, 2007), but it is not yet known whether or not students prefer one form over the other, or if they choose different kinds of inscriptions to support different types of claims. Finally, we chose 2 inscriptions for each type (8 in total) because it is an amount that is feasible to handle yet maintains enough variation.

Procedures

We administered the pretest early in the fall semester (i.e. mid-October 2013) and the post-test following the semester (i.e. late January 2014). Presentation of tasks is counter-balanced: participants are randomly divided into two groups in each class, with half of them completing the energy use task as the pretest and the GMO task as the posttest, and the other half completing the tasks in the opposite order. This counter-balancing mitigates potential effects from task ordering or difficulty. Twenty-seven students completed the energy task at pre-test, then the GMO task; 29 completed the GMO task first, followed by the energy task.

We gave oral instruction to participants prior to the test. They were then asked to read the single-paragraph scenario and make a personal decision (either “if you lived in this city, which power plant would you vote for” or “would you buy GMO foods on a regular basis”), write down their decision and explanation, and use inscriptions provided as many as necessary to support their decisions. No discussions were permitted during the testing, but they were allowed to ask us any questions regarding task comprehension.

Analysis and Findings

In what follows, we present our findings of content analysis on the written tasks. For descriptive purposes, we show separate data in each task as well as combined data to delineate a comprehensive picture of students’ pre-to-post performance.

What Inscriptions Do Students Use?

For the first question, we counted the frequencies of students’ citing each inscription. As we predefine 4 types of inscriptions and distinguish scientific ones from editorial ones, we documented how often they use each type,

and specifically, which they use the most. These frequencies depict a general picture of students' use of inscriptions in the task, and in particular, whether students tend to use inscriptions that are more relevant to everyday life (e.g. photos, drawings). The 56 students cited the given inscriptions 117 times in the pretest (63 in the energy task and 54 in the GMO task) and 127 times in the posttest (64 in the energy task and 63 in the GMO task; see Table 1). Across both tasks, 31% of cited inscriptions in the pretest and 32% in the posttest were scientific. As a group, students were more likely to use editorial inscriptions than scientific ones in both the pretest, $\chi^2(1, N=56) = 10.470, p < .01$, and the posttest, $\chi^2(1, N=56) = 9.646, p < .01$. Thus, students preferred editorial forms of evidence to scientific ones, which suggests an orientation to the problems that is more social, in which the relevant science is subordinated to other kinds of concerns (e.g., economic, moral, etc.).

Table 1. Frequency and percentage of types of inscriptions cited across Energy and GMO tasks

	Diagram	Table	Drawing	Photo	Total
Pretest	30 (26%)	11 (9%)	35 (30%)	41 (35%)	117
Posttest	31 (25%)	15 (12%)	26 (20%)	55 (43%)	127

More specifically, in the energy task, students frequently cited: 1) a drawing portraying nuclear reactors as green leaves with the title "Nuclear Energy Clean Air" (14 times in the pretest), 2) a photo of working coal-burning power plants in Ukraine (18 in the posttest), and 3) a photo of nuclear reactor explosions in the 2011 earthquake in Japan (13 in the pretest and 14 in the posttest). In the GMO task, on the other hand, students frequently cited a line chart (i.e. diagram) that illustrates the increasing rates of inflammatory bowel disease after GMOs were introduced in U.S. public diet (16 in the pretest and 18 in the posttest). Often cited was also a photo in which a little girl was holding up a sign during a street protest against GMO. Overall, Table 1 shows students tended to use photos the most, and tables the least, perhaps because they viewed photos as more common in and relevant to their everyday life. A semester of instruction did not change such perceptions.

How Do Students Use Inscriptions?

To understand how students deploy specific inscriptions within an argument, we adapted a coding scheme developed to categorize levels of "rhetorical reference" to inscriptions of data (Sandoval & Millwood, 2005). These levels range from less to more appropriate, from a scientific perspective, kinds of reference (examples verbatim from students' responses): 1) *pointing*, referring to some inscription as relevant without saying why (e.g. "I would vote for nuclear plant energy because of image 5."); 2) *description*, describing the content of an inscription without linking it to a claim (e.g. "No [GMOs] because in pic 8, a food turned into a crazy animal. This animal can kill the humans and multiply."); 3) *assertion*, asserting an inscription "shows" or "support" a claim without explaining how (e.g. "I would vote for the coal power plant because it seems safer as shown in image #5 compared to image #8."); and 4) *interpretation*, explicating how specific features of an inscription relate to a claim (e.g. "In picture 2 it states that the rate of inflammatory disease has increased a lot since before GMO started. It pretty much doubled since GMO, and if it keeps up, our community will be full of disease."). A sufficient use of inscription should involve as much interpretation as possible.

The level of rhetorical reference is coded for each citation of an inscription. The scale of these levels is ordinal, so we cannot simply create a sum of scores for each participant. Instead, there are two ways to use this scheme to understand students' references to inscriptions. At the aggregate level, we computed the frequencies of each type of reference across the corpus of written responses (see Table 2).

Table 2. Frequency and percentage of rhetorical use of inscriptions

	Pointing	Description	Assertion	Interpretation	Total
Pretest	65 (56%)	25 (21%)	15 (13%)	12 (10%)	117
Posttest	49 (39%)	28 (22%)	37 (29%)	13 (10%)	127

Chi-squared tests indicate that students were most likely to merely "point" to particular inscriptions, $\chi^2_{pretest}(3, N=56) = 61.427, p < .001$, and $\chi^2_{posttest}(3, N=56) = 21.756, p < .001$. This echoes Sandoval & Millwood's (2005) earlier claim that students may see data as self-evident, with no need for interpretation. Compared with those of the pretest, on the other hand, the percentage of pointing decreases and that of assertion increases, both in a notable way, in the posttest.

We also computed ratio scores of the four levels of rhetorical reference for each student (e.g. one used 5 inscriptions: 60% pointer, 20% description, and 20% assertion). Non-parametric correlations show that, in the pretest, only the level of interpretation is significantly correlated with scientific types of inscriptions ($r = .247, p < .05$). Though quite mild, this might suggest that students see the scientific types of inscription as more open to

interpretation, and thus more in need of being explicitly justified in relation to their claims. The percentage of scientific inscriptions does not significantly correlate with any level of reference in the posttest.

How Appropriate Is Students' Use of Inscriptions?

The third question focuses on the functional role of inscriptions in relation to their content. We anticipated two major functional roles, varying by the type of inscription: eliciting values or emotions vs. providing data. For instance (from students' responses), a claim, "The nuclear power is so dangerous, people will get injured or killed, like in [a photo of nuclear reactor explosion in Japan]," was coded as eliciting emotions or values, for the photo is used to show the horror of a nuclear accident. To be clear, it is not that such a photo is not evidentiary; the value of such images is exactly that they show quite viscerally a risk of nuclear energy. Yet, such an image is used, in this case, to elicit an emotion rather than to convey evidentiary information. An example of providing data is: "It also looks as in [a diagram that shows the increasing rates of inflammatory bowel disease after GMOs were introduced in U.S. public diet] that GMO foods are possibly causing inflammatory bowel disease." The inscription provides data to support the causal claim.

Our assessment of appropriateness entails an explicit correspondence between contents and functions; that is, using scientific inscriptions to provide data about the state of the world, and using editorial inscriptions to elicit values, emotions, or moral concerns. Taking up this stance, we counted as appropriate all instances where a claim corresponded to the inscription cited as evidence, shown in Table 3 for both tests.

Table 3. Frequency and percentage of functional use of inscriptions (types of inscriptions : functions)

	Appropriate Use			Inappropriate Use		
	Sci : Data	Edit : Values	Total	Sci : Values	Edit : Data	Total
Pretest	27 (23%)	23 (20%)	50 (43%)	13 (11%)	54 (46%)	67 (57%)
Posttest	35 (28%)	40 (31%)	75 (59%)	11 (9%)	41 (32%)	52 (41%)

Chi-squared test on the 4 combinations of functional use in the pretest shows that students were more likely to use editorial inscriptions to provide data, $\chi^2 (3, N=56) = 31.479, p < .001$. It reveals a mismatch between the content of inscription a student use and its function being deployed at the beginning of study. Yet their appropriate vs. inappropriate use in total did not significantly differ. In the posttest, on the other hand, students were more likely to use inscriptions in appropriate ways, $\chi^2 (1, N=56) = 4.165, p < .05$. From Table 3, it is clear that the increase of appropriate use could be attributed to the increase of using editorial inscriptions to elicit values and the decrease of using them to provide data, suggesting students' improvement in functionally deploying editorial inscriptions in their arguments.

In addition, non-parametric correlations reveal that, in the pretest, appropriateness of functional use is significantly correlated with percentage of scientific inscriptions ($r = .459, p < .01$), which indicates that students who used more scientific inscriptions were more likely to use inscriptions appropriately. The appropriate use is also modestly correlated with the rhetorical level of interpretation ($r = .283, p < .05$), which may suggest a covariant tendency of coordinating the content of an inscription with its function and interpreting an inscription to support a claim. Counterpart analyses on the posttest, again, do not yield significant results.

Discussion

This study addresses an immediate question that has not been explored in sufficient depth in studies of SSI reasoning. That is, how do students deploy different types of inscriptions in their arguments about SSI, given that in everyday situations, SSI often come with a variety of inscriptions? Our analyses show that at the beginning of the semester, students tended to: 1) use editorial inscriptions, especially photos, more than scientific ones, 2) only "point" to an inscription without making clear how it is related to a claim, and 3) use editorial inscriptions as data to support a claim. A semester of science instruction that highlighted scientific argumentation, coordinating evidence and claims in particular, did not significantly change *what* inscriptions students tended to use, but, more or less, *how* students used inscriptions to argue about SSI.

Agreeing with Nielsen (2012a, 2012c) on SSI argumentation, we do not expect students to use only scientific inscriptions and make only scientific arguments. In everyday situations, people make decisions not about "what is true," but about "what to do," in which non-scientific considerations play a significant role. Yet we do argue that school should teach students to explicitly distinguish evidentiary demands in arguments, supporting the veridicality of claims, from other rhetorically meaningful uses of inscriptions, and to understand the strengths and weaknesses of kinds of inscriptions for supporting particular kinds of claims.

Our analyses shed needed light on children's understanding of different forms of evidence and their uses. The prevailing view in science education is that students see data as objectively self-evident, although young students are capable of learning the need to justify evidence-claim relations (Ryu & Sandoval, 2012). This study extends this work, through our contrast between scientific and editorial inscriptions. The findings we

present here suggest that, students may be capable of learning to use editorial, if not scientific, inscriptions more appropriately, and to take the self-evidence of data less for granted.

This study opens up a new direction of research on public engagement with science focused on the features of potential forms of evidence and their influence on how students coordinate claims and evidence. Although our design does not enable causal inferences from classroom instruction to changes in pre-/post-tests, it takes a very first step to understanding use of inscriptional evidence in arguing about socioscientific issues. As educators become more concerned with how students warrant claims, scientific and otherwise, this study aligns with efforts to understand children's ideas of what makes a good warrant for particular kinds of claims. Because science knowledge is inevitably subordinate in socioscientific issues, by their nature, it is important to develop detailed accounts of how students make sense of what counts as "scientific" and what does not. This, in other words, is to explore children's capacity for becoming competent outsiders to science (Feinstein, 2011) by identifying relevant scientific resources and deploying them in their own everyday judgments and decisions.

References

- Albe, V. (2008). When Scientific Knowledge, Daily Life Experience, Epistemological and Social Considerations Intersect: Students' Argumentation in Group Discussions on a Socio-scientific Issue. *Research in Science Education*, 38(1), 67–90.
- Cobb, P. (2002). Reasoning With Tools and Inscriptions. *Journal of the Learning Sciences*, 11(2-3), 187–215.
- Feinstein, N. W. (2011). Salvaging science literacy. *Science Education*, 95(1), 168–185.
- Grace, M. M., & Ratcliffe, M. (2002). The science and values that young people draw upon to make decisions about biological conservation issues. *International Journal of Science Education*, 24(11), 1157–1169.
- Hogan, K. (2002). Small groups' ecological reasoning while making an environmental management decision. *Journal of Research in Science Teaching*, 39(4), 341–368.
- Jiménez-Aleixandre, & Pereiro-Munoz, C. (2002). Knowledge producers or knowledge consumers? Argumentation and decision making about environmental management. *International Journal of Science Education*, 24(11), 1171–1190.
- Kolstø, S. D. (2001). Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socioscientific issues. *Science Education*, 85(3), 291–310.
- Latour, B. (1990). Drawing things together. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice* (pp. 19–68). Cambridge, MA: MIT Press.
- Michaels, S., & O'Connor, C. (2012). *Talk science primer*. Cambridge, MA: TERC.
- Nielsen, J. A. (2012a). Arguing from Nature: The role of "nature" in students' argumentations on a socio-scientific issue. *International Journal of Science Education*, 34(5), 723–744.
- Nielsen, J. A. (2012b). Co-opting Science: A preliminary study of how students invoke science in value-laden discussions. *International Journal of Science Education*, 34(2), 275–299.
- Nielsen, J. A. (2012c). Science in discussions: An analysis of the use of science content in socioscientific discussions. *Science Education*, 96(3), 428–456.
- Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049–1079.
- Roth, W.-M., & McGinn, M. K. (1998). Inscriptions: Toward a Theory of Representing as Social Practice. *Review of Educational Research*, 68(1), 35–59.
- Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children's epistemic understanding from sustained argumentation. *Science Education*, 96(3), 488–526.
- Sadler, T. D., & Fowler, S. R. (2006). A threshold model of content knowledge transfer for socioscientific argumentation. *Science Education*, 90(6), 986–1004.
- Sandoval, W. A., & Millwood, K. A. (2005). The Quality of Students' Use of Evidence in Written Scientific Explanations. *Cognition and Instruction*, 23(1), 23–55.
- Shrader, G., Williams, K., Walker, L., & Gomez, L. (1999). Work in the "work-circle": A description of collaborative design to improve teaching practice. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Wu, Y., & Tsai, C. (2007). High School Students' Informal Reasoning on a Socio-scientific Issue: Qualitative and quantitative analyses. *International Journal of Science Education*, 29(9), 1163–1187.
- Yang, F.-Y. (2004). Exploring high school students' use of theory and evidence in an everyday context: the role of scientific thinking in environmental science decision-making. *International Journal of Science Education*, 26(11), 1345–1364.
- Zeidler, D. L., Sadler, T. D., Simmons, M. L., & Howes, E. V. (2005). Beyond STS: A research-based framework for socioscientific issues education. *Science Education*, 89(3), 357–377.

An Interactional Analysis of Gaze Coordination during Online Collaborative Problem Solving Activities

Murat Perit Çakır and Selin Deniz Uzunosmanoğlu
Informatics Institute, Middle East Technical University, Ankara, Turkey
perit@metu.edu.tr, sdeniz@metu.edu.tr

Abstract: Simultaneous tracking of eye gaze patterns of two or more students while they are engaged in a collaborative learning activity has recently attracted increasing interest in the learning sciences literature. The dual eye tracking paradigm allows researchers to quantify the degree of gaze overlap over the course of a learning activity, which is often treated as an estimate of the degree of joint attention among peers. In this paper we employ the dual eye tracking paradigm in the context of online collaborative geometry problem solving to quantify the degree of joint attention evidenced in gaze patterns, and explore what qualitative factors facilitate gaze coordination by focusing on the interactional organization of sequences in which gaze recurrence takes place.

Introduction

Investigating ocular correlates of learning with eye tracking methods has become an important focus of research in the learning sciences, particularly in the area of multi-media learning (van Gog & Scheiter, 2010). The majority of eye tracking studies in this field consider the individual as the primary unit of analysis. Such studies often control for the content and the pace of the instructional materials presented to the learners, measure the eye fixation patterns of the learners to identify what parts of the interface they allocate their attention, and finally relate these measures to learning outcomes to identify which presentation strategies were more effective. Since fixation sequences by themselves do not reveal what subjects are thinking or whether they are really paying attention, think-aloud protocols are frequently employed to interpret the eye-tracking data. However, since learning to see relevant visual structures at a scene and to associate them with appropriate linguistic terminology are primarily socially shaped processes (Goodwin, 1994), think-aloud alone may not be adequate to understand how students learn to allocate their visual attention to relevant objects in a particular domain.

The dual eye tracking paradigm, where the eye gaze of multiple subjects are monitored simultaneously while they are collaborating on a shared task, provides important opportunities to investigate the mechanisms underlying joint attention (Nüssli & Jermann, 2012). The *degree of overlap* or *cross-recurrence* among fixation sequences of interlocutors provides important information regarding to what extent the participants mutually orient to each other and to the objects in the shared scene (Richardson, Dale & Kirkham, 2007). A general finding in this emerging literature is that there is a positive correlation between the degree of gaze recurrence and collaborative learning outcomes (Jermann et al., 2011). In the context of well-defined tasks, it may even be possible to predict whether a dyad will succeed in solving a puzzle based on the patterns of synchronization detected in their raw gaze and speech data via machine learning algorithms (Nüssli, 2011). These findings also motivated the design of learning environments where users' gaze information is visualized on the screen in real-time as an awareness mechanism. A recent study on such a joint learning environment reported significant gains in gaze coordination and learning outcomes (Schneider & Pea, 2013).

Current studies that employ the dual eye tracking paradigm in the context of collaborative learning focus on devising quantitative metrics that reveal the type and the quality of collaboration, without particularly focusing on the role sequential organization of actions/utterances play in the achievement of joint attention. In this paper we focus on the relationship between changes in gaze recurrence and the sequential organization of interaction in a chat environment called Virtual Math Teams (VMT), where we asked dyads to collaboratively work on dynamic geometry problems. In this short paper we present our preliminary findings regarding (a) how measures of gaze recurrence relate to the overall success of collaboration in the context of collaborative math problem solving online, and (b) how variation in gaze recurrence relate to specific interactional events.

Methodology

This study was conducted with 18 volunteered university students at METU in Ankara, Turkey. Participants were grouped into 9 pairs and asked to collaboratively work on 10 geometry problems by using the VMT environment (Figure 1). Before the experiments participants received a short training that illustrated basic features of VMT. VMT is a CSCL environment that provides chat, shared whiteboard and wiki features to support discussions on math topics online (Stahl, 2009). Most recent version of VMT supports GeoGebra-based dynamic geometry constructions. The tasks were selected among activities presented in Stahl (2013), which included ruler and compass constructions such as constructing a perfect equilateral triangle, an isosceles triangle, and a perfect hexagon in GeoGebra. Participants coordinated their work and discussed strategies via the

chat, and took turns to construct the desired dynamic geometry representations. Dyads were also asked to come up with an informal proof detailing why they thought that their construction was correct upon completion of each problem. Teams posted their proofs on a separate summary tab. During the online sessions eye movements of the participants were recorded with two desktop eye trackers (Tobii T1750 and Tobii T120) at a resolution of 50Hz.

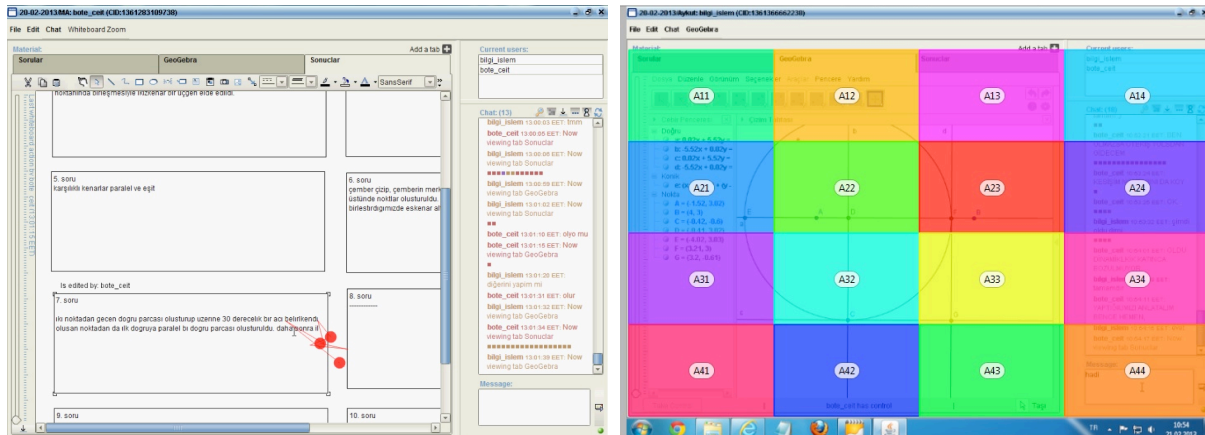


Figure 1. A screen shot of the VMT environment is presented on the left. The screen shot on the right shows the 16 areas used for partitioning the screens into areas of interest.

Eye-tracker data were analyzed both quantitatively and qualitatively. Gaze recordings were split into synchronized excerpts in the Tobii Studio Software. The screen recording of each participant was then split into a 4x4 matrix of non-overlapping areas of interests (AOIs) as shown in Figure 1. 12 of the AOIs covered the area where the dynamic geometry representations were constructed and 4 AOIs covered the chat component on the right. The time course of gaze duration falling over each AOI was then visualized by a custom software written in Java. The visualization provides a scarf plot that represents which AOI participants looked at, when their eye gazes overlapped, and over which AOI (Figure 4). The third row in the visualization indicates gray bars that highlight those time points where both participants' gaze fall on the same AOI. The whitespace indicate those instances where the participant is not looking at any of the AOIs (e.g. during typing).

Each of these 16 regions was considered as an approximation of the part of the screen over which the participants were attending to at any given time. While monitoring the shared scene users either moved their eye gaze with saccadic movements or fixated on specific locations by keeping their eyes still over a location. During a fixation event the fovea, which is the part of the retina that has the highest concentration of light sensitive cells, is oriented towards the fixated location. The fovea covers approximately 1-2 degrees of visual field, and at a distance of 65 cm from a screen, 1-2 degrees of visual field corresponds to a circular area with a diameter of 2.2 cm on the screen (Duchowski, 2007). The visual attention span is considered to cover a larger area covered by the foveal projection, as evidenced in dual-task experiments (Holmqvist et al., 2011). Since 17 inch displays with 4:3 aspect ratios were used during our experiments, the width and the length of the screen was 35 cm and 26 cm respectively. Splitting this area into 16 equal non-overlapping rectangular AOIs covers an area approximately 9 cm wide and 7 cm long. In this study this rectangle was considered as a rough approximation of where the person is attending to at any given time. Using the same AOI definitions on both screens over synchronized eye-tracking data allowed us to quantify gaze overlaps. Since the screens were divided equally, the probability that one of the participants allocate their attention on a given AOI is 1/16. Assuming independence of random gaze events, the possibility that two people allocate their attention on the same AOI is $1/16 \times 1/16 = 1/256$. So, the likelihood of having systematic gaze overlap among 2 people by chance is 0.004, despite the low spatial resolution provided by 16 AOIs.

In addition to the scarf plot, the software also computes a gaze recurrence plot for each problem attempted by every pair. The plots computed for each pair were then combined into a single recurrence percentage plot by taking the average of all corresponding data points coming from the plots for individual problem solving segments (see Figure 2). This yields a recurrence plot that summarizes the gaze patterns over all problem solving segments of the particular pair. In the combined summary plot, data points range from -4000 msec to +4000 msec with a 100 msec resolution. Point 0 indicates the recurrence percentage for synchronous gaze overlaps, -200 indicates the recurrence percentage in which B's gaze follows A's with a 200 msec delay with respect to A, and vice versa. The blue line shows the baseline recurrence level, which is computed over randomly shuffled gaze data. The vertical lines are the standard error bars, which indicate the amount of deviation in the data for the corresponding time.

Results and Discussion

In order to test the relationship between group performance and the degree of gaze overlap, 9 pairs were split into 3 achievement groups in terms of the number of problems they could solve correctly, whether they worked on tasks in a coordinated way and whether both participants contributed to the discussion. The degree of gaze overlap observed during each session was used as an indicator of the level of joint attention achieved by each group. Previous studies conducted with voice enabled computer-mediated communication systems found that participants took on average approximately 2 seconds to focus their attention on an object after it was mentioned by her partner (Richardson & Dale, 2005). In the present study the communication among partners is mediated by a chat and a shared drawing tool. One may expect that reading a chat utterance and then allocating one's attention to the referred object on the drawing board take more than two seconds. However, the gaze recurrence plots with various lag combinations indicated that the highest degree of gaze overlap occurs within a similar time interval. Therefore, those instances in which one subject looks at the same area of the screen that his partner looked at within two seconds were treated as instances of gaze overlap.

The high achievement group exhibited on average 31% gaze overlap, which is followed by the medium and low achievement groups with 24% and 13% gaze overlap respectively. A one-way ANOVA conducted over gaze overlap values indicated that this difference is statistically significant, $F(2,8) = 11.917$, $p < 0.001$, $\eta^2 = 0.341$. Games-Howell post hoc tests found a significant difference between low and medium achievement groups ($MD = -12.32$, $p < 0.05$), as well as between low and high achievement groups ($MD = -20.19$, $p < 0.01$). The difference between medium and high achievement groups was not significant. In other words, higher achieving groups exhibited significantly higher gaze coordination during collaborative problem solving sessions.

Figure 3 shows the overall gaze recurrence plots corresponding to the entire session of two different dyads. The plot on the left of Figure 3 corresponds to the best performing dyad in our sample, whereas the plot on the right corresponds to the worst performing team. The better performing team's recurrence plot differ from the other team's plot in several ways. First, the mean percent recurrence values are much higher for the better team. Second, the better team's recurrence plot significantly deviates from the random baseline, whereas the other team's recurrence plot cannot be distinguished from the baseline. The recurrence plot for the better team has two peaks, one around -1500 msec, whereas the other at 1500 msec. This suggests that both partners equally followed each other's gaze, indicating a high degree of coordinated behavior. Such a pattern is not visible for the other team.

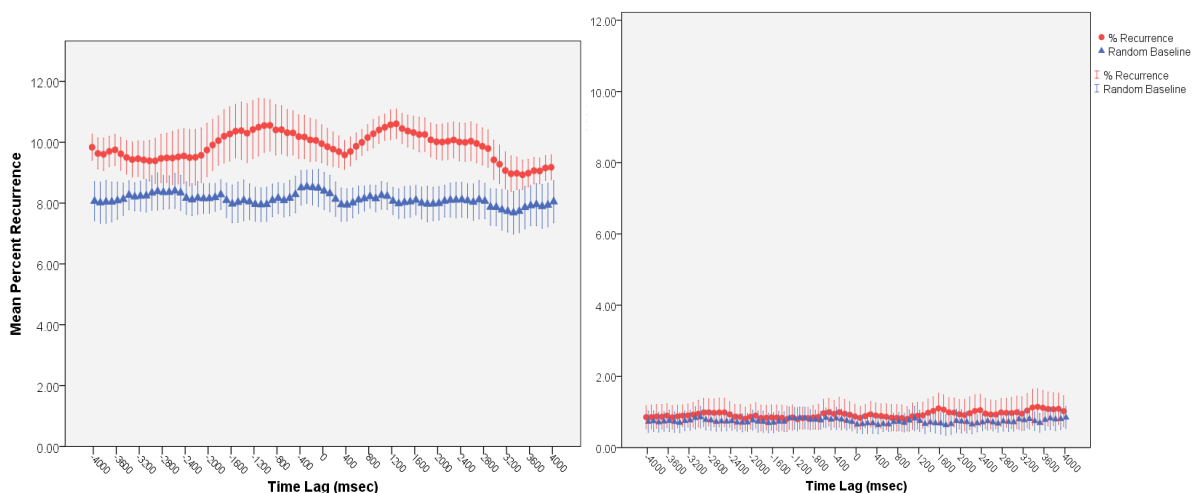


Figure 2. Overall gaze recurrence plots for the highest achieving (left) and the lowest achieving (right) dyads.

In an effort to explore some of the factors underlying these global gaze recurrence patterns, we conducted an interaction analysis of a sample excerpt obtained from both groups' chat sessions. We used the scarf plot in Figure 4 as a guide to index those moments in collaboration with greater degree of gaze overlap. The excerpt shown in table 1 shows the high achieving dyad's work on question 6, which asked them to construct a perfect hexagon by using points, circles and line segments only. In the first three lines, A orients his partner to the new problem. In line 4, after reading the question (as indicated by fixation patterns in the video recording) B solicits A's help for ideas. Next A raises a question which proposes a possible way to characterize a hexagon in terms of 6 equilateral triangles. B's responses in the next two lines suggest that he is not sure about the proposed relationship. Next A takes control of the drawing area and adds a circle. Then, he posts a chat message stating that he does not know how to proceed. He also mentions that he thought about using angles but that did not help him proceed any further.

While A was constructing the circle, B was looking at the unfolding construction on the GeoGebra tab. Since he was following these steps right after they were committed into the system and broadcasted to all

clients, there is a slight lag between A's and B's gaze patterns. Once A seems to be done with his drawing, B makes a suggestion that A should consider drawing an equilateral triangle. As A is drawing the triangle, B looks at places where he expects the next line segment to appear. Such *anticipatory gaze patterns* highlight a stronger degree of common ground or indexical symmetry between the two partners as compared to a situation where one partner merely follows the drawing actions of the other partner.

Table 1. Excerpt from the high achieving dyad

#	User	Time	Chat Message / Drawing Action
1	A	23:43.2	6ya geçelim (<i>Lets move to 6</i>)
2	B	23:46.0	her şey silindi bende (<i>everything is erased</i>)
3	A	23:46.0	ben sildim 6ya bakalım :) (<i>I did that lets take a look at 6:)</i>
4	B	23:57.0	nasıl yaparız? (<i>how can we do it?</i>)
5	A	24:32.6	Şimdi altıgenin içinde 6 tane eşkenar üçgen var di mi (<i>now there are 6 equilateral triangles inside the hexagon, right?</i>)
6	B	24:46.3	6 mı (6?)
7	B	25:16.9	bilmem (<i>I don't know</i>)
8	A	25:19.7	dur bi başlayalım (<i>let's just start</i>)
9	A		constructs a circle with center A and a point B
10	A	26:58.1	bilemedim nasıl yaparız (<i>I don't know how we should do it</i>)
11	A	27:04.5	açı kullancaktım ama yok (<i>I would use angles, but there are none</i>)
12	B	27:10.0	düşünelim (<i>let's think about it</i>)
13	B	27:20.0	yukarda 3 kenar olacak aşağıda da (<i>There should be 3 edges at the top and also at the bottom</i>)
14	B	27:31.9	nasıl buluruz (<i>How can we find it</i>)
15	A		adds point C on circle then draws the diameter connecting B and C (see Figure 3, left)
16	B	29:29.8	eşkenar üçgen oluştursan (<i>how about you draw an equilateral triangle</i>)
17	A		removes diameter, adds another circle sharing the radius with existing circle, draws an equilateral triangle by using the radius (see Figure 3, middle)
18	B	30:46.4	aynısını diğer tarafa da yap (<i>do the same to the other side</i>)
19	A		draws line segments DB and BC, adds a new circle passing through A and B, marks the intersection E (see Figure 3, right)
20	B	31:37.2	aferim (<i>good job</i>)
21	B	31:41.0	tebrkler (<i>congrats</i>)

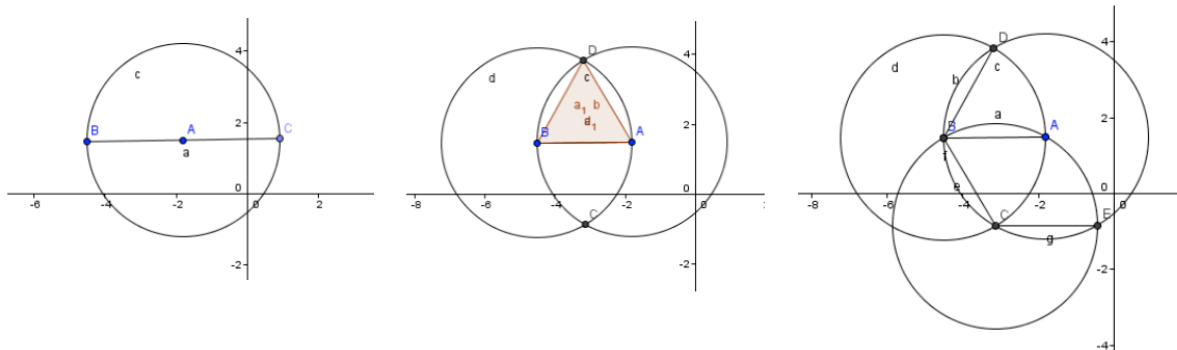


Figure 3. The evolution of A's drawing actions for question 6

Once A completes the triangle (see Figure 3, middle) B suggests him to draw the same triangle on the other side. This is followed by the addition of line segment BC, which is monitored by B. Next, A adds another circle centered at C and sharing a radius with the two other circles. When he defines point E at the intersection of the new circle and the circle on the top right, B recognizes this as the solution. At that moment his eye gaze traverses over the projected new intersections that will become the hexagon once connected by line segments.

Analysis of excerpts from the other team reveals a very different form of interactional organization. As part of their task description, each team was asked to summarize their findings on a separate tab called Summary. Starting with question 2, this team switched to a cooperative mode of operation, where one partner aimed to solve the next question on the GeoGebra tab, while the other partner was busy summarizing the previous solution on the Summary tab. This led to very little transactive interaction among the partners and hence very small degree of gaze overlap. The scarf plot in Figure 4 summarizes the gaze overlap distribution of this team while they were working on one of the questions. The whitespace in the scarf plot also reveals that the first user was busy typing the textboxes in which the team was summarizing their solution for the first question. Occasional gaze overlaps occurred at the chat component when participants exchanged occasional messages.

Overall, the preliminary results of the interaction analysis were consistent with the gaze recurrence analysis. Gaze recurrence analysis is successful for giving a global view of gaze patterns, such as if there is a

significant amount of gaze cross recurrence and if there is an asymmetry among peers (e.g. one partner tended to follow the eye gaze of the other). For instance, the cross recurrence map for the successful dyad was symmetric with respect to the origin, meaning that both partners almost equally followed each other's gaze, which signals that there is a stronger level of coordination among their actions. As far as qualitative aspects of collaboration is concerned, the quality of collaboration increased together with the increase in gaze overlap if both participants actively contributed to the problem solving process by proposing ideas, constructing geometric diagrams and monitoring each other's actions.

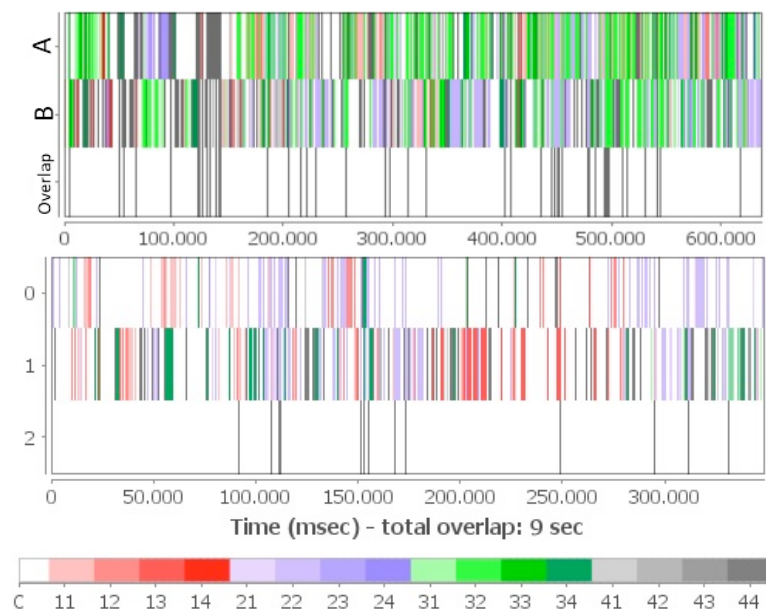


Figure 4. Scarf plot showing the distribution of eye gaze of A and B while they were working on the hexagon problem. The lower graph shows the scarf plot for the lower achieving group.

The degree of gaze overlap was closely related to the achievement of indexical symmetry among partners that renders indexical terms such as “the other side”, “draw an equilateral” etc. meaningful to the participants. In the sequentially unfolding context of their interaction, such terms attain their specificity through the actions of the interlocutors. The presence of indexical symmetry seems to be best evidenced in the case of anticipatory gaze patterns, where one partner fixates on a location where he/she expects the next relevant action to happen. In future work, we are planning to do a more fine-grained analysis of the temporal course of gaze recurrence in an effort to distinguish anticipatory gaze patterns from simple forms of action following. Such features may lead to more fine grained ocular correlates of joint attention and transactive reasoning, which are important factors on the overall success of collaborative learning.

References

- Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice*. New York, NY: Springer.
- Goodwin, C. (1994). Professional Vision. *American Anthropologist*, 96(3), 606-633.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. New York, NY: Oxford University Press.
- Nüssli, M., & Jermann, P. (2012). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task *Proceedings of CSCW'12*, 1125-1134. New York, NY: ACM.
- Nüssli, M. A. (2011). Dual eye-tracking methods for the study of remote collaborative problem-solving. Unpublished PhD thesis, École Polytechnique Fédérale de Lausanne.
- Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29(6), 1045-1060.
- Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination common ground and the coupling of eye movements during dialogue. *Psychological science*, 18(5), 407-413.
- Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *Int Journal of Computer-Supported Collaborative Learning*, 8(4), 375-397.
- Stahl, G. (2013). Topics in Dynamic Geometry for Virtual Math Teams. Retrieved from <http://gerrystahl.net/elibrary/topics/topics.pdf>
- Stahl, G. (Ed.). (2009). *Studying virtual math teams*. New York, NY: Springer.
- van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning & Instruction*, 20(2), 95-99.

Multimedia Educative Curriculum Materials: Designing Digital Supports for Learning to Teach Scientific Argumentation

Suzanna Loper, Lawrence Hall of Science, Univ. of California, Berkeley, CA, sjloper@berkeley.edu

Katherine L. McNeill, Boston College, Chestnut Hill, MA, kmcneill@bc.edu

Raphaela Peck, Lawrence Hall of Science, Univ. of California, Berkeley, CA, ppeck@berkeley.edu

Jeremy Price, Fairmont State University, Fairmont, WV, jeremy.price@fairmontstate.edu

Jacqueline Barber, Lawrence Hall of Science, Univ. of California, Berkeley, CA, jbarber@berkeley.edu

Abstract: We report on work in progress from a research and development project in which we are designing digital supports to help middle school science teachers teach the practice of scientific argumentation. These supports include educative, teacher-facing videos embedded in a digital teacher's guide. In the first phase of the project we developed a framework to guide video development and produced twelve prototype videos. This paper describes the iterative design process for the framework and videos, in which we incorporated evidence from analysis of classroom video, teacher interviews, and teacher focus groups in order to create a design framework aimed to maximize the quality and practicality of the videos.

Introduction

Science education researchers and reform leaders are in alignment with recent U.S. policy documents (National Research Council, 2012) in a call to engage students more deeply and authentically in science practices such as scientific argumentation. Scientific argumentation is a practice that holds great promise for the way in which it draws together conceptual understanding and reasoning and communication skills that are an essential part of the construction of scientific knowledge (Osborne, 2010). Still, implementing the teaching of scientific argumentation poses a great challenge for science teachers (Sampson & Blanchard, 2012). New materials, professional development, and ongoing support are needed (McNeill & Knight, 2013).

Digital tools, including videos and other multimedia and interactive materials, offer potential to provide support for teachers (Lieberman & Mace, 2010). The research and development described in this paper focuses on supporting middle school science teachers in implementing instruction that emphasizes scientific argumentation. This support is provided by multimedia educative curriculum materials (MECMs), including teacher-facing videos that aim to help teachers expand their pedagogical content knowledge (PCK) around scientific argumentation.

Theoretical Framework

Argumentation and the Role of the Teacher

Argumentation has been a focus of research in science education for the past two decades, but has become more prominent in discussions about educational practice through its inclusion in recent standards documents. In our work, we emphasize both dialogic and structural aspects of argumentation. The dialogic aspect of argumentation involves the interactions between multiple individuals as they engage in construction and critique (Ford, 2012). The structural aspect focuses on the products of argumentation, in terms of how claims are supported with evidence and reasoning (McNeill & Knight, 2013).

The teacher plays an essential role in supporting students in argumentation (Zohar, 2008). Teachers' beliefs and PCK for argumentation impact their willingness to incorporate argumentation into instruction and their strategies to support students (Zemba-Saul, 2009). Teachers may lack PCK in how to support students in developing argumentation skills (Zohar, 2008). For example, teachers can struggle with the role of evidence in evaluating competing claims (Sampson & Blanchard, 2012) and they can have difficulty supporting students in classroom discussions in which students critique and question each others' arguments (Berland & Reiser, 2011). In addition, teachers struggle when assessing the strengths and weaknesses of student arguments and in determining strategies to support student learning (McNeill & Knight, 2013). Teachers lack pedagogical strategies, such as how to define an argument or how to provide content-specific examples (Zohar, 2008). In sum, to effectively integrate argumentation into their practice, teachers need greater support.

Digitally Enhanced Educative Curriculum Materials

Educative curricular materials (ECMs) (Davis & Krajcik, 2005) can have a positive effect on teacher knowledge, practice, and student outcomes. By design, ECMs tie new approaches to specific practices, activities, and learning objectives; ECMs thus have great potential for supporting change in teacher practice (Cervetti, Kulikowich, Drummond & Billman, 2013). Multimedia educative curricular materials (MECMs) that link text, representations of student work, and video clips demonstrating pedagogy in action in real classrooms,

hold even greater potential, particularly in an area such as argumentation that incorporates both oral and written modalities. Although there has been little work on the intersection of digital media with educative curriculum, our design of MECMs draws on previous work focused on multimedia representations of practice. *Multimedia representations of practice* can provide learning opportunities for teachers grounded in real life situations that utilize a rich and multi-layered approach to classroom teaching (van den Berg, Wallace & Pedretti, 2008). In particular, video cases specific to a teacher's curriculum can help support the development of pedagogical content knowledge (Roth, et. al, 2011). Multimedia representations of teaching can illustrate the intricacies and subtleties of effective teaching practices. This project builds on research on the affordances of digital curriculum materials, including the possibility for multimedia representations of teaching practice, to help teachers develop the beliefs and PCK that can support their enactment of argumentation instruction.

Research and Development Process

The work described here is part of a five-year project in which we are designing and researching MECMs to support teacher learning of argumentation. We focus here on our early design phase. Our question in this phase was, *How can multimedia educative curricular materials (MECMs) be designed to positively impact teachers' beliefs and pedagogical content knowledge about argumentation?* In this design work, our two main considerations were practicality and quality (Doyle & Ponder, 1978); in other words, we sought a design framework that would support the development of videos likely to be accessed by teachers and to have an impact.

In order to pursue these goals of quality and practicality, our development followed an iterative process (see Figure 1), with input at multiple timepoints from a number of sources. Through this process we created a MECM Design Framework to guide the development of prototype videos. The framework has three components: (1) Learning Goals (our goals for teacher learning); (2) Teacher Needs (types of needs that videos could address); and (2) Video Specifications (guidelines for length, setting and participants, and other elements to be included in the videos). The Learning Goals were derived from review of literature and input of expert advisors, described in more detail in McNeill, Katsh-Singer, Gonzalez-Howard, Price & Loper (2013). The Teacher Needs and Video Specifications were developed based on two main sources of evidence; (1) analysis of videotaped lessons and teacher interviews; and (2) teacher focus groups. These evidence sources are described below.

Evidence Sources

Videotaped Lessons and Teacher Interviews

In 2011-12, we collected data from ten teachers implementing a field trial version of an earth science curriculum with a specific focus on argumentation. For each teacher we collected videotapes of two argumentation lessons and conducted follow-up interviews. Interviews focused on the teachers' instructional practices in relation to the original curriculum and their rationales for instructional decisions specifically around argumentation. Interviews were audio recorded and transcribed and a coding scheme was developed from both our theoretical framework and an iterative analysis of the data (Miles & Huberman, 1994). The findings from this study informed the development of the Learning Goals and Teacher Needs and are reported in more detail in McNeill, Gonzalez-Howard, Katsh-Singer, Price & Loper (2013).

Teacher Focus Groups

In 2012-13, we conducted a series of teacher focus group sessions to inform our development of the MECM Design Framework and the prototype videos themselves. Five teachers were selected from a list of teacher contacts in the local area. These five teachers represented a range of school settings (suburban and urban;

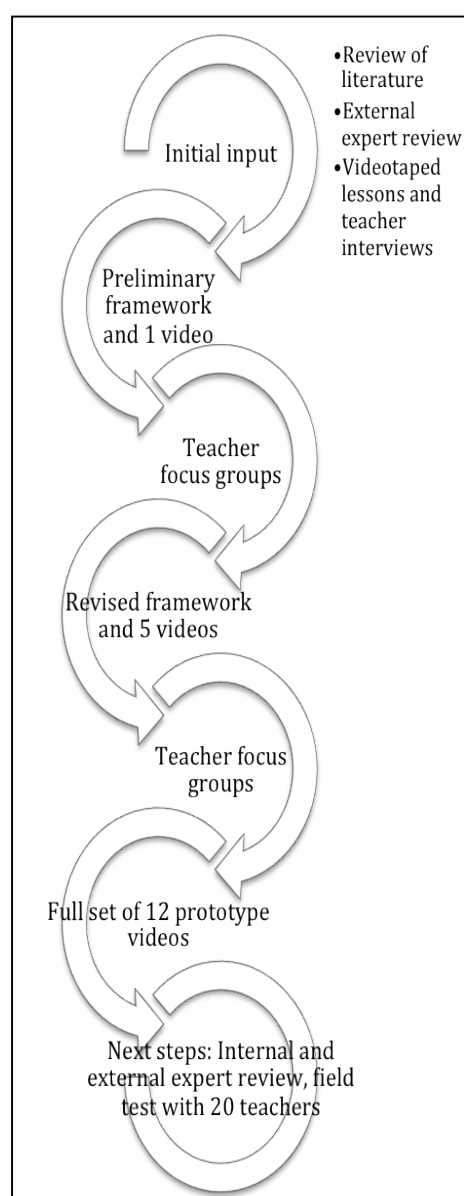


Figure 0. Iterative development of framework and prototype videos

parochial, charter, and public), years of experience teaching in the classroom (2-30 years), and prior experience participating in pilot and field trial studies with our group.

Four focus group sessions were conducted over a period of seven months. First, each teacher participated in an individual interview at their school site. The remaining sessions were conducted as group sessions, with two sessions occurring in face-to-face meetings and one session conducted via web-based conference and the use of Google Docs. Data collected included audio recordings, field notes, research memos and participant artifacts. The data was coded in a recursive and comparative manner (Charmaz, 1994) into categories of findings. The findings from these focus groups informed both the development of the Teacher Needs and the Video Specifications.

Results: Iterative Development of MECM Design Framework

Findings from the evidence sources were used in the iterative development of the MECM Design Framework, which includes three elements: 1) Learning goals, 2) Teacher needs, and 3) Video specifications.

Learning Goals

We identified three critical learning goals for teachers based on our evidence sources (McNeill, Katsh-Singer et. al, 2013; McNeill, Gonzalez-Howard, et. al, 2013) that focus on both the structural and dialogic aspects of argumentation. The phrasing of the goals was refined, based on feedback from the teacher focus groups, to make them clear and appealing to teachers. For example, while these are goals for teacher learning, focus group teachers recommended that the goals nonetheless be phrased to emphasize student achievement. Focus group participants felt that teachers are motivated by a focus on supporting their students, as opposed to a focus on themselves as learners.

Table 1: Learning Goals

Teachers will learn how to...

1. Improve students' use of high-quality evidence.
2. Help students to strengthen their arguments by articulating reasoning (connections between evidence and claims)
3. Create conditions that support student-driven argumentation.

Teacher Needs

Following a design based research approach, three Teacher Needs were developed as an “embodied conjecture” (Sandoval, 2004) drawing on literature review, expert advisor input, and our other evidence sources. For example, in the analysis of videotaped lessons and teacher interviews, we found that different teachers enacted the same lesson plan in very different ways, and yet all felt that they were implementing the lessons faithfully (McNeill, Gonzalez-Howard, et. al, 2013). This suggested a need for a type of video that provides ‘images of practice’ to help teachers better envision the intent represented by the lesson plans. Two other needs were identified: ‘meta-level’ videos that provide teachers with foundational information about what argumentation is and why it is important; and ‘strategy’ videos that provide teachers with concrete actions they can use in the classroom. Based on input from the teacher focus groups, we developed teacher-facing labels for three types of videos addressing each of these needs. Table 2 lists the three categories of videos, with the associated Teacher Needs. The teacher-facing labels for the videos are in quotations.

Table 2: Video Categories and Associated Teacher Needs

1. **Meta-Level Videos (“Building Blocks”)**: Teachers need to understand the elements of scientific argumentation (for example, the components of an argument, or what counts as scientific evidence) and the rationale for incorporating argumentation into science instruction.
2. **Images of Practice (“In the Classroom”)**: Teachers need an image of what an instructional practice looks like in the classroom, and increased confidence that it can be enacted with real students.
3. **Strategy Videos (“Strategies”)**: Teachers need concrete tools, activities and strategies that they can take into the classroom and try right away.

Video Specifications

The teacher focus groups were our primary evidence source for constructing a set of Video Specifications, designed to help us maximize the practicality of the videos, in particular the likelihood that teachers would actually use them (Doyle & Ponder, 1978). We sought input from teachers about their preferences for the elements included in the videos and how they believed they would use the videos. Themes that emerged from the teacher focus groups fell into three categories: 1) Access Options and Potential Uses, 2) On-Screen

Participants and Settings, and 3) Video Details and Supporting Media. These findings from the focus groups resulted in the development of a list of Video Specifications, excerpted in Table 3.

Table 3: Video Specifications

<p><i>Access Options and Potential Uses</i></p> <ul style="list-style-type: none"> • Flexible access system, including both lesson-specific links and top-level menu structures <p><i>On-Screen Participants and Settings</i></p> <ul style="list-style-type: none"> • Diverse classroom settings; avoid classrooms that appear unusually ‘high-resourced’ • Include extended student dialogue • Teacher voices are more desirable than scientists or curriculum developers <p><i>Video Details and Supporting Media</i></p> <ul style="list-style-type: none"> • Video length 3-6 minutes • Include reflection questions for teacher • Include annotation or call-outs to help scaffold teacher interpretation of student footage • Include explicit connections to state and national standards

Development of Prototype Videos

The MECM Design Framework was used to guide the creation of twelve prototype videos, listed in Table 4.

Table 4: Titles of 12 prototype videos

Learning Goal	Building Blocks	In the Classroom	Strategies
1: High quality evidence	• <i>What Is Evidence?</i>	• <i>What Does Evidence Look Like?</i>	• <i>Strategies for Working with Evidence from Text</i>
2: Reasoning	• <i>Why Is Reasoning Important?</i>	• <i>What Does Reasoning Look Like?</i>	• <i>Strategies for Supporting Reasoning with Questioning</i>
3: Student-Driven Argumentation	• <i>Why Argumentation?</i> • <i>What Is a Science Seminar?</i>	• <i>What Does a Culture of Argumentation Look Like?</i> • <i>What Does Working with Multiple Claims Look Like?</i>	• <i>Strategies for Adapting Activities for Argumentation</i> • <i>Strategies for Conducting Science Seminars</i>

The videos contain a mix of classroom and interview footage, text, and voiceovers. Video scripts were developed in an iterative process involving preliminary script development, collection of footage, culling of footage, and revision of scripts based on footage availability. Four filming days were conducted, with two days focused on classroom footage and two days focused on interviews with teachers, scientists and curriculum developers. For each classroom filming day, we spent 1-2 days prior to filming working with the students to prepare them for the activities that would be filmed. Working with a video production company specialized in educational media, we created draft videos that went through several rounds of input and revision.



Figure 1: Frames from *Why Is Reasoning Important?* and *What Does Working with Multiple Claims Look Like?*

Figure 2 shows sample frames from two of the videos. The frame on the left-hand side is from a “Building Block” video. In this case, the video is annotated to help the viewer link the students’ words to the components of an argument. The frame on the right-hand side is from an “In the Classroom” video focused on student-driven argumentation. Here the text at the bottom of the screen highlights this characteristic of argumentation as the viewer observes students engaging in these important interactions.

Discussion

Previous technology innovations have not become widespread in K-12 schools, in part because of their lack of focus on systemic issues such as usability, scalability and sustainability (Fishman, Marx, Blumenfeld, Krajcik & Soloway, 2004). In our work on the development of MECMs to support teachers' learning of argumentation, we take this issue seriously. Our current MECM design framework suggests the importance of focusing on both *quality* and *practicality*: creating materials that could have the desired impact, and developing them in a way that teachers would see as useful and appealing. For example, in the creation of the learning goals, we utilized the current research literature (e.g. McNeill & Knight, 2013; Ford, 2012) and our evidence sources to identify high *quality* learning goals addressing both structural and dialogic aspects of argumentation. Then we modified the language of those learning goals to target supporting students, instead of teachers, because in terms of *practicality* the focus group teachers found these more appealing. We feel that both quality and practicality are essential characteristics of MECMs to support greater scalability. The prototype videos are currently being revised based on further input, and their impact will be investigated in a randomized experimental study.

References

- Berland, L. K. & Reiser, B. J. (2011). Classroom communities' adaptations of the practice of scientific argumentation. *Science Education*, 95, 191-216.
- Cervetti, G. N., Kulikowich, J., Drummond, M., & Billman, A. (2013). Influence of educative curriculum materials on teachers' learning and teaching. Paper presented at the annual meeting of the American Educational Research Association. San Francisco, CA.
- Charmaz, K. (1994). Grounded Theory: Objectivist and Constructivist Methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 509–535). Thousand Oaks, CA: Sage.
- Doyle, W., & Ponder, G. (1978). The practicality ethic in teacher decision-making. *Interchange*, 8(3), 1-12.
- Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, 34(3), 3-14.
- Fishman, B., Marx, R. W., Blumenfeld, P., Krajcik, J., & Soloway, E. (2004). Creating a framework for research on systemic technology innovations. *Journal of the Learning Sciences*, 13(1), 43-76.
- Ford, M. J. (2012). A dialogic account of sense-making in scientific argumentation and reasoning. *Cognition and Instruction*, 30(3), 207-245.
- Lieberman, A. & Mace, D. P. (2010). Making practice public: Teacher learning in the 21st century. *Journal of Teacher Education*, 61(1-2), 77-88.
- McNeill, K. L., & Knight, A. M. (2013). Teachers' pedagogical content knowledge of scientific argumentation: The impact of professional development on k-12 teachers. *Science Education*, 97, 937-972.
- McNeill, K. L., Katsh-Singer, R., Gonzalez-Howard, M., Price, J. F., & Loper, S. (2013). *Factors that impact teachers' argumentation instruction in their classroom*. Paper presented at the annual meeting of the American Educational Research Association. San Francisco, CA.
- McNeill, K. L., Gonzalez-Howard, M., Katsh-Singer, R., Price, J. F. & Loper, S. (2013). *Teachers' beliefs and practices around argumentation during a curriculum enactment*. Paper presented at the annual meeting of the National Association for Research in Science Teaching. Rio Grande, Puerto Rico.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.
- National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts and core ideas*. Washington, DC: The National Academies Press.
- Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, 328, 463-466.
- Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Videobased lesson analysis: Effective science PD for teacher and student learning. *Journal of research in science teaching*, 48(2), 117-148.
- Sampson, V. & Blanchard, M. (2012). Science teachers and scientific argumentation: Trends in views and practice. *Journal of Research in Science Teaching*, 49, 1122-1148.
- Sandoval, W. A. (2004). Developing learning theory by refining conjectures embodied in educational designs. *Educational Psychologist*, 39(4), 213–223.
- van den Berg, E., Wallace, J., & Pedretti, E. (2008). Multimedia cases, teacher education and teacher learning. In Voogt, J. & Knezek, G. (Eds.), *International handbook of information technology in primary and secondary education*. (pp. 475-487). New York, NY: Springer, LLC.
- Zemba-Saul, C. (2009). Learning to teach elementary school science as argument. *Science Education*, 93, 687-719.
- Zohar, A. (2008). Science teacher education and professional development in argumentation. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research*. (pp. 245-268), Dordrecht: Springer.

How Good Is This Evidence?

Students' Epistemic Competence in Evidence Evaluation

Clark A. Chinn, Leah C.-C. Hung, Randi M. Zimmerman, Ravit Golan Duncan
Graduate School of Education, Rutgers University, 10 Seminary Place, New Brunswick, NJ 08901
clark.chinn@gse.rutgers.edu, randi.zimmerman@gse.rutgers.edu, leahcchung@gmail.com,
ravit.duncan@gse.rutgers.edu

Abstract: Inquiry environments in science classes have increasingly incorporated more features of authentic scientific practice. However, relatively few environments have incorporated a critical feature of scientific practice: evaluation of evidence quality. This paper reports results from two studies in middle-school life science classes that investigate seventh graders' competence in evaluating evidence. Overall, we found evidence that students have strong evaluative capabilities that can be built upon in instruction.

Introduction: Integrating Evidence Evaluation into Inquiry Environments

Chinn and Malhotra (2002) argued that, for inquiry-oriented instruction to be effective in improving students' reasoning, it is important that these environments be as authentic as possible. Oversimplified inquiry environments can promote oversimplified student epistemologies that are counterproductive in the complexity of real-life settings. For example, when students learn to control variables in oversimplified laboratory tasks, they may come to think of scientific inquiry as simple and algorithmic, just a matter of checking off a set of provided variables, failing to appreciate the difficulties of deciding what variables to control in real-life inquiry, devising how to control them, and so on. Chinn and Malhotra (2002) touted more complex inquiry environments such as the BGuILE environment (Reiser et al., 2001) that afforded students the opportunities to reason in more realistic ways about more complex problems of the sort that they will encounter out of school.

However, there is one important aspect of authentic scientific inquiry that remains relatively infrequent in inquiry environment designs, especially those developed for younger students such as middle-school students. Specifically, in authentic science, scientists not only use evidence to evaluate theories; they also evaluate the evidence itself. They consider how strongly to weight the evidence based on their evaluations of evidence quality (e.g., whether a study has an adequate sample size, properly controlled variables, valid and reliable measures, and so on) (Staley, 2004). They appropriately discount low-quality evidence. But most inquiry environments have not featured low-quality evidence with the intent of helping students learn how to reason about such evidence. There are plausible arguments for holding off on introducing extensive discussions of methodological features of studies. Students may have difficulty learning to identify methodological flaws; if this is so, then there is a danger that students will draw incorrect conclusions from evidence. For example, a student presented with methodologically flawed evidence against global warming might fail to notice its flaws and use the evidence as a centerpiece of an argument that global climate is not warming. Given worries such as these, there is a need for research that investigates students' abilities to evaluate evidence quality, as a first step in understanding how methodological diversity in evidence can be introduced to inquiry environments.

There is relatively limited research on how students evaluate the methodological processes used by scientists (Allchin, 2011). Some research has examined students' understanding of error in simple scientific experiments (Masnick & Klahr, 2003), but there is much less work examining evaluation of more complex, realistic studies. Those educational studies that provide students with details of research reports often simply code whether participants refer to "methods," and do not provide more detailed analyses of what exactly participants have to say about such methods. Our research attempts to help us begin to build an understanding of the strengths and weaknesses of students' reasoning about evidence quality.

In this paper, we report the results of two exploratory investigations of students' skill in evaluating evidence quality in science inquiry environments. This is part of a larger project investigating students' ability to learn to coordinate theories and evidence while simultaneously evaluating evidence quality. In the spirit of the "Reports & Reflection" category of ICLS papers, we report some critical initial findings that can provide grounding for future design work. We view this work as an essential first step in developing inquiry environments that are effective at promoting the critical epistemic practice of evaluating evidence quality.

Context of Research: PRACCIS

The context of our research is the PRACCIS (Promoting Reasoning and Conceptual Change in Science) project (Chinn, Duncan, Dianovsky, & Rinehart, 2013). In this model-based inquiry project, students use evidence to develop and revise models, as well as to decide between two or more models. We use a variety of scaffolds to promote thinking about theories and evidence, including the MEL (Model-Evidence Link) diagram (Rinehart, Duncan, & Chinn, in press), in which students use arrows of different types to indicate whether evidence

supports, strongly supports, contradicts, strongly contradicts, or is irrelevant to one or more models. Like most other inquiry environments, our early implementations of PRACCIS used only “good” evidence that students were expected to take seriously as they developed models. Previous to the studies reported in this paper, we made no attempt to design “bad” or low-quality evidence into the curriculum so that students could learn to reason about such evidence. The present paper outlines our exploratory work to authentically incorporate “bad” evidence into our curricula so that students have an opportunity to learn to evaluate evidence.

Study 1: Initial Explorations Using an Inquiry Unit with Low-Quality Evidence

Our initial exploration developing inquiry units with low-quality evidence occurred late in a school year after students had engaged in model-based inquiry for over 7 months. We developed a food web unit in which seventh graders used evidence to choose between two alternative models of the food web of an arctic habitat involving seals, foxes, polar bears, salmon, and plankton. We employed evidence of different quality, ranging from low-quality to high-quality evidence, to encourage student reflection and discussion about evidence quality. Two teachers implemented this unit in their classes.

In one low-quality piece of evidence, a hotel owner interviewed on a simulated Alaskan radio show reported that 350 guests at her hotel reported seeing 11 arctic foxes this year and no salmon in the ocean, but the guests a year ago reported seeing 16 arctic foxes and two salmon. This evidence involved haphazard, unrepresentative samples both of guests and observations. A second piece of evidence showed four photos snapped by a fisherman, each showing one or two polar bears eating a seal. Other evidence involved more systematic observations and population counts of different species.

We wondered whether students would detect the poor quality of some of the evidence without extensive coaching from the teacher, or indeed without being told in any way that any evidence was problematic. The following transcript excerpts present one class’s first discussion of these issues:

- Teacher: So what does this evidence tell us about how many foxes and salmons were in the area this year and last year? Blair?
- Blair: Uh, that there are more arctic foxes living uh in the area near the hotel but there was more the year before because it uh went down the next year. And the salmon, there aren't a lot in the ocean or there aren't a lot around the ocean where people are looking.
- Teacher: [to the class] You need to listen to him, and you need to respond to him. I'm not saying anything.
- Regina: Um, well, um, what I noticed in this evidence is where there are less arctic foxes, um, there weren't any, um, no one saw any salmon.
- Teacher: No, no, just look at this information right here and say, what does the evidence show.
- Regina: Um well, the less arctic foxes there were, there weren't any salmon and the less arctic foxes there were a few more. And um, I conclude from this evidence that um, if there is less arctic foxes there is, um, more ring seals. And where are there more ring seals,

To this point, there was no hint that any students recognized that the evidence might be of such low quality that it should be ignored or severely discounted. Instead, these two students (Blair and Regina) seem to have assumed that the evidence was good evidence, and that they needed to account for this evidence when choosing which food web was better. However, as the discussion continued (below), two students--without any hints or coaching by the teacher--commented on problems with the evidence.

- Teacher: Okay. Anybody want to say anything to her? Yes, go on.
- Andy: Well it's not really easy to see salmon in the ocean either, because they're underwater and stuff. So, you'd probably be pretty lucky to see any so just 'cause somebody saw two last year and there are more arctic fox don't mean that because there are more fox. There would be more salmon. 'Cause doesn't arctic fox eat salmon anyway?
- Teacher: Did anyone have some thoughts about what she was saying?
- Candace: Um, in the ocean, and they say that they saw two fish or whatever number they saw. If there's two, two fish in the water, there are more fish in the water because like how is there just going to be two fish? Like, that just makes no sense.
- Teacher: Okay so what she was saying. So they start to give you up and down. Right? What was going up and what was going down. But does this tell us what was really happening with all the salmon?
- Candace: No.
- Teacher: They're reading beyond that. They're now, Andy and Candace are making, critiquing; they are critiquing evidence rather than trying to draw a conclusion. Yes? And what else?

In this excerpt, Andy spontaneously questioned whether hotel visitors can reliably spot salmon in the ocean

(thus noticing the possibility of error in observations). Candace added that the number of salmon that can be seen tells you little about the salmon that cannot be seen; in effect, Candace has detected a problem with inferring population sizes from a limited sample of observations. Only at this point did the teacher explicitly articulate that Andy and Candace are critiquing the evidence rather than trying to draw conclusions from it.

A few turns later, James argued that the “fox population is going down,” and he offered this reason:

James: Because uh, last, because last March, there were only 11 foxes sighted. But the, but the year before there were a total of 16 sighted.

But other students responded to James by continuing their critique of the evidence:

Teacher: Sighted, okay. Yes, um, Erica?

Erica: That really doesn't tell you anything though, because it's just guesses. Like walking around, there, there's not like actual people going out trying to find these things. And it might not even be an arctic fox. They might just think it's an arctic fox.

Brianna: Also maybe like the arctic fox is kind of like, um, hibernated to another part of the forest, and they're not near the hotel anymore.

Teacher: You had your hand up?

Natalie: I was going to say that also, yeah, people could like not be looking for them or mistake them for an arctic fox which cannot be an arctic fox because they may not really know.

Regina: Um, well I don't think arctic foxes um will um intentionally um show themselves to people. So if people only saw 11, maybe there are more. They just saw it, there doesn't have to be 11. It could be um anywhere else. And....

Teacher: Keep going.

Regina: So um, if you just saw it, that doesn't mean there are only 11 there. Because I don't think arctic foxes will just show themselves.

Thus, students argued strongly against James's argument that the number of foxes sighted provides information about the actual fox population in this habitat. In addition to Andy's and Candace's earlier reasons, the new reasons given by the students include: (1) walking around without the intent to observe carefully is unreliable (Erica and Natalie); (2) observers might be mistaken in their observations due to poor ability to classify species seen (Erica and Natalie); (3) hotel guests might look in the wrong places, with foxes being elsewhere (Brianna); and (4) human observers underestimate the number of foxes because foxes hide from people (Regina). (Later in the discussion, a comment by James indicated that he accepted these criticisms as valid.)

Collectively, the students accurately (and in our estimation, impressively) identified a broad range of flaws with this evidence. The teacher did little to guide this discussion. Later, as the discussion continued, and students continued to give reasons to doubt the evidence, the teacher asked if the hotel guests' observations gave an accurate population count, and many students answered in chorus, “No.” Thus, through the discussion, many students appear to have concluded that this evidence was poor quality and should be discounted.

This discussion and similar observed discussions from Study 1 provided encouraging preliminary evidence that seventh graders can reason effectively about authentic evidence sets that include low-quality as well as high-quality evidence. We explored students' capabilities in more detail in a second study.

Study 2: Seventh Graders' Evaluations of Evidence Quality

Encouraged by our experience with the food web unit, we developed a set of model-based inquiry units that systematically encouraged students to reflect on the quality of evidence. As students developed, revised, and evaluated models based on evidence, they also evaluated the evidence itself. In the analyses reported here, we focus on students' evidence evaluations in a six-week unit on organelles.

Method

We carried out a yearlong study involving over 500 seventh grades in 20 New Jersey science classes taught by 5 teachers implementing a life-science model-based inquiry curriculum that we designed. Two classes taught by each teacher were video recorded throughout this time. Our analyses draw on written data provided by 38 students in two of these teachers' video-recorded classes.

The students engaged in a variety of inquiry activities designed by the research team in extensive, iterative collaboration with participating teachers. The units engaged students in reasoning about evidence of widely varying quality. The data we discuss in this paper are all from an early unit in the curriculum, the organelles unit, in which students used evidence to develop a model of what chloroplasts do, used evidence to determine which of two models of mitochondria function is better (mitochondria produce “energy” versus mitochondria produce “movement”) and used evidence to decide which of two models of nuclear function is better. In each unit, students were encouraged to reflect on and evaluate evidence quality. For example:

Chloroplast observations. After observing chloroplasts through a microscope, students considered how

well their observational evidence could be taken to support any conclusions that they wanted to draw.

Hamster evidence. In the mitochondria lesson, students read a multimedia blog by a hamster owner (a college student) in which he performed microscopic observations of his dead hamster to determine whether the hamster's chronic low energy was caused by having too few mitochondria. The blog related several methodological problems experienced by the hamster owner as he carried out his investigation in a university lab.

Flagella evidence. In the mitochondria lesson, students learned about another study on another student's blog about observations of single-cell organisms with flagella. Several daily blog entries by the student provided a lively description of the procedures and findings of the study (more mitochondria were clustered near the flagellum than in other parts of the cell), but the results from only two cells were shown.

Cat evidence. The nucleus unit included a computer-based animation showing a study in which nuclei from glowing jelly fish were transplanted into cats, which produced the same protein that made the jellyfish glow, and yielded glowing cats. This was a well conducted study, albeit with a small sample size.

Students used these and other evidence to make their model-based judgments in each lesson. At several points, they also answered explicit questions about the quality of evidence. For example, students responded to this prompt about the cat evidence: "How good or bad is Evidence #2 [the cat evidence]? Write your reasons for your answer. Write to someone who might disagree with you." Our analyses focused on the students' responses to questions of this sort about the hamster evidence, the flagella evidence, and the cat evidence.

Research Questions

To support improved design of inquiry environments, we sought to understand the strengths and weaknesses of students' evaluation of evidence. The main questions were: (1) What is the range of students' responses when they are asked to evaluate evidence quality? (2) More specifically, how sensitive are students to methodological strengths and weaknesses in evidence? Given that scientists' evaluations of evidence quality typically focus strongly on how reliable the methodological processes were in the evidence, we wondered particularly whether students would pick up on reliable and unreliable methodological processes described in the evidence.

Coding

Two coders coded all of the data along two primary dimensions, discussing to reach agreement on all items. The first dimension captured the different categories of response to the data. The resulting categories from this dimension are presented in Table 1. The second dimension addressed how elaborated the responses were, on a 1 to 3 scale. Responses at Level 1 briefly mentioned features of the study but did not explain why those features made the evidence good or bad. Level 2 responses provided somewhat more detail than Level 1 responses. Level 3 responses pointed out specific features of the evidence and explained how these features supported an evaluation of the evidence as good or bad.

Table 1: Students' responses to prompts to evaluate how good the study is

Category	Example	% of students who gave response		
		Hamster evidence	Flagella evidence	Cats evidence
Student comments on study results.	It is OK evidence because it does sound like a low number of mitochondria.	14%	14%	11%
Student comments on conclusions to drawn from the study.	I think this was bad evidence because it did not relate to the model or talk about the nucleus. This is why it's bad evidence.	17%	14%	11%
Student identifies a feature of the evidence presentation	It has examples, pictures, and labels.	11%	64%	36%
Student refers to an aspect of the methodological processes				
Observation/measurement issues	... he actually looked through a microscope to get to his conclusion.	56%	47%	0%
Inference issues	... the picture was bad, and there are a lot of inferences based on her observation on the hamster.	11%	8%	6%
Design issues (e.g., controls)	I think evidence #2 is good evidence because the experiment was controlled.	14%	11%	17%
Execution and procedural issues	1. They [went] step by step. I really liked it. 2. ...the images were unclear and fuzzy.	3%	0%	0%
Sample size issues	... their sample size wasn't so great. They only used one cat to put jellyfish gene in.	8%	0%	29%
Researcher credibility	... he said he hoped he didn't do anything wrong so he could have messed up.	14%	3%	11%
Corroboration from others	Furthermore he observes the flagellum and approves of it with other classmates to have all of the info covered.	0%	17%	0%

Results

A brief summary of several key results is as follows:

First, students' overall evaluations of all three studies roughly reflected the quality of the evidence as intended by the design team. The hamster evidence (intended to be the weakest evidence) was evaluated positively overall by 31% of the students, whereas the stronger flagella evidence and cat evidence were rated positively by 67% and 62%, respectively.

Second, in their responses, students frequently evaluated studies on the basis of the studies' conclusions or the results, rather than focusing on the methodological processes of the study. For example, one student wrote: "It is OK evidence because it does sound like a low number of mitochondria," focusing on the results of the study rather than the methods. These are not appropriate responses, in our view. Scientists, too, may view a study as good in part because it has a clear-cut, compelling result that informs theory (Staley, 2004).

Third, most students (over 75%) gave fairly detailed, level 3 responses to the questions. Thus, students were able not only to give responses; they gave reasoned responses providing justifications for their ideas.

Fourth, students collectively provided a broad range of comments focused on strengths and weaknesses of the methodological processes used. Table 1 displays the responses by students in selected categories for each piece of evidence. Collectively, students successfully identified most weaknesses and strengths of each study.

Fifth, students' responses exhibited some sensitivity to the particular strengths and weaknesses of the three pieces of evidence. Although space precludes a detailed analysis in support of this point, the different percentages of responses in each category across the three pieces of evidence supports the claim that students were sensitive to particular feature of each piece of evidence.

Sixth, a large majority of students' responses were normatively appropriate. That is, a majority of students' identified methodological strengths can indeed be regarded as strengths, and their identified weaknesses can be regarded as weaknesses.

Discussion

The two studies provide evidence that seventh graders are collectively aware of important criteria for evaluating the quality of evidence. Study 1 shows that it is possible for students in class discussions to collaboratively develop accurate evaluations of evidence quality, pooling the insights of different students. Study 2 shows that students collectively have a wide range of productive insights about evidence evaluation that can be so pooled. In other work (Authors, 2010, 2011b), we have found that when students collectively identify a broad range of important criteria for evaluating models, they develop class norms that successfully combine their productive individual understanding. Thus, these results provide a warrant for designing inquiry environments that encourage seventh graders to share their ideas about evaluating evidence so as to promote greater skill.

References

- Allchin, D. (2011). Evaluating knowledge of the nature of (whole) science. *Science Education*, 95, 1-25.
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic reasoning in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86, 175-218.
- Chinn, C. A., Duncan, R. G., Dianovsky, M., & Rinehart, R. (2013). Promoting conceptual change through inquiry. In S. Vosniadou (Ed.) *International Handbook of Conceptual Change* (2nd ed.) (pp. 539-559). New York: Routledge.
- Masnick, A., & Klahr, D. (2003). Error matters: An initial exploration of elementary school children's understanding of experimental error. *Journal of Cognition and Development*, 4, 67-98.
- Reiser, B. J., Smith, B. K., Tabak, I., Steinmuller, F., Sandoval, W. A., & 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.
- Rinehart, R. W., Duncan, R. G., & Chinn, C. A. (in press). A scaffolding suite to support evidence-based modeling and argumentation. *Science Scope*.
- Staley, K. W. (2004). *The evidence for the top quark: Objectivity and bias in collaborative experimentation*. Cambridge: Cambridge University Press.

Acknowledgements

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.

Mathematical Tasks as Boundary Objects in Design-Based Implementation Research

Raymond Johnson, Samuel Severance, Heather Leary and Susan Miller

University of Colorado Boulder, Boulder, Colorado, USA

raymond.johnson@colorado.edu, samuel.severance@colorado.edu, heather.leary@colorado.edu,
susan.miller@colorado.edu

Abstract: This paper describes a collaboration involving school district curriculum supervisors, mathematics teachers, university researchers, and web engineers engaged in design-based implementation research of the adaptation of an Algebra 1 curriculum to meet the demands of the Common Core State Standards for Mathematics. In this effort, mathematical tasks operated as a boundary object, acting to organize cooperative work despite a lack of consensus regarding their meaning or purpose. Exposing and understanding this lack of consensus during joint work at community boundaries provided opportunities for learning in the form of changed practice and activity. Evidence for learning through coordination is presented, such as the communication of curricular vision and the routinization of a task rating process. Where lack of consensus led to potential conflict, as in the case of task adaptation, evidence for learning in the form of new practices was less evident.

Major Issue

The adoption of the Common Core State Standards for Mathematics (CCSSM) has created a need for many teachers and school districts to consider changing their curricular materials. For the 44 U.S. States that have adopted these standards, they generally represent a more focused and demanding target than prior standards (Porter, McMaken, Hwang, & Yang, 2011). While some schools will purchase new curriculum materials that better align to the CCSSM, many others will attempt to modify and improve existing materials. This kind of adaptation can require considerable effort and expertise, and implementation strategies for the CCSSM are not yet well-developed (Cobb & Jackson, 2011). Collaboration with researchers can be a productive option for not only developing quality curriculum, but for better understanding the processes by which standards-based curriculum is designed and implemented.

A useful unit of mathematics curriculum is the *task*, which is sometimes defined as “a classroom activity, the purpose of which is to focus students’ attention on a particular mathematical idea” (Stein, Grover, & Henningsen, 1996, p. 460). As intended by the CCSSM, high-quality mathematical tasks are likely to be inquiry-based activities that engage learners in multiple Standards for Mathematical Practice (NGA Center/CCSSO, 2010). While adapting curricula to include such tasks may be a goal shared by teachers, district curriculum supervisors, researchers, and other reform agents, the meaning and significance of a mathematical task may vary depending on one’s position and perspective. In this way, mathematical tasks are a potential *boundary object* (Star & Griesemer, 1989; Star, 2010). In this paper, we ask: How do mathematical tasks operate as boundary objects, and what do mechanisms for learning related to boundary objects and boundary crossing imply for collaborative efforts to adapt and implement new curricula?

Contextualization and Significance

Our research is an example of design-based implementation research (DBIR; Penuel, Fishman, Cheng, & Sabelli, 2011), an approach that expands classroom-based design research to consider perspectives of other stakeholders in an educational system. DBIR leverages iterative and collaborative development of processes and products to enhance a system’s capacity to sustain and scale change related to persistent problems of practice.

Participants in this DBIR project represented four communities: university researchers, curriculum supervisors from an urban school district, high school algebra teachers from the district, and an engineering team coordinating with the researchers to develop a web-based catalog of curricular resources. The district supervisors selected teachers with the goal of representing varying levels of teaching experience and expertise with curriculum development. Approximately ten teachers participated at any one time on a Teacher Design Team (TDT), with some leaving and others being added as the project changed phases or as individual circumstances demanded. This project is significant not only in its purpose to support curricular change in response to the CCSSM standards, but in its DBIR approach and researcher access to both the district’s teachers and their supervisors. The focus of this paper is the initial phase of the project, a period from July 2012 through the first TDT meeting in December of 2012, a full-day workshop involving all four communities. During this phase there were weekly meetings of the research team as well as weekly phone conferences between the researchers and the district supervisors. One member of the engineering team attended these weekly meetings in a significant liaison role.

Theoretical and Methodological Approaches

Boundary objects are so called because they exist at and cross the boundaries of social worlds, such as the boundaries between communities of teachers, researchers, and curriculum supervisors. Star and Griesemer (1989) described boundary objects as “objects which are both plastic enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites” (p. 393). In addition to an interpretive flexibility across social worlds, Star (2010) clarified that boundary objects allow groups to cooperate without a consensus regarding their meaning or use, and they have a dynamic between ill-defined and well-structured uses. For this paper, we are particularly interested in how boundary objects represent the “stuff of action” (Star, 2010, p. 603) being the thing around which work is organized.

Work organized around boundary objects can represent learning in the form of changes in identity and/or participation. In their review of boundary crossing and boundary objects, Akkerman and Bakker (2011) identified in the literature four mechanisms of learning at boundaries: identification, coordination, reflection, and transformation. Akkerman and Bakker grouped identification and reflection together as mechanisms that reflected meaning-oriented learning processes with implications for people’s identities and perspectives, and grouped coordination and transformation together as practice-oriented learning processes with implications for activity. In this DBIR project, mechanisms related to practice and activity are most relevant for discussing learning. *Coordination* involves communication, efforts of translation, boundary permeability, and routinization, all of which contribute to people or objects moving smoothly across boundaries. Alternatively, *transformation* typically involves confrontation, recognition of shared problem spaces, hybridization resulting in new cultural forms, and crystallization of routines or procedures that embody new learning and require continuous joint work at boundaries. The in-between practices that can emerge in transformation are often held as goals but tend to be difficult to achieve and sustain, while coordination maintains the nature of boundary objects in ways that avoid the disruption of power structures or organizational hierarchies.

The primary sources of data used in this analysis are field notes taken in meetings of the research team, the researcher-supervisor planning meetings, and the TDT meeting. Results from an end-of-year teacher survey were also considered. The analysis here focuses on meetings where boundaries were exposed, such as in researcher-supervisor meetings and the TDT meeting. Iterating over the meeting notes, we identified arcs of work (Strauss, 1985) related to mathematical tasks, then identified design tensions (Tatar, 2007) within each arc.

Arcs of Work in Design-Based Implementation Research

The project’s initial arc of work established mathematical tasks as a potential boundary object. DBIR often begins with a search for an appropriate boundary object (Penuel, Coburn, & Gallagher, 2013) around which work can be organized. The object is not a “given” because the goals of a project emerge through collaboration. Though many of the participants in this project had collaborated before, identifying a common aim—and a new potential boundary object—was a key initial step in the work. The previous collaboration resulted in an online catalog of digital objects for supplementing a new science curriculum. The district supervisors’ initial stance was to develop a similar website and set of resources for Algebra 1, while researchers focused more on issues of curriculum implementation aligned with current research agendas, rather than the curriculum itself. Researchers’ early conversations suggested “productive adaptation” of curriculum (meeting notes, July 23, 2012), curriculum adaptation and authoring tools for teachers (meeting notes, August 28, 2012), analyzing teachers’ use of teacher-created materials (meeting notes, September 4, 2012), and a need for teachers to do task evaluation in a way that was simple but rooted in learning sciences research (meeting notes, August 28, 2012).

District supervisors pressed to pursue task evaluation and requested a selection of task rubrics, including rubrics for cognitive demand and language (meeting notes, September 17, 2012). The researchers suggested in the following meeting that the work focus primarily on learning trajectories, which was responded to by a supervisor with “Why not just use the curriculum guides we already have?” and “I don’t want to sound too pedestrian, but I want us to help teachers identify and use tasks that extend our current program” (meeting notes, September 24, 2012). This request, combined with an underdeveloped research base for learning trajectories in high school algebra (Sztajn, Confrey, Wilson, & Edgington, 2012), led to organizing project work around mathematical tasks. The focus on tasks fulfilled the supervisors’ desire for new, high-quality curricular materials, and the prospect of a task rating process would allow researchers to focus on productive curriculum adaptation and other aspects of teacher practice based in mathematics education and learning sciences research.

The second arc of work included the selection of task qualities teachers would use to rate tasks. Beginning with the supervisor-requested qualities of cognitive demand (Stein, Smith, Henningsen, & Silver, 2009) and language (Moschkovich, 2012), researchers assembled task evaluation guides and rubrics, attempting to consider what teachers actually do in implementing tasks (meeting notes, September 25, 2012). Additional potential rubrics and guides included alignment to standards and district learning goals, task “launch” (Jackson, Shahan, Gibbons, & Cobb, 2012), cultural relevance, and use of technology (meeting notes, October 9, 2012).

The main tensions in this arc of work centered on task adaptation and implementation support. The draft of the language rubric suggested task modifications for English learners, which raised concerns from a

supervisor: “One thing I worry about is, how will a teacher know if a task is appropriate for modification? Or if it has no guide for modification?” (meeting notes, October 22, 2012). Rather than modify the task, the supervisors requested supporting materials for ELLs that could support all tasks, including those in the district-adopted textbook. It was agreed that the development of these and other supporting materials would be pursued in a future phase of the project and that task qualities would apply to tasks only as written. This decision was applied later in the meeting when discussing the evaluation of tasks for cultural relevance. This rubric was set aside, with one supervisor suggesting that cultural relevance was related to teacher planning, not task quality. A different supervisor agreed, stating, “If teachers determine it’s a worthwhile task, there ought to be a place to make some notes about how that task is supported,” which again suggested a future phase of work focused on supporting the implementation of tasks. At meeting’s end, the list of task qualities to consider with the TDT were limited to the alignment with CCSSM and district goals, cognitive demand, language, and technology.

Following the October 22 meeting, a district supervisor sent researchers four tasks selected from the Mathematics Assessment Project (<http://map.mathshell.org>). The tasks had been scheduled into the Algebra 1 pacing guide for teachers to use at specific points in the curriculum. Each task represented a full lesson and included supports developed by the Mathematics Assessment Project such as student materials, slides, and discussion guidance. The district supervisors had added a cover sheet to each task that outlined a lesson plan and aligned the task to district content and language standards, but the task itself was unmodified from the original.

A third arc of work involved the design of a process for rating tasks within a group professional development activity. The researchers recognized there would be a “need [for] tasks for [teachers] to analyze using the draft rubric” (meeting notes, September 25, 2012) with sufficient practice with each rubric for teachers to have a “calibrated/shared understanding on that area of the rubric” (meeting notes, October 9, 2012).

Negotiations over plans for the TDT meeting surfaced tensions related to aims for tasks and relevant task qualities. In a November 5th meeting with supervisors, the researchers suggested structuring activities so that teachers “directly address their own questions about aligning to the goals of the CCSSM,” which was met with a supervisor concern that the conversation could devolve into irresolvable details of practice that might distract from the goals of the activity. Instead, the supervisor expressed a broader vision: “It would be nice to see a task that truly represents the kind of task that students should be capable of after three years of high school math.” A researcher altered the TDT agenda to reflect this vision of an “ultimate goal.” A supervisor also wanted teachers to work with the rubrics enough to be comfortable with both CCSSM alignment and cognitive demand, stating, “A task might have great cognitive demand, but if it doesn’t align to the standards, it doesn’t serve much of a purpose for us.” In a November 13th meeting, the liaison from the engineering team presented plans to show the TDT the curriculum website, how rated tasks might eventually be presented, and how other digital library resources were being algorithmically brought in to align with the district scope and sequence.

Teachers voiced their own perspectives at the TDT meeting, asking to consider their perceptions of students’ capabilities to engage in the task. For example, when discussing cognitive demand, teachers indicated their ratings depended on where in the curriculum they might use the task, or if the task was to be used with a relatively higher- or lower-ability group of Algebra 1 students. Seeking consistency in the rating process and consensus amongst raters, the researchers encouraged teachers to evaluate the tasks only as written and their “qualities independent of the particular groups of students” (meeting notes, December 1, 2012). This tension over task adaptation to classroom contexts persisted to the year-end teacher survey. When asked what factors influenced their use of tasks *not* captured in the rubrics, answers included “individual student abilities,” “the needs of my students,” “whether the task will be engaging/interesting to my students,” and “level of engagement from the students.” When asked how they would design PD around the CCSSM, responses included:

Teacher 2: “I would want a focus on how these resources can be used in my unique situation”

Teacher 7: “I really just wanted to focus on creating better tasks ... I don’t really care too much about the rubric”

Teacher 11: “[I would give] teachers resources that would enable them to create their own tasks”

Discussion

Figure 1 summarizes how each community interpreted mathematical tasks. For supervisors, tasks were a vehicle to add cognitive demand into the intended curriculum. The data shows the supervisors’ resistance to encourage teacher modification of tasks, preferring instead to identify high-demand tasks and to support their use. Teachers, by showing a desire for selection and adaptation of tasks, were grounded in their daily need to use or adapt available resources in the written curriculum meet students’ particular needs. For the researchers, tasks represented opportunities to engage teachers in using the rubrics. The particular task was somewhat unimportant, as the higher goal was to instill teachers with a set of curriculum design principles to apply to all tasks in the transformation between the written and intended curriculum. Lastly, for the web engineers, tasks were interpreted as a digital resource in a collection of resources that comprised a written curriculum.

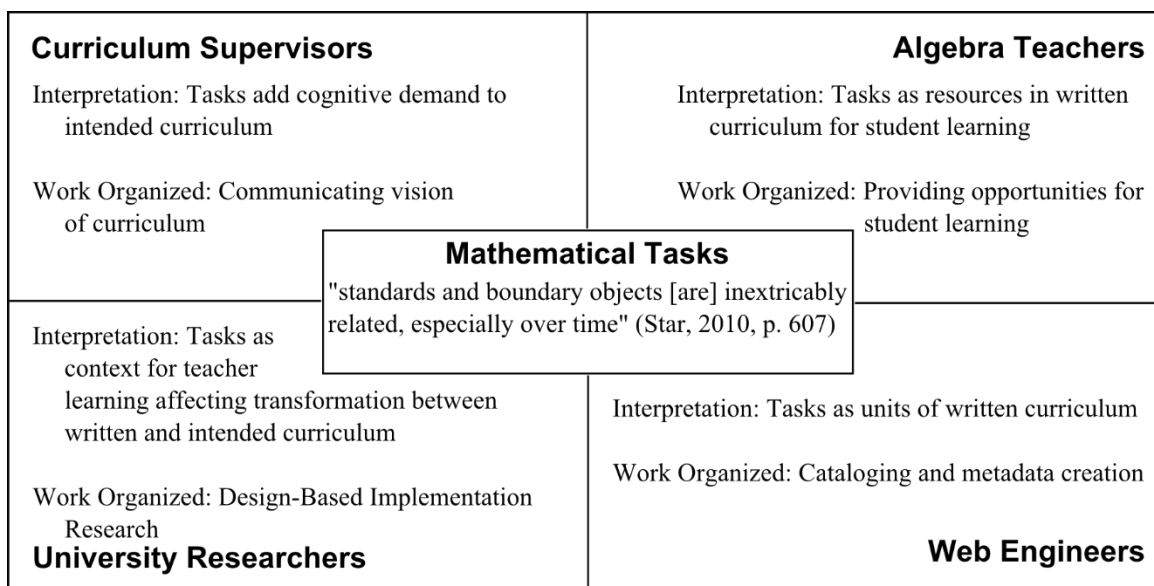


Figure 1. Mathematical tasks as a boundary object

It is also useful to consider how mathematical tasks organized work *within* each community in ways Star (2010) described as “invisible” or “back stage” work. The typically unseen work of the curriculum supervisor is the coordination of content and professional development to sculpt their vision for curriculum and instruction. Carefully chosen tasks are one way for supervisors to communicate that vision. For teachers, task-oriented back stage work includes how the unique demands of each classroom drives the search for tasks wherever they might be found, or the creation of new tasks when necessary. These unique demands are a reason why fidelity-based approaches to curriculum implementation tend to be problematic, despite supervisors’ and curriculum designers’ attempts to anticipate the needs of teachers. For DBIR researchers, invisible work includes collecting data of task analysis and enactment, as well as theorizing from patterns in that data. In this study, this invisible work would sometimes manifest itself in the form of a researcher-designed process to address a problem (e.g., “here’s a recommended strategy for encouraging teachers to participate in a webinar about task use”) when supervisors had more direct solutions available (e.g., “just tell us who needs to participate and we’ll call people we trust to see that it gets taken care of”). For the web engineers, the back stage work with tasks was the meticulous cataloging of the tasks and the creation of metadata records to describe their location in the curriculum, their alignment to standards, and how they were rated.

Akkerman and Bakker (2011) identified coordination and transformation as practice-oriented mechanisms for learning at boundaries. The evidence presented in this paper primarily describes mechanisms of coordination, as the arcs of work consisted largely of communication and translation between communities rather than confrontation and shared recognition of problem spaces. Continued research in this DBIR project should yield evidence of whether the supervisors’ vision for curriculum establishes a *communicative connection* across all spaces, if researchers’ *efforts of translation* of the scholarly literature impacts perceptions of tasks, or if task rating can become a *routinized* teacher practice in individual lesson planning and in teachers’ professional communities. Nine of 11 teachers reported in the year-end survey that their experience had a positive impact on their practice, with several indicating an influence of the task rating process into their daily selection and use of mathematical tasks. While this is a promising sign, more work at the boundary is required to understand how teachers’ participation is changing, and how change can be made scalable and sustainable.

There are opportunities for learning to occur as transformation in this project, particularly where confrontation persists around mathematical tasks. Task adaptation is one such area, but the existence of *hybridization* or *crystallization* is unclear as conflict was largely avoided or postponed. New forms of participation are tempered by histories and structures that exist across and within communities, such as teachers’ experiences with professional development and the roles of building administrators. Transformation may require a greater “coevolution” of teacher participation across contexts (Kazemi & Hubbard, 2008) and new theorizations within DBIR for addressing issues of power within educational systems. Transformation in this project could evolve as the crystallization of a task rating process that overcomes tensions between (a) teachers’ need to situate tasks in their classrooms, (b) supervisors’ need to provide curricular vision, (c) researchers’ need for scalability and sustainability of processes, and (d) web engineers’ need to classify and catalog tasks.

Conclusion

Collaborations between school districts, teachers, researchers, and other reform agents in design-based implementation research is a promising approach for building scalable and sustainable change in educational systems. This paper described how one such collaborative partnership organized their mathematics curriculum reform work around mathematical tasks. As a boundary object, mathematical tasks facilitated coordination despite a lack of consensus regarding their meaning. The lack of consensus exposed during joint work at the boundary of communities revealed opportunities for and achievement of learning, evidenced by changed practice and activity made possible mostly through the mechanism of coordination. In the case of task adaptation, a lack of consensus became a point of conflict. Successful confrontation of this conflict in the cooperative design process could yield new hybridized or crystalized task adaptation practices that, while difficult to achieve, have sustainable impact.

References

- Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, 81(2), 132–169. doi:10.3102/0034654311404435
- Cobb, P., & Jackson, K. J. (2011). Assessing the quality of the Common Core State Standards for Mathematics. *Educational Researcher*, 40(4), 183–185. doi:10.3102/0013189X11409928
- Jackson, K. J., Shahan, E. C., Gibbons, L. K., & Cobb, P. (2012). Launching complex tasks. *Mathematics Teaching in the Middle School*, 18(1), 24–29.
- 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. doi:10.1177/0022487108324330
- Moschkovich, J. N. (2012). Mathematics, the Common Core, and language. *Understanding Language: Language, Literacy, and Learning in the Content Areas*. Retrieved from <http://ell.stanford.edu/publication/mathematics-common-core-and-language>
- NGA Center/CCSSO. (2010). *Standards for mathematical practice*. Retrieved from <http://www.corestandards.org/Math/Practice>
- Penuel, W. R., Coburn, C. E., & Gallagher, D. J. (2013). Negotiating problems of practice in research–practice design partnerships. In B. J. Fishman, W. R. Penuel, A.-R. Allen, & B. H. Cheng (Eds.), *Design-based implementation research: Theories, methods, and exemplars* (Vol. 112, pp. 237–255). New York, NY: National Society for the Study of Education.
- Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher*, 40(7), 331–337. doi:10.3102/0013189X11421826
- Porter, A. C., McMaken, J., Hwang, J., & Yang, R. (2011). Common Core standards: The new U.S. intended curriculum. *Educational Researcher*, 40(3), 103–116. doi:10.3102/0013189X11424697
- Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values*, 35(5), 601–617. doi:10.1177/0162243910377624
- 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. doi:10.1177/030631289019003001
- Stein, M. K., Grover, B. W., & Henningsen, M. A. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, 33(2), 455–488. doi:10.3102/00028312033002455
- Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E. A. (2009). *Implementing standards-based mathematics instruction: A casebook for professional development* (2nd ed., p. 182). New York, NY: Teachers College Press.
- Strauss, A. (1985). Work and the division of labor. *The Sociological Quarterly*, 26(1), 1–19. doi:10.1111/j.1533-8525.1985.tb00212.x
- Sztajn, P., Confrey, J., Wilson, P. H., & Edgington, C. (2012). Learning trajectory based instruction: Toward a theory of teaching. *Educational Researcher*, 41(5), 147–156. doi:10.3102/0013189X12442801
- Tatar, D. (2007). The design tensions framework. *Human-Computer Interaction*, 22(4), 413–451. doi:10.1080/07370020701638814

Acknowledgments

This work was supported by a grant from the National Science Foundation (Award #1147590). The opinions expressed herein are those of the authors and do not necessarily reflect those of the NSF.

Evolution of Communities of Learning Practice in Higher Education: Collective Units of Analysis

Filitsa Dingyloudi and Jan-Willem Strijbos, Ludwig-Maximilians-University Munich,
Department of Psychology and Munich Center of the Learning Sciences
Email: Filitsa.Dingyloudi@psy.lmu.de, Jan-Willem.Strijbos@psy.lmu.de

Abstract: This study explores the development of Communities of Learning Practice (CoLP) over time, using Wenger, McDermott and Snyder's (2002) indicative developmental model of communities of practice (CoP) as a point of reference. The participants were 23 international graduate students from two different cohorts enrolled in a two-year Learning Sciences master's program. Students from each cohort formed a different community, with Community 1 having thirteen participants and Community 2 having ten participants. A mixed methods approach was adopted to capture community development from different angles. Data were collected through video recordings of community meetings and a needs analysis questionnaire. Results show that although the stages of development suggested by Wenger, McDermott and Snyder (2002) are apparent in a similar fashion in both communities, they do not follow a linear and smooth sequence, which signifies that communities go through the stages but in their own "living" way.

Introduction

Socioculturalism became particularly prominent in the 1980s, when researchers applied Vygotskian premises to understand learning, which emphasize the importance of socially shared activities in individuals' knowledge construction (John-Steiner & Mahn, 1996). Socioculturalists are not interested in the confined individual, but in "(...) events, activity and practice, and they are considered to be irreducible to properties of individuals" (Sawyer, 2002, p. 285). Therefore, the unit of analysis for socioculturalists is the socially situated practice, since the individual and the group cannot be studied in isolation due to their interrelation (Hatano & Wertsch, 2000; Sawyer, 2002). The individual is perceived as a singular plural (Nancy, 2000), which emphasizes the idea that "(...) individual actions always constitute concrete realizations of collective actions possibilities" (Roth & Lee, 2006, p. 30). As Vygotsky (1986) also claims, "(...) unlike elements, units are capable of retaining and expressing the essence of that whole being analyzed" (p. 211).

Even though behavioral researchers in the 1960s explored group formation and processes, Tuckman (1965) observed that the temporal changes that might occur over time were neglected. While building upon this observation, Tuckman (1965) and Tuckman and Jensen (1977) considered the temporal changes in group development and were particularly interested in the stages of small group development. On the basis of a review of fifty-five articles (Tuckman, 1965) and twenty-two articles (Tuckman & Jensen, 1977) on group formation and development, they suggested a group development model consisting of five stages, namely *forming*, *storming*, *norming*, *performing* and *adjourning*. These stages are characterized by different group structures (also: interpersonal realm) and tasks (also: task-activity realm). Although a full description and operationalization of this model and stages would have been insightful, it moves beyond the scope of this paper.

With respect to the development of Communities of Practice (CoPs), Wenger (1998b) proposed a developmental model consisting of five stages. The degree to which Wenger (1998b) built his developmental stages on the stages of small group development suggested by Tuckman and Jensen (1977), has not been explicitly clarified in any of his works (yet a small note at the back of the Wenger, McDermott and Snyder (2002) book signifies some awareness of these small group stages). The term CoPs refers to groups of people who share a common interest in a domain of human endeavor and mutually engage in a process of working together to accomplish enterprises like knowledge exchange and construction, or complex problem solving within their shared domain of interest (Lave & Wenger, 1991; Wenger, 1998a). According to Wenger (1998b), communities move through various life phases and stages of development with their own rhythm, which is associated with the members' interaction levels and the kinds of activities they undertake. Wenger (2000) and Wenger et al. (2002) emphasize the developmental nature of CoPs by adopting a humanization approach: CoPs grow and develop like humans, going through different stages in their lives but still remaining the same persons. Wenger (2000) highlights that although the reasons for staying together over time, may differ from the reasons that initiated the community formation, all these reasons are relevant to community members.

Wenger (1998b) initially suggested a developmental model, which after some variations in the labeling of the stages, came to be what is shown in Figure 1 (Wenger et al., 2002). According to Wenger (2000) and Wenger et al. (2002), the model has indicative and not prescriptive connotations and value. CoPs do not necessarily progress through all suggested stages, since they may skip, merge, revisit or not even reach some of

the stages. They grow in their own rights, and their developmental cycle may vary. Regardless of the shape of the developmental cycle, CoPs do have a lifecycle and they should not be perceived as one-state stable entities.

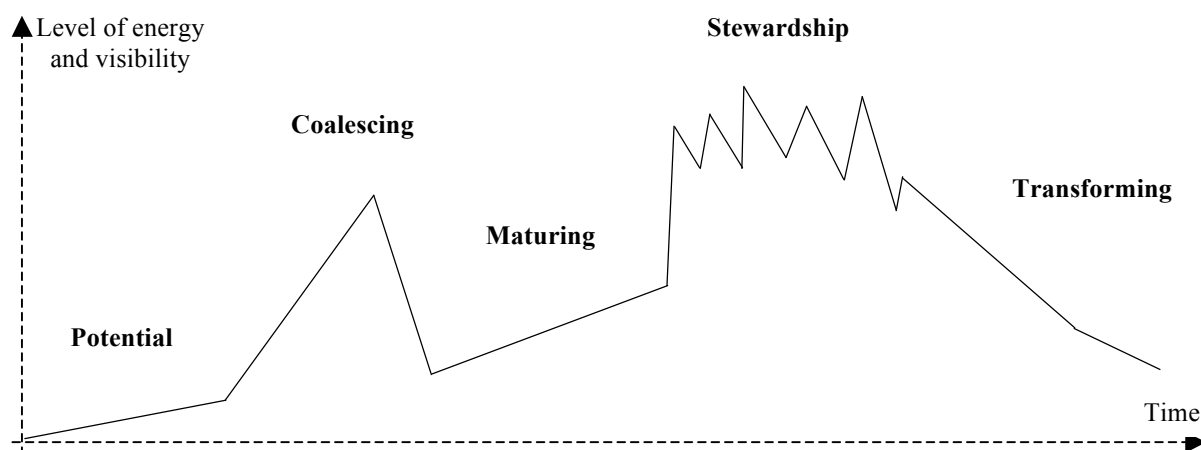


Figure 1: Stages of development of Communities of Practice (Adapted from Wenger et al., 2002)

Based on Wenger et al. (2002), the first stage, called *potential*, refers to a network of individuals who identify an interest in a domain or deal with a similar situation. Within this network of people, the idea of community formation emerges and individuals start to view their domain of interest from a collective perspective, which further intensifies the need for interaction among the potential members. Once individuals identify the potential for the community to form, the critical *coalescing* stage in the lifecycle of the community follows. At this stage, the community members start building relationships, establishing a public (i.e., community meetings) and private (i.e., one-to-one discussions) rhythm and finding opportunities to help each other. With the coalescing stage the community becomes more tangible in the experiences of its members. After establishing its formation, the community proceeds to the *maturing* stage. During this stage the community forms its own identity by exchanging information, developing focused activities, establishing ways for addressing problems, and sharing responsibilities. The community then broadens its scope, through the *stewardship* stage, during which community aims to solve new problems and increases the complexity of the activities, relationships and interests involved. Newcomers may also enter the community at this stage, and through this modified membership the community is prompted to readapt to the dynamic needs of its old and new members. During this stage it is time for the community to question and reflect on the gained values and as a result a need for reconfiguration might arise. The community also explores and tries to establish connections and implications of gained knowledge beyond the community boundaries. The final stage of community development is the *transforming* stage, when the community has fulfilled the purposes of its existence and members move on into other CoPs that are more relevant to their new developmental trajectories.

Research questions

Despite the theoretical identification of the developmental stages for CoPs by Wenger et al. (2002), there are to date no studies that empirically examined whether – and to what degree and/or which temporal order – communities develop along the lines of these indicative stages. This study aims to address this gap by exploring the occurrence of the stages within two communities of learning practice (a detailed theoretical account is provided in Dingyloudi & Strijbos, 2014) that developed in parallel to a Learning Sciences masters' program. The following research questions will be addressed: (1) How do communities develop over time in the public space?, (2) To what extent are the theoretical stages of community development represented in the communities under study?

Method

The present study applied a mixed-methods approach. In mixed methods there is concurrent implementation of both qualitative and quantitative data collection, with qualitative data being more dominant in the present study, due to the main research interest, which is to deeply explore a composite phenomenon within a particular setting (Creswell, 2008). Wenger, Trayner and DeLaat (2011) also highlight the importance of being able “(...) to attribute observable outcomes to community and network activities so that one can establish enough causal links to go beyond mere correlations between distinct data streams.” (p. 8).

Participants

The participants were twenty-three international graduate students ($M_{\text{age}} = 25.1$, $SD_{\text{age}} = 2.3$) enrolled in a two-year research oriented Learning Sciences master's program. The 23 students are part of two different student cohorts and constitute two communities of learning practice in parallel to the master's program: Community 1 consisted of thirteen participants ($M_{\text{age}} = 25.2$, $SD_{\text{age}} = 2.8$) out of the twenty-seven students enrolled in their cohort. Community 2 consisted of ten participants ($N = 10$, $M_{\text{age}} = 25$, $SD_{\text{age}} = 1.8$) out of the thirty students enrolled in their cohort. Participation in both communities was voluntary and participants were free to join or withdraw from any community meeting.

Design

The communities of learning practice were formed by students, who gathered together as plain peers in order to address their problems and help each other with respect to academic skills that were relevant to their study program. Peers voluntarily participated in the public space of the community through community events (i.e., informal face-to-face community meetings), which lasted approximately two hours each. There were six and five events in Community 1 and Community 2, respectively. The community events were co-organized and co-structured among the community members and a non-peer facilitator, who was present at every community event in order to facilitate the members' interactions and community activities. The sharing mechanism, used by the community members and coordinated by the community facilitator, was face-to-face peer feedback on work in-progress that was relevant to the study program. The lifespan of the community was one semester, with Community 1 taking place during the second semester of their study program and Community 2 during the first semester.

Instruments

Video recording. Video recording allows for retrospective analyses of group interactions (Brown, 1992; DiSessa & Cobb, 2004). In total eleven community events (i.e., face-to-face community meetings) were recorded: six events for Community 1 and five events for Community 2. The events lasted approximately 2 hours each, resulting into 22 hours of video data. A coding scheme was developed to identify and explore the occurrence of the five developmental community stages suggested by Wenger et al. (2002). In the first step three coders identified relevant episodes for each video within which potential thematic units constituting the five stages could be observed (Krippendorff's $\alpha = .83$). In the second step, several coding trials were conducted on the video of two events. After refining the coding scheme, all 22 videos were analyzed by two independent coders and interrater reliability was calculated. The interrater reliability for both communities was high (Community 1: Krippendorff's $\alpha = .85$; Community 2: Krippendorff's $\alpha = .81$).

Needs analysis. A needs analysis was conducted via a questionnaire designed for this study, and administered on the cohort level before the initiation of each community. The questionnaire items were context dependent and aimed to capture "(...) the gap between real and the ideal that is both acknowledged by community values and potentially amenable to change" (Reviere, Berkowitz, Carter, & Gergusan, 1996, p. 5). The items covered three topics: (a) background information, (b) needs realization, and (c) feedback preferences. Background information addressed potential community members' former participation in extracurricular groups and the experienced value, the frequency of seeking help from peers in general, and extent of acknowledgement of improvement after receiving peer feedback. Needs realization addressed students' needs regarding academic skills they would like to practice and receive support in an extracurricular group, and needs regarding frequency of community events. Feedback preferences addressed the extent to which students perceive their peers as valuable feedback source, and appreciate their peers' opinions on academic work. The needs analysis aimed to uncover the potential for a community to emerge.

Findings

The needs analysis was considered vital to identify the potential stage of development, since it indicates the possibility for a community to emerge based on collective goals. Since the community sharing mechanism was peer feedback, it was of major importance to also identify any previously experienced value of peer feedback (see Table 1).

Table 1: Main results of needs analysis prior to community formation for both cohorts

	Cohort 1 ($N = 15$)	Cohort 2 ($N = 28$)
Seeking help from peers on academic work	80%	89%
Appreciation of peers' opinions on academic work	79%	86%
Extent of acknowledgement of improvement after receiving peer feedback	73%	85%
Extent to which they perceive peers as valuable feedback source	57%	82%
Willingness to participate in extracurricular support groups	100%	100%

The results of the needs analysis indicate that there is a potential for students to address their needs regarding the academic requirements through formation of and participation in a community together with their peers. This is reflected through the extent to which they had sought help from peers, appreciated peers' opinions and acknowledged improvement in their performance after receiving peer feedback. In addition, most of the students perceived peers as valuable feedback sources and they all expressed their willingness to participate in an extracurricular group that could support them with academic skills.

The analysis of the video data for both communities revealed the presence of the developmental stages described by Wenger (2000). Also, both communities seem to follow a similar, but not identical, developmental cycle. As illustrated in Figure 2, analysis of the first community event of Community 1 revealed that 88% of the identified thematic units referred to the *potential* stage and 12% to *coalescing*. Respectively, content analysis of the first community event of Community 2, revealed that 60% of the identified thematic units referred to the *potential* stage, 30% to *coalescing* and 10% to the *maturing*. The content analysis of the second community event of Community 1 revealed that 70% of the identified thematic units referred to *coalescing* stage, 20% to *potential* and 10% to *maturing*. Respectively, content analysis of the second community event of Community 2, revealed that 75% of the identified thematic units referred to *coalescing* and 25% to *potential*. The content analysis of the third community event of Community 1 revealed that 50% of the identified thematic units referred to the *potential* stage and 50% to *coalescing*. Respectively, the content analysis of the third community event of Community 2 revealed that 50% of the identified thematic units referred to the *stewardship* stage, 33% to *potential* and 17% to *coalescing*. The content analysis of the forth community event of Community 1 revealed that 33% of the identified thematic units referred to the *potential* stage, 33% to *maturing*, 22% to *coalescing* and 11% to *stewardship*. Respectively, the forth community event of Community 2 revealed that 36% of the identified thematic units referred to the *coalescing* stage, 27% to *potential*, 18% to *maturing*, and 18% to *stewardship*. The content analysis of the fifth community event of Community 1 revealed that 60% of the identified thematic units referred to the *stewardship* stage and 40% to *coalescing*. Respectively, the content analysis of the fifth and final community event of Community 2 revealed that 62% of the identified thematic units referred to the *transforming* stage and 38% to *coalescing*. Community 1 continued its public development in a sixth community event, and its content analysis revealed that 73% of the identified thematic units referred to the *transforming* stage, 18% to *coalescing* and 9% to *stewardship*.

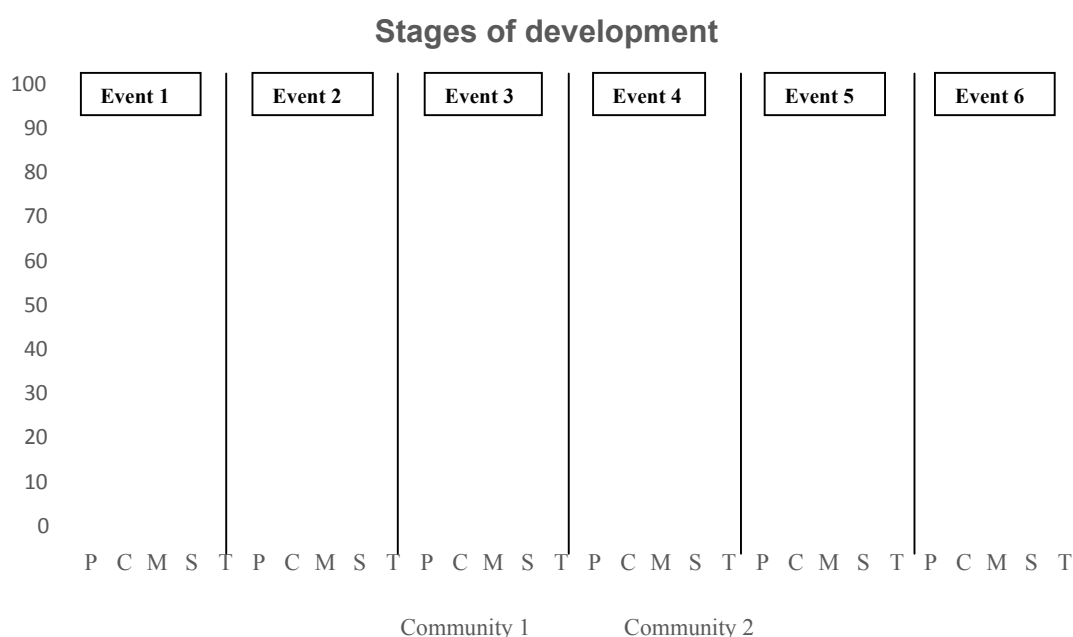


Figure 2: Stages of development of Community 1 and Community 2

Discussion and Conclusion

The present study explored how communities of learning practice (i.e., extra-curricular communities in an educational setting) develop over time. In particular, whether and to what extent the five theoretical stages of development (Wenger, 2000; Wenger et al., 2002), were represented in the developmental cycles of the observed communities. Both communities of learning practice went through the five stages of development, but not in a linear sequence. This is in line with the evidence for the non-linear and context-dependent order of the stages for small group development as groups may return to a previous stage or even skip a stage (Tuckman & Jensen, 1977). Furthermore, both communities' developmental cycles evolved in a similar fashion, with slight

variations. The most prominent stage in both communities was the *potential* stage, since it emerged and re-emerged in every community event up until the fourth event for both communities. This prominence reflects that either the community (a) continuously reformed itself by identifying different potential interests and ways to connect to each other, or (b) struggled to find a rhythm and build relationships strong enough for the community to further develop. The least prominent stage pertains to *stewardship*. This might be attributed to the way the communities structured themselves in this particular setting, i.e. without involving experts as oldtimers and novices as newcomers, which is a typical feature of CoPs but not for the present CoLPs (see Dingyloudi & Strijbos, 2014). Aspects of prominence and relevance of specific stages in various contexts should be further examined.

Overall, the findings of the present study address the gap between the theory of five stages of development in CoPs (Wenger, 2000; Wenger et al. 2002) and their actual exploration by researchers and practitioners. This study provides some empirical evidence for the five theorized developmental stages at a community-level as well as some empirical evidence that these stages can be inferred from a needs analysis and video-recordings of community meetings. With respect to theoretical implications, this study confirms the applied value of the theoretical stages of development for communities in an educational setting, although an adaptation of the model can be explored further in order to capture elements of the CoLP community model (see Dingyloudi & Strijbos, 2014). Future research will explore social network data and value creation stories (see Wenger et al., 2011) – collected as part of the research project, but their inclusion is beyond the scope of the present paper – which may aid interpreting this prominence. Our findings suggest researchers and educators to observe how communities under study or in implementation develop, in order to further facilitate the development and aliveness of communities, while taking into consideration the stage or combination of stages communities experience at some given time.

References

- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, 2, 141-178.
- Creswell, J. W. (2008). *Research design: Qualitative, quantitative and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Dingyloudi, F., & Strijbos, J. W. (2014). Communities of learning practice: Balancing emergence and design in educational settings. In *Proceedings of the ICLS 2014 conference "Learning and becoming in practice"*. Boulder, CO: International Society of the Learning Sciences.
- DiSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *The Journal of the Learning Sciences*, 13, 77-103.
- Hatano, G., & Wertsch, J. V. (2001). Sociocultural approaches to cognitive development: The constitutions of culture in mind. *Human Development*, 44, 77-83.
- John-Steiner, V., & Mahn, H. (1996). Sociocultural approaches to learning and development: A Vygotskian framework. *Educational Psychologist*, 31, 191-206.
- Nancy, J.-L. (2000). *Being singular plural*. Stanford, CA: Stanford University Press.
- Reviere, R., Berkowitz, S., Carter, C. C., & Gergusan, C. G. (Eds.) (1996). *Needs assessment: A creative and practical guide for social scientists*. Washington, DC: Taylor and Francis.
- Sawyer, R. K. (2002). Unresolved tensions in sociocultural theory: Analogies with contemporary sociological debates. *Culture & Psychology*, 8, 283-305.
- Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin*, 63, 384-399.
- Tuckman, B. W., & Jensen, M. C. (1977). Stages of small group development revisited. *Group and Organizational Studies*, 2, 419-427.
- Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.
- Wenger, E. (1998a). *Communities of practice: Learning, meaning and identity*. Cambridge: Cambridge University Press.
- Wenger, E. (1998b) Communities of practice: Learning as a social system. *System Thinker*, 9, 1-5.
- Wenger, E. (2000). Communities of practice: Stewarding knowledge. In C. Despres & D. Chauvel (Eds.), *Knowledge horizons: The present and the promise of knowledge management* (pp. 205-224). Boston, MA: Butterworth-Heinemann.
- Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Boston: Harvard Business School.
- Wenger, E., Trayner, B., & De Laat, M. (2011). *Promoting and assessing value creation in communities and networks: A conceptual framework*. Heerlen, the Netherlands: Ruud de Moor Centrum.

Recasting the Textbook: Student Creation of Interactive Digital History Textbooks with Primary Source Documents

Molly B. Zieleszinski and Paul Franz, Stanford University
Email: mbullock@stanford.edu, pefranz@stanford.edu

This research report details phase 1 of an ongoing design based research project. Motivated by the current context of the emerging participatory youth media ecology and national support for digital textbooks in the hands of all students, we set out to determine the feasibility of students creating their own interactive digital textbooks using primary source documents. In collaboration with a history teacher, the research team designed a theory driven curriculum supplement and implemented this with four 10th grade history classes. The process and preliminary findings are presented and discussed.

Today's youth commonly engage with social media, media distribution, and digital media production as part of a new kind of shared culture (Ito, 2009). Engagement in these friendship driven practices and interest driven practices are associated with the development of new forms of media literacy skills and informal peer learning environments (Ito 2009). While teens are coming of age in this new media ecology, their high school history teachers sit on the opposite end of the digital native/digital immigrant spectrum often lacking technological fluency or technological pedagogical content knowledge (TPCK) to introduce technology into the curriculum. Complicating the situation further is the proliferation of digital technologies in U.S. classrooms; since 1991, the number of Internet connected devices available per student in schools has risen steadily (Warschauer, 2010) and the end is not in sight. Most states are gearing up for digital administration of common core aligned tests, a feat that all but requires a nationwide increase in student to device ratio. Furthermore, U.S. Secretary of Education Arne Duncan has set the national goal to have digital textbooks in the hands of all U.S. students within the next five years. So the stage is set: youth computer access and use is higher than ever, access to technology in schools is increasing, and technological skills are integral for those entering the labor force in the next twenty years (Warschauer, 2010). How will the story unfold? Well, if we continue on our current trajectory technology will be "oversold and underused" (Cuban 2001) and as a result curriculum and practice will remain mostly unchanged by this perfect storm. Alternatively, students may continue to experience school-as-usual but with modifications in method of delivery. As current analogue practices are being furiously digitized. In this version of the future, tests and textbooks, worksheets and essays, will be delivered and completed digitally but otherwise remain largely unchanged. A third possible outcome entails students engaging with technology for learning in new, authentic, and innovative ways that simultaneously aid in the development of technological fluency, content specific practices and content knowledge. How will this perfect storm of competing contexts in U.S. unfold in American Education? How might textbooks be used in innovative ways as tools for learning?

The research presented in this report is specifically interested in education's digital future in the high school history class. We consider historical thinking and the ways in which technology can be used in development of this skill. Wineburg (2001) introduces the notion that "Historical thinking, in its deepest forms, is neither a natural process nor something that springs automatically from psychological development". Students in high school history classes are unlikely to develop mature historical thought through instruction that relies on memorization of facts and dates. Despite this, many history teachers emphasize student learning of exhaustive lists of historical facts (Barton & Levstik, 2003). Monologic instruction is characterized by authoritative voice (Bhaktin 1981) and tacitly demands that students accept one true history without actually engaging with the information. In contrast, the development of true historical thinking requires a learner to linger in the ambiguity, navigating an underlying tension between the strangeness of history and its familiarity (Wineburg 2001). Historians achieve a necessary balance between understanding historical moments in their true context (strangeness) and determining their relevance and application to our world (familiarity). This work is characterized by evaluation of evidence, argumentation, and the application of historical knowledge (Beyond the Bubble 2013) and it is often conducted by reasoning alongside primary source documents that represent a multiplicity of perspectives. As such the adoption of document based teaching and learning in the high school history class is germane to the development of historical thinking.

Although many history classrooms engage in document based instruction, there is a gap between the abundant digital resources available through archives, museums, and libraries and the dearth of resources that are readily available in high school history class. Motivated by the current context of educational technology, this research project explores what might happen if we bridged that gap, connecting students to available resources and asking them to engage in the work of historians. Given the rise in digital media production and distribution in youth culture, we explore the feasibility of digital authoring tools as a means to move students beyond uncritical content consumption and build their capacity for historical thinking. This report is a presentation of

the methods and preliminary findings from the first phase an ongoing design based research agenda. This initial phase of the project was motivated by the following research questions. RQ 1) To what extent is it feasible for high school students to craft their own digital interactive history textbooks using primary source documents? RQ 2) To what degree did students engage in the practices of authoring history in crafting their textbook (i.e., sourcing, contextualizing, multiple perspective taking, choosing relevant primary sources)? RQ 3) To what degree did students utilize the affordances of the technology specifically in service of historical thinking? To what degree did students utilize the affordances to share multiple perspectives on historical events?

Conceptual Framework

This study has two key constructs. The first construct of interest in this work is *historical knowledge*. The development of historical knowledge is a by-product of historical thinking. As such, we treat the process of authoring an interactive digital history textbook as a potential act for historical thinking and treat the student created artifacts as a demonstration of historical knowledge. We acknowledge that this a conceptual leap and recognize the underlying assumption that to demonstrate historical knowledge one must engage in some degree of historical thinking. As this was a preliminary study in an ongoing project, we chose to focus on the feasibility of the medium for the presentation of historical knowledge and planned for phase 2 to collect data more explicitly tied to historical thinking. We draw on 3 main components from the framework of van Drie and van Boxtel (2004) in our operationalization of historical knowledge, and as such conceptualize historical knowledge with regards to use of historical questions, use of primary sources, and contextualization. Additionally, we consider multiple perspective taking, to be a specifically relevant in this research given potential of the medium for displaying a multiplicity of historical voices. Multiple perspective taking is the presentation and exploration of different ways of thinking about moments in history as they were thought of at the time and is demonstrated through a combination of sourcing and contextualizing. Future research will examine the remaining components of the van Drie and van Boxtel (2004) framework: argumentation, use of substantive concepts, and use of metaconcepts. The second construct of interest is the *affordances of technology*. We operationalize this explicitly in connection with the iBooks author, the tool chosen for digital authoring. Each of the interactive widgets afforded by iBooks author is has a unique set of affordances that can be utilized during authorship. We seek to answer our research questions by taking up these constructs first separately and then together through examination of specific affordances of the technology as the co-occur with demonstration of historical knowledge.

Research Design

At the onset of the project, we set out to utilize design based research methods to find out if it was feasible for high school students to craft their own digital interactive history textbooks using primary source documents. In the early stages of the project, the research team engaged in the selection, purchase, organization, and deployment of the hardware and software identified as having the greatest potential value for history learning with primary sources. Concurrently, we conceptualized and created a two-week curriculum supplement integrating best practices in history instruction grounded in existing scholarship. After selecting a site for implementation, we began a partnership with the teacher chosen to collaborate in the study. We conducted multiple observations of his history class and documented via field notes and memos. We worked with the teacher to identify the most conducive conditions for students' collaborative construction of multimodal interactive digital history textbook chapters. We identified alignment between affordances of our chosen digital authoring software (such as callouts, galleries, interactive quizzes, and 3D models) and use cases for sharing multiple perspectives and designed curricular components intended to support the students and teacher in exploring and exploiting these tools specifically for the purposes of authoring history with a thick narrative. One strategy for this included the design and deployment of templates to scaffold student groups in the task of authoring a historical narrative that integrates multiple perspectives, with each perspective introduced via one or more primary source documents. Another tool developed to support students in this task was a model or exemplar chapter. Support tools and curriculum were refined over multiple iterations.

To address the gap in availability of digital resources, the research team liaised between the collaborating history teacher in our study and a local archive. This resulted in a visit by the archivist to each of the four tenth-grade classes. A university student (with expertise in this area) conducted a follow-up lesson covering important considerations for working with primary source materials. Finally, the research team compiled a well-curated database of primary source documents that was organized to meet the needs of the learning environment (as identified in multiple interviews and correspondences with four tenth-grade history teachers). These resources were introduced as part of our curriculum.

95 10th grade students from a northern California charter school participated in a 14-day our curricular intervention. 3 members of the research team and the collaborating teacher implemented the curriculum supplement. During the intervention, the research team collected a plethora of data (over 3 terabytes) including but not limited video files of each lesson, audio files of students working in small groups, screen capture and

keystroke data captured as students worked on computers, field notes, research memos, student surveys, summative transfer task, and the final artifacts. Findings in this research report were generated via a content analysis using the student created digital history textbook chapters as the message source. Concurrent analyses of a transfer task, multiple student surveys, and in depth qualitative look at a subset of textbook chapters are in progress and the results of all will be used to inform the next iteration and implementation of our curriculum supplement.

Preliminary Findings

History as Usual: Pre-Implementation Observations

A comparison of field notes from observations during the normal history class to field notes taken during project implementation indicate that our intervention allowed for a higher frequency of technology use, content creation and creativity. During five observations of the normal history class the only use of technology observed was laptop and projector use by the teacher. Content creation by students observed prior to the intervention consisted primarily of completion of lecture-aligned worksheets. While we did not observe specific instances of creativity during these observations, student work hung on the walls indicates this to be an occasional part of the normal history class. Student generated maps; propaganda posters and research projects were prominent in the classroom. Finally, pre-observations revealed an aspect of classroom culture related to communication. Students frequently engaged in turn and talk activities and we observed a consistently high on task dialogue and rate of participation in whole group discussion.

History Recast: Digital Media Production

One unexpected outcome of our implementation was the adoption of our curated database of standards aligned primary source documents. Even before we began the intervention, we received requests for access from other members of the history department. This was followed by requests by other teachers in the later school network.

Figure 1 presents a summary of the formal characteristics of the interactive digital textbook chapters authored by the students who participated in our intervention. The students were extremely generative during the two-week implementation writing nearly 14,000 words distributed across 73 chapters. During class wide project debriefs held after implementation, students lamented the time constraints that limited further productivity. They also expressed what we perceive to be a high degree of enthusiasm and engagement both throughout the intervention and during the final debrief.

Figure 1: Summary of student-generated content during phase 1 implementation

Group	# Students	# Groups	Words	Paragraph	Media (video)	Media (images)	Pages	Chapters	Words/Chapter	Chapters/Group
Class 1	25	13	5759	68	0	35	115	20	288	1.5
Class 2	26	12	4645	119	0	37	109	14	332	1.2
Class 3	25	13	1675	85	3	52	133	24	70	1.8
Class 4	22	11	1804	70	2	36	83	16	113	1.5
Total	98	49	13883	342	5	157	433	73	190	1.5

Our content analysis of student artifacts includes coding for 15 different instantiations of historical knowledge and 16 variables related to technological affordances. Currently the research team is in the process of coder training in order to establish inter-rater reliability for each of the 31 variables in our theory driven codebook. A complete recoding of the data set by multiple coders is scheduled to follow. As such, the subset of findings presented below is centered on trends identified by the authors related to the perspective taking variable. Specific counts and relative frequencies have been temporarily omitted (pending the establishment and calculation of group coding reliabilities) but will be included upon presentation of the paper.

RQ 1) To what extent is it feasible for high school students to craft their own digital interactive history textbooks using primary source documents? Having completed phase 1 of this work in four high school classes it is determined that it is definitely feasible for students to create interactive digital history texts utilizing primary source documents.

RQ 2) *To what degree did students engage in the practices of authoring history (and therefore historical thinking) in crafting their textbook (i.e., sourcing, contextualizing, multiple perspective taking, choosing relevant primary sources)?* Preliminary coding of student created artifacts indicated that students' demonstration of historical knowledge varied widely within the sample. This was less so for the key variable of interest- perspective taking. A vast majority of the digital textbook chapters limited their portrayal of historical events to two perspectives. Furthermore, while presentation of two voices is technically multiple perspective taking, it is the most shallow form. Presentation of two perspectives is not in itself necessarily lacking depth, thus we describe artifacts as shallow when they mirror the traditional monologic treatment of history that is common in high school history textbooks. Shallow artifacts authoritatively present typical binary classifications of historical events (pro ally vs pro enemy). This was at the expense of a more dialogic approach to construction of deep historical knowledge through exploring a multiplicity of voices and/or a introducing the notion of who is left out of their analyses. Although this deep historical thinking was the primary design goal of the researchers for this preliminary study, a very small subset of the sample achieved this goal. One possible explanation for this is insufficient instruction related to this skill or insufficient time for skill development. The students had been introduced to the concept of perspective taking prior several times during the year prior to implementation of our study but these instances were demonstrated by the teacher rather than opportunities to do so themselves. Another possible explanation for this is a misalignment between the goal of the researchers and the evaluation of the student work by the teacher. During whole post project debriefs, students from two of the four classes independently reported a tension between their eagerness to use the tools to present a multiplicity of historical voices and an urge to focus on what they were being graded on. Grades were assigned based on the presentation of two perspectives. Each of these interpretations has considerable implications for the next iteration and implementation of our intervention.

RQ 3) *To what degree did students utilize the affordances of the technology specifically in service of historical thinking?* To what degree did students utilize the affordances to share multiple perspectives on historical events? Instances of multiple perspective taking co-occurred with affordances of the software in a majority of the student artifacts. Provision of templates, models and coaching supported students' use of interactive widgets as resources for displaying multiple perspectives. This was evidenced by the fact that student authors only occasionally demonstrated multiple perspectives in forms that varied from one of the numerous templates provided. This raises the question of whether the templates were a factor limiting the innovative use of the affordances in demonstration of historical knowledge. Prior to our next implementation, we intend to develop some insights as to whether this was the case through a comprehensive interrogation of our data and analysis of interactions between the variables from each of our constructs.

Summary

We set out to investigate the feasibility and outcomes of digital textbook authorship by high school students. Preliminary findings from phase 1 implementation indicate that while feasible, the intervention requires considerable redesign prior to phase 2. While this is true, our rich multimedia data set of student generated content using primary and secondary sources when taken in conjunction with anecdotal evidence of the enthusiasm of the students, teacher, and extended school network provide a warrant for the continuation of our research. It is noteworthy although not unexpected that small contextual factors have huge implications for the utility and uptake of technological affordances for document based historical thinking and demonstration of historical knowledge. Some of the major contextual factors shaping these outcomes include time to complete task, establishing and setting norms for technology mediated learning, modeling desired outcomes with exemplars, and use of templates designed to promote interrogation of a multiplicity of historical voices. Finally, alignment between tools, tasks, and how the students are being evaluated is essential for success in the adoption of digital authoring tools as means to support document based reasoning. Our next phase of analysis has promise for contributing to theory and scholarship regarding the ways in which specific technological affordances support and/or limit specific aspects of the construction and demonstration of historical knowledge. This information has potential for informing history teaching practice and technology design for history learning.

References

- Barton, K. C., & Levstik, L. S. (2003). Why don't more history teachers engage students in interpretation?. *Social Education*, 67(6), 358-358.
- Beyond the Bubble: A New Generation of History Assessment- Our Approach. (n.d.). Beyond the Bubble. Retrieved November 1, 2013, from <http://beyondthebubble.stanford.edu/our-approach>
- Itō, M. (2009). *Living and learning with new media: Summary of findings from the Digital Youth Project*. MIT Press.
- Larry, Cuban. (2001). Oversold and underused: Computers in the classroom.
- van Drie, J., & Van Boxtel, C. (2004). Historical reasoning: A comparison of how experts and novices

- contextualise historical sources. *International Journal of Historical Learning, Teaching and Research*, 4(2), 84-91.
- 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.
- Wineburg, S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past*. Temple University Press.

Acknowledgements

Thank you to our advisors: Sam Wineburg and Roy Pea; our funder- MediaX, especially Martha Russel; our cooperating teacher: Bobby Cupp and his amazing history students, members of the Digital History Research Group: Max Alexander, Tanner Veal, Jeremy Jiminez, and Laura Moorhead; and others who have advised us along the way including Rob Lucas, Joel Breakstone, and the members of the SHEG group.

Sources of Affect around Interdisciplinary Sense Making

Benjamin D. Geller,* Julia S. Gouvea,* Vashti Sawtelle, and Chandra Turpen

Department of Physics, University of Maryland, College Park, College Park, MD 20742

Email: geller@umd.edu, juliasvoboda@gmail.com, davisvas@gmail.com, chandra.turpen@gmail.com

* Authors contributed equally to this paper

Abstract: We unpack an episode in which a case study student in an Introductory Physics for Life Scientists (IPLS) course experiences positive and negative affect coupled to sources of frustration and satisfying resolution. We argue that the positive affect that the student experiences stems from an alignment between his identification as a sense-maker and his epistemological view of this physics course as one that values sense-making. Conversely, we attribute his frustration to a tension between his identity and his epistemological view of his biology courses as descriptive and fact-driven. We discuss some implications of this model for students engaged in interdisciplinary sense-making more generally. In particular, we suggest that Gavin's frustration with the lack of attention to mechanism in his biology courses might ultimately serve to *strengthen* his sense of interdisciplinary connectedness and satisfaction.

Introduction: Different Reasons for Appreciating Physics

In recent years, efforts have been made to design introductory physics courses that are specifically tailored to life science students (e.g. Meredith & Redish, 2013). A number of objectives are routinely cited for doing so, including the increased importance of physical modeling and quantitative approaches in upper-division biology coursework, the need to train future physicians in methods and technologies developed in the physical sciences, and a general recognition that science disciplines are increasingly integrated and dependent on each other for inspiration and innovation. These objectives often align with national calls for substantial reform of the undergraduate curricula for life science and pre-health profession students (National Research Council, 2003).

In practice, instructors recognize the importance of affect as both a mediator of participation in and an outcome of Introductory Physics for Life Scientists (IPLS) courses. Meeting the interdisciplinary objectives stated above seems to require that students be open to participating in physics. Because life science students often have negative orientations towards physics, helping life science students come to “appreciate” or “like” physics is seen as an important component of interdisciplinary learning. This goal is sometimes made explicit, but more often is conveyed implicitly through efforts to include content that biologists would find “exciting” or of particular “interest.” However, the sources and consequences of affective responses have not been well researched.

There are many reasons why life science students might profess to like or appreciate physics. Perhaps, for example, students see utilitarian value in their physics class in preparing them for the MCAT exam or in later coursework. Perhaps students like physics because they find it easy and it makes them feel competent or confident as a learner. Perhaps students like physics because they find it intrinsically satisfying to make sense of why objects behave as they do. These different sources of positive affect can have consequences for how and if students participate in physics and potentially for their participation in other science disciplines. A better understanding of the pathways that lead to affective responses can help instructors be more intentional about the kinds of emotions they are trying to foster in their students.

The focus in this paper is on affective responses that are triggered by moments in which physics is helping a student *make sense* of phenomena previously encountered in his biology courses by seeking coherent mechanistic accounts of these phenomena (Hammer et al., 2005). By focusing on how and why these moments of explanatory coherence generate positive affective response, we hope to suggest ways of inviting students to participate in physics as well as foster an appreciation for thinking and learning across disciplinary boundaries.

In order to better understand both the sources and consequences of affective responses for participation and learning in science, we attend to interactions among affective displays and the ways in which students identify with and understand the epistemologies of the disciplines. We present an episode from an interview conducted with “Gavin,” a case-study student in an IPLS course. We examine how aspects of Gavin's identity interact with his epistemological orientation toward physics and biology to generate both positive and negative emotion and unpack the *source* of his affect in these moments. In turn, we consider how this affective response influences and is shaped by epistemological views of coherence in the natural world.

Theoretical and Methodological Approach

In our analysis we attend to an instance in which positive and negative affect are coupled to sources of frustration and resolution. The unit of our analysis is not the individual but rather a series of moments in the context of this interview. We do not claim, for example, that Gavin *always* exhibits these particular emotions for the reasons illustrated by the episode we have chosen. Rather, the episode serves to highlight how Gavin's epistemic resources and ways of positioning his identity were coordinated in these moments. For reasons we describe below, these ways of coordinating epistemology and identity may influence Gavin in the future.

This framework views neither identity, nor epistemology, nor affect as stable entities that an individual carries with him from moment to moment. Instead, each of these dimensions is influenced by the different contexts in which an individual participates (Gupta et al., 2010; Hammer & Elby, 2002; Nasir & Saxe, 2003). This framework does not preclude the possibility that some of these constructs may be more or less consistently activated across a variety of contexts. It simply starts from the assumption that these constructs are sensitive to context and leaves the determination of whether they are more or less stable across context to empirical investigation.

Our focus is on the ways in which an interaction between *identity* and *disciplinary epistemology* is responsible for Gavin's *disciplinary affect* in an interview about his experiences in this IPLS course. Disciplinary epistemology here refers to ways of knowing and learning associated with a particular discipline (Hammer & Elby, 2003). For students, disciplinary epistemologies are likely to be closely tied to their course experiences (Watkins & Elby, 2013). For example, a student might develop an understanding of biology as "complex and difficult to model in a simple way," or of physics as "abstract and idealized" from his biology and physics coursework respectively. Different course experiences could contribute to the development of different sets of epistemological resources. That same student might develop an understanding of biology as elegant and mathematical or physics as uncertain and messy from another set of course experiences.

Identity, like epistemology, is dynamic and context dependent. As Esmonde (2009) writes, "identities may shift in meaning or salience as one moves from one context to the next." The way a student positions herself relative to a discipline can also vary from moment to moment in more or less consistent ways. A student may over time begin to define herself as "a biology person," but may in other moments feel alienated from or excluded from that discipline (Nasir & Saxe, 2003), particularly in comparison to experts. Another student may identify as "someone for whom physics is really difficult," but may, at times, position herself as more aligned with the discipline. There may be ways of identifying that are broader than any particular discipline, but that interact with the individual disciplines in influential ways, such as one's view of himself as "a hands-on kind of person" or "a person who is good with symbols."

Epistemology, identity, and affect are related in various and nuanced ways. Danielak et al. (2013) describe how identity and personal epistemology are coupled for a student who identifies as a certain kind of knower. Disciplinary identity and disciplinary epistemology sometimes evolve together, as when researchers come to position themselves as aligned with or distant from particular disciplinary practices (Osbeck & Nersessian, 2010). Affect, in turn, can stabilize or destabilize epistemic orientations and aspects of identity or may result from a match or mismatch between them. Someone might feel anxious in a context in which his enacted identity is not valued, or proud in a context where it is. At times these constructs of identity, epistemology, and affect reinforce each other, and at times they are in tension. This is particularly true in an interdisciplinary setting in which more than one set of disciplinary identities, epistemologies, and affective responses may be at work.

Gavin's Story: Satisfaction from a Mechanistic Explanation of Diffusion

One goal of our IPLS course is to unpack the physical mechanisms underlying biological phenomena that are only described phenomenologically in typical introductory biology and chemistry courses. An example of this is the diffusion of particles or gases along a concentration gradient, a phenomenon with which life science students become familiar but for which they are often not provided a mechanistic explanation in their introductory courses (Redish & Cooke, 2013).

Gavin finds the unpacking of diffusion in mechanistic terms to be highly satisfying and references the example in describing the role this interdisciplinary course plays in his education more generally:

- [1] Gavin: This [IPLS] class was very good about telling us about thermodynamics and entropy's role in the universe... And I think diffusion was when everything started to click; when we talked about how molecules go from higher concentration to lower concentration because they're bumping into each other so much, and so these Newtonian interactions were able to move particles away from one another... there was less collisions and stuff like that... And so I felt like that's when things started to click [*snaps*]

fingers]... I was like that's why molecules go from higher concentration to lower concentration...

- [2] Interviewer: So you already knew that it happened?
- [3] Gavin: I knew that it happened but then I was like how the hell do they know where the lower concentration is?! And in biology we never explain that [*brushes arm across his chest*]. And I think that biology has done obviously very brilliant things and I love biology, but as far as the professors, they're very knowledgeable but they have to go over so much stuff that they don't really take time to explain why things happen. And I'm a very "why" kind of person; I want to understand why does this happen. And that's why I struggle with [organic chemistry] so much, because it's like 'memorize the mechanisms and take the test' [*throws up his hands*]... well how the hell do I know why the mechanism is happening in the first place?!
- [4] Interviewer: How do the molecules know what to do...
- [5] Gavin: Exactly. And why do they do this bouncing thing [*moves hands back and forth*] and it was never explained to me very well, and then when I take this [IPLS] class and understand, oh, this is why molecules interact the way they do, this is why you are going to have this expansion of particles over space.
- [6] Interviewer: Yeah
- [7] Gavin: It's because they collide less often when they're further apart than when they're together. And they are going to want the least colliding orientation which is going have the most microstates which is therefore going to have the greatest entropy.
- [8] Interviewer: So it connected... you knew that it wanted the greatest entropy, and it connected sort of underneath it what was causing?
- [9] Gavin: Right it gave me a foundation...
- [10] Interviewer: And that was satisfying to you?
- [11] Gavin: That was very satisfying... understanding the why really gave me the confidence in order to go into tests and be able to rationalize why things work the way they do and what to look for.

In turn [3] of this exchange, Gavin reflects on his experience in prior biology (and organic chemistry) courses. He diagnoses biology as descriptive and fact-driven ("they have to go over so much stuff"), as placing too great of an emphasis on memorization of factual information and too little of an emphasis on the explication of "why things happen." Gavin's epistemological orientation toward biology in this moment is one in which he sees the discipline as failing to take up mechanistic explanations of the sort that his IPLS course provides for diffusion (or perhaps even failing to ask questions for which a mechanistic answer is appropriate). Gavin's reflection on his experiences in biology is accompanied by markers of frustration. He is exasperated that his biology instructors "don't really take time to explain why things happen" and that he is asked to "memorize the mechanism and take the test." In this moment both his words and hand gestures convey frustration. We describe Gavin's frustration as stemming from a disconnect between his identity as a sense-maker – "I am a why kind of person" – and what he finds to be an unsatisfying preoccupation with *knowing* (as opposed to *explaining*) in biology (left side of Figure 1).

Whereas Gavin's epistemological orientation toward biology is in tension with his identification as a "why kind of person," the epistemological view of physics that he articulates in this moment aligns with that identification. Gavin describes physics as a discipline where mechanistic sense-making is commonplace. In turn [1] he credits physics as a place where he came to understand "why reactions proceed the way they do," and in turn [5] as the place where he finally came to understand "why molecules interact the way they do... why you are going to have the expansion of particles over space." In turn [9] he labels the explanatory base he feels he acquired in physics as a "foundation." These descriptions of his epistemological orientation toward physics are accompanied by markers of positive emotion and excitement. He describes his IPLS course in turn [1] as the place where "everything started to click," and in turn [11] agrees with the interviewer that the conceptual foundation that he feels he established is "very satisfying." Where Gavin's epistemological orientation toward biology is in tension with his identification as a "why kind of person," his epistemological view of physics as a place where sense-making happens *aligns* with this identity in such a way that his affective response to physics is notably more positive (right side of Figure 1). Gavin also attributes his greater comfort on tests to improved facility with mechanistic explanation of the sort emphasized in his IPLS course, saying that such an understanding provides "the confidence in order to go into tests and be able to rationalize why things work the way they do."

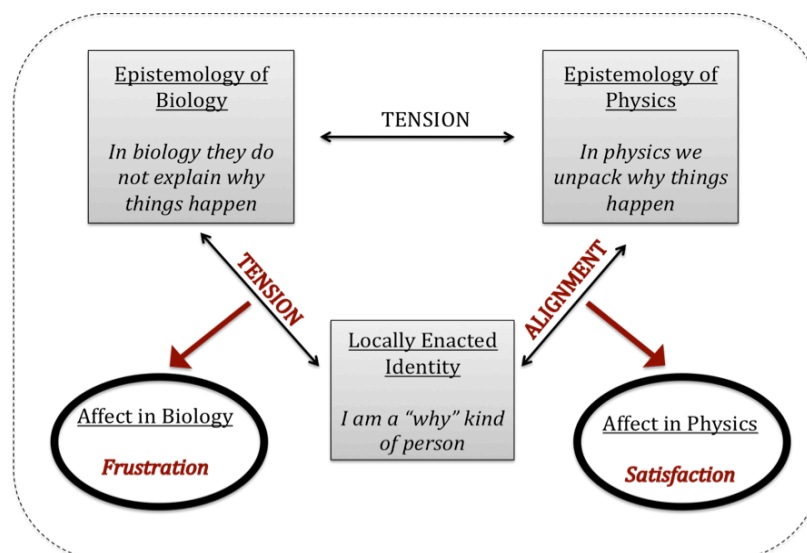


Figure 1. Modeling the sources of Gavin's affect during an interview in IPLS.

Implications for Interdisciplinarity

The frustration stemming from tension between Gavin's self-identification as a "why kind of person" and his epistemological view of biology is coupled to the satisfaction that Gavin achieves in his IPLS course. While we do not know if Gavin would or would not have appreciated physics in its own right had he not first encountered phenomena in biology for which he desired further explanation, his sense of resolution in IPLS *can* be attributed in part to his dissatisfaction with explanations in biology. Gavin positions his satisfaction with the role that physics is playing in his understanding of natural phenomena in direct *comparison with* and in direct *contrast to* his dissatisfaction with the incompleteness of explanations in biology.

Similarly, Gavin's frustration with the lack of attention to mechanistic explanation in his biology courses might actually serve to *strengthen* his ultimate sense of interdisciplinary connectedness and satisfaction. Because he sees *physics* class as a place where he was encouraged to develop explanations, Gavin may actually be more likely in the future to view physics as relevant and important for understanding the living world. The frustration that Gavin feels in association with biology may not only be productive in the sense that it enables him to more fully appreciate and experience the power of mechanistic explanation when he does eventually encounter it, but also in the sense that it allows him to appreciate a role that physics can play in the life sciences. A student who is never troubled by a lack of mechanistic reasoning in biology (or a student who sees biology as descriptive but actually *likes* that aspect of the discipline) might see superficial connections between biology and physics in an IPLS setting, but that student is less likely to see physical models as *essential* for answering some interesting questions in biology.

Figure 1 represents disciplinary affect as an *outcome* of the interaction (either tension or alignment) between identity and disciplinary epistemology. It is also plausible, however, that the alignment between Gavin's identification as a why kind of person and his disciplinary epistemology is *stabilized* by the positive emotions resulting from that alignment. The feelings may reinforce Gavin's belief that he is a sense-maker who values mechanistic explanation, and it is not unreasonable to predict that he may seek out opportunities in his future courses to do more of that sense making. We would hope, for example, that Gavin might begin to look for opportunities to make sense of biological phenomena in mechanistic ways that are authentic to the discipline of biology. Alternatively, if Gavin does not have this kind of opportunity in his subsequent life science courses, he might consider leaving biology for a field that he views as more conducive to mechanistic sense making (Danielak et al., 2013). By way of comparison, an IPLS student who experiences neither Gavin's frustration with biology nor his satisfaction in having explained something important in his physics course may be less likely to seek out connections between physics and biology beyond the confines of the IPLS environment.

In a course that has as one of its goals the dismantling of disciplinary silos, positive affect associated with the role that physics can play in unpacking biological phenomena is of particular importance. Future work is required to establish whether such affect indeed does stabilize a student's orientation toward interdisciplinary sense making, and to determine if such affect makes it more likely that that student will cross disciplinary boundaries in the future. But the positive affect is also an end *in and of itself*. Many prominent scientists have attributed their motivation to participate in science to those rare but powerful feelings of satisfaction, pleasure, and beauty that accompany the successful reconciliation of various pieces of conceptual understanding. It is possible for our students to experience similar satisfaction. When defining what it is that we hope our students will learn in interdisciplinary courses, we would be well served to consider also what we hope they might *feel*.

References

- Danielak, B. A., Gupta, A., & Elby, A. (2013). The marginalized identities of sense-makers: reframing engineering student retention, 1–76. Submitted for publication.
- Esmonde, I. (2009). Ideas and entities: Supporting equity in cooperative mathematics learning. *Review of Educational Research*, 79, 1008-1043.
- Gupta, A., Danielak, B. A., & Elby, A. (2010). Understanding students' difficulties in terms of coupled epistemological and affective dynamics. Proceedings of the 2010 IEEE Frontiers in Education Conference, Red Hook, NY: Curran Assoc. Retrieved from <http://arxiv.org/abs/1003.4238>
- Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. Hofer & P. R. Pintrich (Eds.), *The Psychology of Beliefs and Knowledge about Knowing*. Mahwah, NJ: Erlbaum.
- Hammer, D., & Elby, A. (2003). Tapping epistemological resources for learning physics. *The Journal of the Learning Sciences*, 12(1), 53–90.
- Hammer, D., Russ, R., Mikeska, J., & Scherr, R. (2005). Identifying inquiry and conceptualizing students' abilities. In Duschl, R. & Grandy, R. (Eds.), *Teaching scientific inquiry*. Rotterdam, Netherlands: Sense Publishers.
- Meredith, D. C., & Redish, E. F. (2013) Reinventing physics for life science majors. *Physics Today*, 66:7, 38-43.
- Nasir, N. S., & Saxe, G. B. (2003). Ethnic and academic identities: A cultural practice perspective on emerging tensions and their management in the lives of minority students. *Educational Researcher*, 32(5), 14–18.
- National Research Council (US). (2003). Bio 2010: Transforming Undergraduate Education for Future Research Biologists. National Academy Press.
- Osbeck, L. M., & Nersessian, N. J. (2010). Forms of Positioning in Interdisciplinary Science Practice and Their Epistemic Effects. *Journal for the Theory of Social Behaviour*, 40(2), 136–161.
- Redish, E. F. & Cooke, T. J. (2013). Learning each other's ropes: Negotiating interdisciplinary authenticity. *Cell Biology Education – Life Science Education*, 12, 175-186. doi:10.1187/cbe.12-09-0147.
- Watkins, J., & Elby, A. (2013). Context Dependence of Students' Views about the Role of Equations in Understanding Biology. *CBE-Life Sciences Education*, 12(2), 274–286.

Acknowledgments

We appreciate the feedback from the UMD Physics and Biology Education Research Groups and are grateful for the support of Joe Redish, and for substantive discussions with Ben Dreyfus. This work was supported by NSF-TUES DUE 11-22818 and the HHMI NEXUS grant.

The Impact of Principle-Based Reasoning on Hands-on, Project-Based Learning

Marcelo Worsley and Paulo Blikstein, Stanford University, 520 Galvez Mall, CERAS 102, Stanford, CA 94305
mworsley@stanford.edu, paulob@stanford.edu

Abstract: Prior research suggests that experts and novices employ markedly different approaches to engineering design tasks. For example, novice designers commonly use trial and error, which researchers liken to backward-reasoning. Experts use forward-reasoning, which allows them to accurately predict the impact of certain decisions. In this paper, we present a complementary conceptualization for how experience affects design approaches. We liken backward-reasoning to example-based reasoning, and forward-reasoning to principle-based reasoning. In study 1 (N=13) students complete an engineering design activity. A qualitative analysis shows clear instances of example- and principle-based reasoning strategies. Study 2 (N=20) compares the efficacy of the two approaches by using a between subject design. We find that principle-based reasoning improves the quality of designs ($p < 0.01$) and learning of important engineering principles ($p < 0.001$). This suggests that hands-on learning environments may benefit from encouraging students to employ principle-based reasoning.

Introduction

With the rise of Makerspaces and Fablabs, constructionist (Papert 1980) and project-based learning is experiencing a revival. Schools, community centers and museums around the world are spending millions of dollars to create spaces where students can engage in design and invention activities (New York Hall of Science 2012, Blikstein 2013). Amidst this resurgence, however, are questions about how to effectively implement hands-on learning in a way that foments Science, Technology, Engineering and Mathematics (STEM) learning. For museums, they must address questions about how to promote changes in students STEM intuitions through short-lived constructionist activities (see “Keychain Syndrome”, Blikstein 2013 for an example). At the same time, schools have to consider how to produce measurable learning gains through engaging, hands-on activities. In this paper, we begin to address these concerns through two studies. In Study 1 we describe two common approaches to solving open-ended engineering tasks. We refer to these approaches as *example-* and *principle-based reasoning*. These two approaches share resemblance to prior literature on learning by analogy and expert/novice design strategies (e.g. Ahmed Wallace and Blessing 2003, Gentner 1997, Cardonell 1983). Study 2 shows that principle-based reasoning produces superior designs and greater recognition of engineering principles among non-expert students relative to example-based reasoning.

To situate the reader this paper draws from studies of engineering education (e.g. Atman and Bursic 1998, Russ et al 2008, Lau, Oehlberg and Agogino 2009, Atman and Bursic 2007, Worsley and Blikstein 2011), the role of analogy in supporting learning (e.g. Gentner and Holyoke 1997, Cardonell 1983, Lakoff and Nunez 2000, Polya 1945) and prior work on expertise (e.g. Ericsson, Krampe, & Tesch-Römer, 1993, Chi, Glaser and Rees 1981, Cross and Cross 1998, Atman et al. 1999).

Study 1

Study 1 examines engineering design approaches of students with different levels of experience. Students are presented with a mechanical engineering challenge that requires them to use their intuitions and prior knowledge about forces.

Ten 9th- through 12th-grade students and three graduate students participated in this study. This variance in experience was a primary component in pre-defining student experience levels, i.e. low, medium, high and expert. A detailed description of the user population can be found in Worsley and Blikstein (2013).

Students worked individually using common household materials: one paper plate, one ping-pong ball, a roll of tape, four drinking straws and five wooden Popsicle sticks. The objective was to use the materials provided to create a structure that could support approximately 2-3 pounds. Participants were also asked to support the weight as high off the table as possible. Students were instructed that they would receive ten minutes to complete the task, but were permitted to work for as long as they wanted. Total participation time ranged from eight minutes to fifty-two minutes.

Engineering Strategies

As we analyzed each video we paid particular attention to the post-activity interviews. During this interview students were asked questions about the motivation for their design. We were intrigued by how some students' designs were motivated by of real-world objects, while others relied on geometric properties and principles from

engineering. As such our goal in this section is to briefly demonstrate the existence of the example-based and the principle-based approaches. To be clear, these two approaches do not represent all strategies that students used, but constitutes the majority. In what follows we present short case studies of each approach.

Example-Based Reasoning

The first example that we present is of a student who successfully modeled their structure after a chair (Figure 1). More importantly, though, we learn that this was not modeled after just any chair, but after a chair that the student has at home. The student emphasizes this point when responding to a question about the inspiration for the design: “It’s like a form of a chair... Plus. Um. Just like a chair, because I have a chair like that... at home.”

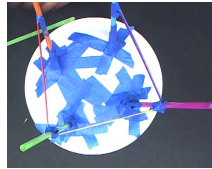


Figure 1. Successful Example-based Design

This student has drawn a connection to a real-world object from his home as a way for approaching the task. As the student continues to describe the motivation for the design, he indicates that he had briefly entertained another idea. This other idea very closely resembled his current structure but did not have wooden sticks connecting the legs. When asked why he didn’t pursue the other design he says that the other idea “was dumb.” While he may have been hinting at principles in engineering design, he does not articulate this point. As we examine the line of reasoning offered by students who employed principle-based reasoning, we will see an apparent difference in the extent of mechanistic reasoning that the students employ (Russ et al. 2008).

Principle-Based Reasoning

Among students that exhibited principle-based reasoning it was common to find structures laden with triangles and circles. This was the case for structure pictured in **Error! Reference source not found.** and Figure 5.

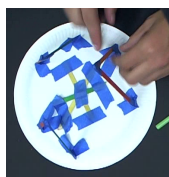


Figure 2. Successful principle-based design

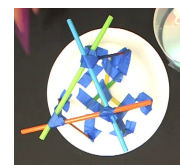


Figure 3. Successful principle-based design

Figures 2 and 3 contain the underside of his structure. In Figure 2 we see the early stages of the base. This base features two levels of triangles. The first is the shape of each leg. The second is the triangular base that the three legs define. Figure 3 makes the second level of triangles more explicit with the addition of three straws that form a triangle. When asked what inspired the design the student responded, “Well triangles are strong. And so, I decided to use as many triangles as I could.” Upon further probing about the importance of triangles the student offered the following explanation: “It’s the most secure shape because, uhh, none of the angles can change once you have three sides in place. Whereas a lot of other shapes, they can tilt around and change.” The student is very confident in his reasoning, and provides a strong justification for his design. Furthermore, in contrast to the example-based designs, this structure bears little resemblance to any real-world object.

Discussion

Study 1 was designed to identify the ways that students would naturally go about handling an open-ended engineering design challenge. It offers a glimpse into the types of experiences that students may have when approaching unscripted, hands-on learning activities. Examining these practices is of particular relevance given the rise of making that is taking place in Maker Spaces, museums, FabLabs and schools. Within these constructionist environments, some students attain success by drawing on real-world examples (Gentner 1997, Cardonell 1983, Kolodner 1997). Identifying real-world examples gives students an entry point into constructing their design. It also challenges students to consider ways for generalizing the objectives of the challenge to structures that are salient in their individual lives. Other students approach the task from the perspective of engineering principles. When both designing and troubleshooting their structure, they draw upon engineering principles to guide their thinking and their actions. Like the example-based approach, the principle-based approach yields mixed results. Nonetheless, these two approaches appear to align with prior literature on expertise in open-ended design activities (Ahmed, Wallace and Blessing 2003). The example-based approach is

akin to backward reasoning, where the individual begins with a structure and works backwards to deduce its feasibility and quality. The principle-based approach is in line with forward reasoning - students identify the interworkings of a design before concerning themselves with their larger, overall design. While it was not our objective to align these practices with levels of expertise, this is something that loosely emerged from the qualitative observations. This perceived difference in who is using the different techniques motivates the design of Study 2.

Study 2

In Study 2 we prompt participants to use either an example- or principle-based approach when completing the challenge. This allows us to examine the impact of the two reasoning strategies on student success and on student learning. Study 2 also differs from Study 1 because participants work in pairs, instead of individually. This change in design was adopted to foster more natural verbalization of one's thoughts and strategies.

Prior to the study, we had two main hypotheses. The first hypothesis was that principle-based reasoning prompt would cause students to be more aware of the mechanisms that conferred stability. Moreover, as we observed in Study 1, we anticipated that when students focused on mechanisms, they would be more cognizant of weaknesses in their design. This, in turn, would result in the principle-based reasoning condition being more successful in building their structures. At the same time, we hypothesized that the example-based group would mirror the novices in Study 1 and create structures that resembled real-world structures. However, in so doing, we expect for many students to overlook one or more important engineering principles because they are primarily thinking at the macroscopic level.

Twelve 9th- through 12th-grade students and eight undergraduate students participated in this study. Pairs of students were randomly assigned to each condition, after controlling for prior education experience.

Students worked with common household materials: one paper plate, 4 ft. of garden wire, four drinking straws and five wooden Popsicle sticks. Students were also given scissors. The objective was to use the materials provided to create a structure that could support a weight of approximately half a pound. Participants were also asked to support the weight as high off the table as possible. Students were instructed that they would receive fifteen minutes to complete the task, but were permitted to work for as long as they wanted. Changes from the Study 1 protocol were adopted to avoid some of the superficial trouble shooting strategies observed in study one. For example, tape was taken away because during Study 1 several students used excessive amounts of tape in order to compensate for unstable structural designs.

Activity Sequence

The set of activities that students completed includes: a pre-test; introducing the design challenge; an intervention, i.e. one of the two conditions; a preliminary design drawing; a hands-on, paired, building activity; a post-test; and reflection. For both the pre- and post-test, students were asked to generate as many ways as possible, to easily reinforce an unstable structure. For the post-test, students were permitted to both review and refer back to their pre-test responses.

During the intervention, students were first shown a picture of a bridge, a ladder and an igloo. In the example-based condition students were asked to generate three ideas of relevant structures from their home, community or school that would be useful in thinking about completing the current task. Students received three minutes to generate and draw three ideas. In the principle-based condition students were asked to generate three mechanisms, or engineering principles, that cause one or more of the three items pictured (the bridge, the ladder and the igloo) to be structurally sound. Again, students were given three minutes to generate three mechanisms.

One key observation from Study 1 was the failure to recognize important engineering principles, or mechanisms, among the example-based reasoning group. Most common was for students to overlook the importance of connecting the legs of the structure, especially when using slanted legs. In Study 2 we adopted a coding scheme that explicitly classifies pre- and post- tests based on the presence of connected legs and slanted legs. In order for an item to be coded as having connected legs, there must be something that connects the legs at some place other than the top or bottom of the structure. Identifying the presence of slanted legs was based on the presence of a non-90-degree angle between the legs of the structure and the upper portion of the structure. When coding student pre- and post-tests, the presence of slanted legs resulted in a scoring of -1, and the presence of connected legs resulted in a scoring of +1. Accordingly a given test could be coded using three possible values: -1, 0 or 1. Each design drawing was coded by two research assistants. An analysis of inter-rater reliability analysis yielded a Fleiss' Kappa of 0.76. In addition to coding the pre- and post-test, an explicit metric of success was used to rate the quality of each structure. In order to be deemed a success, the structure was required to hold the half-pound mass for at least one minute.

Results

We first examine the correlation between success and the two conditions. A binomial test (probability of success = 0.1 – based on prior work) confirms that the principle-based condition (M: 0.6, SD: 0.52) significantly

outperforms the baseline probability ($p < 0.01$), while the example-based reasoning condition ($M:0.2$, $SD: 0.42$) does not outperform the baseline probability ($p \sim 0.40$).

The other value that we compared is coded values for slanted legs and connected legs. When we perform a student t-test on this data we find a significant difference ($t(18)=3.46$, $p < 0.003$). Students in the principle-based reasoning condition were more likely to include only connected legs, while the example-based condition was more likely to only include slanted legs. This is important because the slanted leg configuration in the absence of connected legs will typically result in failure. On the other hand, using connecting legs has utility across several leg configurations and orientations. This is significant because it means that after doing the hands-on activity students in the example-based reasoning condition were more likely to propose a design solution that included slanted legs, without connecting the legs. This suggested that they were less likely to attribute structural failures to having slanted legs or having unconnected legs, whereas the principle-based reasoning group was more likely to correctly make this association.

One concern is that this observation is a function of the student's prior knowledge of structural engineering. Accordingly, we coded the pre-tests as well as the initial design drawings (recall that each student completed a design drawing immediately after having finished the intervention). From both of these, we find no significant difference in the averaged coded value for slanted legs-connected legs either immediately before the intervention or immediately after the intervention (pre-test - $p < 0.23$, initial design drawing - $p < 0.13$). Of particular note is that connecting the legs was more common in the example-based reasoning group ($p=0.6$) than in the principle-based reasoning group ($p=0.4$) when comparing design drawings made immediately following the intervention and preceding the building activity. This suggests that the observed difference on the post-test is not a function of the intervention alone, nor is it a function of prior knowledge, but is mediated through the hands-on activity.

Discussion

Through studies 1 and 2, we highlighted the existence and importance of different approaches for completing engineering design tasks. We found that students who employ principle-based reasoning were more successful in completing the task, and also identified important mechanisms in the post-test. On the other hand, students who used example-based reasoning were less likely to succeed in their designs. We attribute these differences to students in the example-based reasoning condition overlooking important engineering principles that were integral to the stability of their example structure. As an example of this, consider students Mike and Tom. These two students modeled their structure after a water tower. They include several of components of a water tower in their structure. However, their structure failed because they do not recognize the importance of connecting the legs of their structure together. Ironically, these two included connected legs in their design drawings, but still did not recognize their importance. Unfortunately, we do not have sufficient post-interview data to confirm that all students in the principle-based reasoning condition that included connected legs fully understood the implications of including reinforcements or braces. Anecdotally, we observed that a number of students in the principle-based reasoning condition made reference to connecting the legs during the activity, but because this was not explicitly tested for all participants this inference is only speculative.

As we consider next steps one area for further research is to leverage an approach from our prior research that more explicitly studies user action patterns (Worsley and Blikstein, 2013). Within this paradigm, we would be interested in studying how the process differed between the two conditions. As a part of this analysis, we would look to identify if the example- and principle-based reasoning cause students to use substantively different iteration cycles and troubleshooting strategies. Moreover, several previous research studies have highlighted differences in how experts and novices complete engineering design tasks. While we have seen that principle-based reasoning helps students be more successful, there is the additional question of whether or not this approach actually helps students behave more like experts when examining more fine-grained action sequence. Related to this is another possible study that tries to more definitively delineate the mechanics that mediate success for the principle-based reasoning group, again through process-based measures.

Conclusion

In this paper we presented a pair of studies that investigate how to improve the quality of hands-on learning experiences. In Study 1 we provided a short qualitative analysis of two engineering design approaches: example- and principle-based reasoning. We argued that these two approaches are in line with prior research in engineering education, expertise and learning by analogy. Moreover, we described how example-based reasoning is commonly used among novices as they confront new engineering design tasks. Study 2 was a controlled study where we compared the efficacy of the two approaches in terms of the quality of designs, and in terms of what students learn. Here we found that principle-based reasoning is associated with better quality designs, and better recognition of important engineering mechanisms. As described in this paper, priming students to use principle-based reasoning could conceivably be enacted in the variety of constructionist learning environments that are currently found in community centers, museums and schools. As such, adopting this

approach could foster significant improvement to the quality of constructionist learning. More generally though, this paper echoes previous research that highlights the importance of pushing students to draw on the depth of their conceptual intuitions. While some prior work has cautioned that students' intuitions may contain non-scientifically valid principles and concepts, in the case of this study, this concern was far less impactful than the positive benefits conferred by promoting principle-based reasoning.

References

- New York Hall of Science. (2012). *Design-Make-Play: Growing the Next Generation of Science Innovators*. New York, NY: New York Hall of Science.
- Ahmed, S., Wallace, K. M., & Blessing, L. T. (2003). Understanding the differences between how novice and experienced designers approach design tasks. *Research in Engineering Design*, 14(1), 1-11.
- Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. *Journal of Engineering Education*, 96(4), 359-379.
- Atman, C. J., & Bursic, K. M. (1998). Verbal protocol analysis as a method to document engineering student design processes. *Journal of Engineering Education*, 87(2), 121-132.
- Atman, C. J., Chimka, J. R., Bursic, K. M., & Nachtmann, H. L. (1999). A comparison of freshman and senior engineering design processes. *Design Studies*, 20(2), 131-152.
- Atman, C. J., Deibel, K., & Borgford-Parnell, J. (2009). The process of engineering design: A comparison of three representations. *International Conference on Engineering Design: Stanford University*.
- Blikstein, P. (2013). Digital Fabrication and 'Making' in Education: The Democratization of Invention. In J. Walter-Herrmann & C. Büchling (Eds.), *FabLabs: Of Machines, Makers and Inventors*. Bielefeld: Transcript Publishers.
- Carbonell, J. G. (1983). Learning by analogy: Formulating and generalizing plans from past experience (pp. 137-161). Springer Berlin Heidelberg.
- Chi, M. T., Glaser, R., & Rees, E. (1981). Expertise in problem solving. *Learning Research and Development Center, University of Pittsburgh Pittsburgh*.
- Cross, N., & Cross, A. C. (1998). Expertise in engineering design. *Research in Engineering Design*, 10(3), 141-149.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, 100(3), 363.
- Gentner, D., & Holyoak, K. J. (1997). Reasoning and learning by analogy: Introduction. *American Psychologist*, 52(1), 32.
- Guzdial, M. (2008). Education paving the way for computational thinking. *Communications of the ACM*, 51(8), 25-27.
- Kolodner, J. L. (1997). Educational implications of analogy: A view from case-based reasoning. *American psychologist*, 52(1), 57.
- Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. Basic books.
- Lau, K., Oehlberg, L., & Agogino, A. (2009). Sketching in design journals: An analysis of visual representations in the product design process. *Engineering Design Graphics Journal*, 73(3)
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc.
- Polya, G. (1945). *How to solve it: A new aspect of mathematical model*.
- 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.
- Worsley, M., & Blikstein, P. (2011). What's an expert? Using learning analytics to identify emergent markers of expertise through automated speech, sentiment and sketch analysis. *EDM*, 235-240.
- Worsley, M. and Blikstein, P. (2013). Toward the Development of Multimodal Action Based Assessment. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK '13)*. ACM, New York, NY, USA, 94-101.

Emotional Engagement in Agentive Science Learning Environments

Andrew Morozov, Leslie Rupert Herrenkohl, Kari Shutt, Phonraphee Thummaphan, Nancy Vye, Robert Abbott, and Giovanna Scalone, University of Washington, Seattle, WA, USA
Email: {amorozov, leslih, shuttk, phonrt, nancyvye, abbottr, gscalone} @u.washington.edu

Abstract: As students engage in ambitious intellectual activities, they inevitably manage social interactions and emotional responses in the classroom. In this research, we studied fifth graders' emotional engagement in elementary science classes redesigned to offer more ambitious, agentive learning. We examined student self-reports of affective states to address the following questions: (1) Do students in "Agency" classes report more positive or negative emotional experiences, compared to students in classes that use traditional, kit-based science units? (2) To what extent do students' emotional experiences vary as a function of class? Findings indicate that, after taking classroom differences into account, students in Agency classes generally felt more positive about their experiences in science than students in kit-based classes. Next steps include exploring differences across classrooms and synthesizing research on learning, affect and identity.

Background

Since 2006, we have been engaged in design-based implementation research (Penuel, Fishman, Cheng & Sabelli, 2011) on a model of "agentive" science learning to advance students' science inquiry learning, social-emotional outcomes, and practice-linked identities. Our work involves adapting hands-on science kits in ways that offer greater student agency: Students are provided with the opportunity to direct their own inquiry on a complex problem over a sustained period of time, thereby broadening the range of investigative practices in which they engage, helping build knowledge by connecting experiences across time and place, in and out of school (NRC, 2009) and offering opportunities for original thought (Metz, 2000; Bereiter & Scardamalia, 2009; Herrenkohl & Mertl, 2011). Per Engle and Conant (2002), we encourage students "to be authors and producers of knowledge with ownership over it rather than the mere consumers of it, and [encourage] teachers and other members of the learning community to position students as stakeholders by publicly identifying them with the claims, approaches, explanations, designs, and other responses to problems that they pursue" (p. 404).

Our redesigned science kits are challenge-based (Schwartz, Lin, Brophy, & Bransford, 1999): Each redesigned unit 1) has an overarching challenge (problem) that connects unit lessons; 2) links to students out-of-school science practices and experiences using student self-documentation; 3) offers sustained, in-depth study of science concepts/practices in the context of the overarching problem; 4) involves student-guided inquiry (with teacher scaffolding), and; 5) culminates in a public performance and "call to action" on a local science-related issue. The investigative practices that we embed in these units encompass many of the NGSS practices; for example, students generate their own investigative questions, design investigations to address these questions, engage in argumentation from evidence, construct explanations and used data from investigations to solve problems.

Theoretical Framework and Significance of the Work

Research suggests that providing students with greater agency to wrestle with science questions and problems can benefit learning (e.g., NRC, 2000). Additionally, metacognitive benefits may accrue as students monitor and reflect on their knowledge and progress in the context of solving problems about which they care (Brown, 1978; Herrenkohl, et al., 1999), and if learners are appropriately scaffolded in their efforts (White & Frederiksen, 1998).

Emotional Engagement in Learning

In addition to learning and metacognition, we speculate that there may also be important social-emotional benefits of agentive learning. As students engage in these ambitious intellectual activities, they inevitably manage social interactions and emotional responses in the classroom. Students may develop skills such as motivation, grit and persistence, adaptability and communication, self-concept and identification as science learners, social belonging, and self-regulation together with their conceptual and epistemological skills. Recent economic studies suggest that these social-emotional skills play an important role in long-term social and economic success related to schooling and career choices, employment, and wages (e.g., Heckman, Stixrud, & Krzua, 2006).

Several learning sciences researchers have begun over the past decade to theorize about the connections between social emotional constructs, such as identity, and learning (Herrenkohl & Mertl, 2011; Nasir, Lee,

Roseberry & Warren, 2006; Engle et al, 2008; Esmonde & Langer-Osuna, 2007; Gresalfi, 2009). Some have coined terms like “intellective identities” (Greeno, 2002) to emphasize how students’ sense of competence intersects with their conceptual and epistemological abilities.

Building on interpretations of Spiro (1982) and Dreyfus (1984), Holland{ XE "Holland" }, Lachicotte{ XE "Lachicotte" }, Skinner, & Cain (1998) examined how expertise and interest develops together in a non-scientific domain. Holland et al.’s study of young women’s approaches to romance demonstrated that “salience, identification, and savoir faire appeared to develop together in an interrelated process – a process that was continually supported and shaped in the context of social interaction{ XE "social interaction" }” (p.116). In other words, as women developed greater competence in navigating the romantic scene, their evaluation of the importance of romantic pursuits, their emotional{ XE "emotional" } investment in this activity, and their desire to continue intensified. Holland and her colleagues suggest that positive emotional engagement and expertise function as a system, with one supporting the development of the other. This theoretical stance suggests that the development of expertise in scientific inquiry is inextricably linked to the development of interest and the management of positive emotional responses that support students to seek out further opportunities to engage in scientific practices.

Although these aspects of interest and emotional response are theoretically supported, there is very little empirical data available to examine these relationships. In the present study the relationship of student agency and affect in science is explored. This paper derives from a larger comparative research study conducted in the 2012-13 school year in which three 5th grade Full Option Science System (FOSS) science units (representing a school year long intervention) were redesigned to offer students greater agency for their learning. Outcomes, in the present case affective outcomes, were then compared to outcomes for students participating in the original FOSS units. The study addresses the following two questions: (1) Do students in the redesigned Agency unit report more positive or negative affect, compared to students taking the FOSS unit? (2) To what extent do these emotional experiences vary with the units if classroom effects are taken into account?

Method

Participants

Participants were from an urban/suburban school district in the Pacific Northwest that serves approximately 18,000 students. We collaborated with teachers from several elementary schools in the district to study the effects of the redesigned science units (Agency units), as compared to FOSS’ kit-based units.

The participants in this study were 180 fifth grade students from 5 Agency classes across 2 schools, with 21 to 24 students per class (n=113) and 3 FOSS classes across 2 other schools, with 21 to 24 students per class, (n=67). The percentages of Free and Reduced Lunch students were 44% and 25% for the schools with the Agency classes, and 37% and 46% for the schools with the FOSS classes. The percentages of students receiving ELL services were 20% and 37% for schools with the Agency classes, and 16% and 29% for schools with the FOSS classes. The Agency classes were taught by two teachers, one of whom (a Science Specialist) taught four classes at one of the schools. The FOSS classes were taught by different teachers.

Measures

To measure students’ affect related to their experiences in science, students were periodically asked to complete an Exit Card (EC) at the end of science class. It contained 10 adjectives: Excited, Frustrated, Challenged in a Good Way, Interested, Bored, Confused, Confident, Helpful, Like a Scientist and Creative. Students were asked to circle all adjectives that described *how they felt today in science class*. The EC took students 1-2 minutes to complete.

Procedure

The findings reported here are based on data collected in January-March 2013 comparing a traditional, kit-based, 5th grade FOSS *Landforms* earth science unit, and one of our redesigned units. In the redesigned version of the FOSS *Landforms* unit, My Skokomish River Challenge (MSRC), students used stream tables to study the impact of various factors on erosion. The generation of empirically testable questions was completed as a whole-class activity. This is followed by three investigation cycles in which students work collaboratively in small groups to select a question, design and conduct an experiment to answer that question, analyze data, and share their findings. Their findings inform their response to the unit challenge – determining which of three construction sites should be chosen in order to cause as few problems due to erosion as possible. Although the purpose for selecting the sites was hypothetical, the sites themselves were real places within driving distance of the schools.

Agency students completed Exit Cards 7 times distributed over the course of the 12-week redesigned Landforms unit; FOSS students completed Exit Cards on 6 occasions. Teachers administered the EC to students; however, students’ responses were anonymous to their teachers. Students “signed” their ECs with unique codes

that only the research team was able to link with their names/identities.

Analysis

Principal components analysis with varimax rotation was used to calculate two component scores (those with eigenvalues greater than 1) based on the exit card ratings. Hierarchical linear models (HLM) were then used to study the effect of the unit intervention on component scores based on the exit card data, since the ratings by students were nested within classrooms.

Results

Construct Development

Table 1 provides the descriptive statistics of the exit card scores for the overall study sample (n=180). The highest mean score is for the Interest item (Mean=.76), followed by Excitement and Confidence (Mean=.63 and .55 accordingly). The lowest mean scores is for Frustration (Mean=.10), followed by Confusion and Boredom (Mean=.13 and .15, accordingly). Coherence of the exit card measure set is further supported by the strength of the inter-correlations among the individual exit card measures--most variables significantly correlated at the .01 level.

Table 1: Intercorrelations between exit card measures

Exit card measures	Correlation (r)									Mean	SD
	1	2	3	4	5	6	7	8	9		
1. Excitement	-									.63	.375
2. Frustration	-.073	-								.10	.188
3. Challenge	.428**	-.052	-							.44	.364
4. Interest	.555**	-.279**	.377**	-						.76	.314
5. Boredom	-.353**	.330**	-.290**	-.452**	-					.15	.255
6. Confusion	-.068	.441**	.002	-.177**	.171*	-				.13	.234
7. Confidence	.549**	-.156*	.391**	.476**	-.312**	-.158*	-			.55	.388
8. Helping	.494**	-.048	.519**	.384**	-.220**	.027	.586**	-		.49	.372
9. Scientist	.565**	-.118	.578**	.425**	-.337**	-.041	.492**	.479**	-	.52	.389
10. Creativity	.551**	-.144*	.500**	.443**	-.256**	-.094	.509**	.590**	.636*	.51	.394

All tests are 2-tailed.

* $p < .05$, ** $p < .01$

The principal components analysis produced two components, which accounted for 58.7% of the overall variance and was felt to be conceptually appropriate. Component 1 may be conceptualized as *positive-valence feelings*, and includes the items Excited, Challenged in a Good Way, Interested, Confident, Helpful, Like a Scientist, and Creative. Component 2 may be conceptualized as *negative-valence feelings*, and includes the items Frustrated, Bored, and Confused. (Although the learning process may be accompanied by occasional frustration, confusion, or boredom, if too many students in a class feel this way too often, or to a greater extent, this might adversely impact the learning experience over time.) Since the component analysis results found 2 factors: positive and negative domains of feeling, we used each factor as an outcome variable to check whether the student-reported feelings varied significantly depending on unit as well as classroom.

Comparisons Across Units and Classrooms

Finally, HLMs were used to determine whether the differences in component scores were predicted by the unit intervention or not. For this analysis, the exit card component scores were the outcome/dependent variables, and the intervention was the independent variable in level 2. The data analysis involved first modeling the dependent variables of positive or negative feelings across classrooms; this constituted Model I. Model II comprised Model I plus the intervention type (see Table 2).

According to the results, the variance of classroom means in Model I for the positive feeling component is significant at .01 level ($\chi^2 = 22.89$, $p < .01$). This indicates that the positive feelings of students were different among classrooms, with about 9.09% of the variance explained by classroom effects. After taking the treatment (unit intervention) into considerations in Model II, the treatment effect was significant at the .05 level, indicating that the positive feelings of students were significantly higher in the Agency classrooms (Mean=.61), than in the FOSS classrooms (Mean=.47). With regard to the negative feelings component, in

Model II, the treatment effect was not significant, indicating that the negative feelings of students were not significantly different between Agency (Mean=.12) and FOSS classrooms (Mean=.14).

Table 2: Estimated effects of intervention level on positive and negative feeling

	Model I		Model II	
Positive feeling				
Fixed Effect	Coefficient	<i>t</i> -ratio	Coefficient	<i>t</i> -ratio
Intercept, γ_{00}	0.558665	15.409**	0.469465	10.513**
Intervention			0.142309	2.523*
Random Effect	Variance	χ^2	Variance	χ^2
Intercept	0.00728	22.88838**	0.00272	11.03637
Intervention	0.07277		0.07278	
Negative feeling				
Fixed Effect	Coefficient	<i>t</i> -ratio	Coefficient	<i>t</i> -ratio
Intercept, γ_{00}	0.127254	10.316**	0.143120	7.078**
Intervention			-0.025273	-0.990
Random Effect	Variance	χ^2	Variance	χ^2
Intercept	0.00000	4.40228	0.00000	3.41998
Intervention	0.02736		0.02736	

* $p < .05$, ** $p < .01$

Discussion

Conclusions

The exit card measures cluster into two components (based on their co-occurrences): positive-valence feelings (Excited, Interested, etc.) and negative-valence feelings (Frustrated, Bored, and Confused). A treatment (Agency/FOSS) effect was found for the positive feelings component, indicating that students in Agency classes had more positive emotional experiences during science instruction. In addition, although students in the Agency classes tended to report more positive feelings, there were also large classroom-level differences in positivity.

On the other hand, no significant differences between students in the Agency and FOSS classes were observed in terms of negative feelings. These findings suggest that classrooms that support student agency and scientific inquiry are more likely to create conditions for positive emotional experiences among the learners.

Findings from the exit card analysis are also supported by the analysis of interview data collected from students in the Agency and FOSS classes (Scalone, 2014). The Agency students viewed science as playful experimentation. These reasons were coupled with enjoyment, with students proclaiming that they “love” science and that “science is fun” – perceiving science as playful experimentation. In addition, Agency students indicated that they like science more than the FOSS students. Their reasons included getting to “learn new things” and “do experiments.” The Agency students also identified themselves as scientists, signaling their emotional investment in the doing of science.

Implications for Future Research

The study discussed here examines one of the five redesigned units developed for Grades 2 and 5. Further research will focus on examining students’ reports of emotional states across all redesigned and FOSS comparison units. More fine-grained analysis of the exit card data is also under way to better understand students’ specific emotional experiences with different units and in different classrooms. In addition, further analyses using the student interview data will allow us to examine the relationship between student report of emotional states and conceptual and epistemological understanding in science. We also plan to triangulate self-reported affect with videotaped episodes of classroom activities to explore the situative aspects of affective, cognitive, and behavioral engagement during collaborative scientific inquiry, including the role of teaching practices (e.g., Engle & Conant, 2002; Herrenkohl & Guerra, 1998). Ultimately this body of work can contribute to testing an empirical model that links developing interest and expertise as important parts of a system to support science learning.

Relevance to Conference Theme

The theme of ICLS this year highlights a movement over the past decade to increasingly connect to ideas of being and becoming processes that include social emotional dimensions mentioned above. Our research on science learning addresses principles for developing students’ agency to inquire about personally-relevant, socially-consequential science problems, while also supporting them to cultivate powerful intellectual practices

and positive attitudes toward science. The analysis presented in this paper connects to these ideas of learning, being and becoming, by highlighting the emotional dimensions that are often left out of empirical studies.

References

- Bereiter, C., & Scardamalia, M. (2009). Teaching how science really works. *Education Canada*, (49)3, 14-17.
- Brown, A. (1978). Knowing when, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 77-165). Hillsdale, NJ: Erlbaum.
- Dreyfus, H. L. (1984). What expert systems can't do. *Raritan*, 3(4), 22-36.
- Engle, R. A., & Conant, F. (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.
- Engle, R. A., McKinney de Royston, M., & Langer-Osuna, J. (2008). *Toward a model of differential influence in discussions: Negotiating { XE "negotiating" } quality, authority, and access within a heated classroom argument*. Paper presented at the Cognitive Science Society, Washington, DC.
- Esmonde, I., & Langer-Osuna, J. (2007). *Power in numbers: An analysis of how shifts in social { XE "social" } and mathematical power { XE "power" } among students in a cooperative group affect opportunities for learning { XE "learning" }*. Paper presented at the American Educational Research Association, Chicago, IL.
- Greeno { XE "Greeno" }, J. G. (2002). *Students with competence, authority, and accountability: Affording intellectual identities in classrooms*. New York: The College Board.
- Gresalfi, M. S. (2009). Taking up opportunities to learn: Constructing dispositions in mathematics classrooms. *Journal of the Learning Sciences*, 18(3), 327-369.
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behaviors. *Journal of Labor Economics*, 24(3), 411-482.
- Herrenkohl, L. R., & Guerra, M. R. (1998). Participant Structures, Scientific Discourse, and Student Engagement in Fourth Grade. *Cognition and Instruction*. 16 (4), 431-73.
- Herrenkohl, L.R. and Mertl, V. (2011). *How students come to be, know, and do: A case for a broad view of learning*. Cambridge UK, New York City: Cambridge University Press.
- Herrenkohl, L.R., Palincsar, A.S., DeWater, L.S., & Kawasaki, K. (1999). Developing scientific communities in classrooms: A sociocognitive approach. *The Journal of the Learning Sciences*, 8, 451-493.
- Holland { XE "Holland" }, D., Lachicotte { XE "Lachicotte" }, W., Skinner, D., & Cain, C. (1998). *Identity and agency { XE "agency" } in cultural { XE "cultural" } worlds*. Cambridge, MA: Harvard University Press.
- Metz, K. (2000). Young children's inquiry in biology. In J. Minstrell & E. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science*. (pp. 371-404). Washington, DC: AAAS.
- Nasir, N. S., & Hand, V. M. (2006). Exploring sociocultural { XE "sociocultural" } perspectives on race, culture, and learning { XE "learning" }. *Review of Educational Research*, 76(4), 449-479.
- Nasir, N. S., Rosebery, A. S., Warren, B., and Lee, C. D. (2006). Learning as a cultural { XE "cultural" } process. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp.489-504). NY: Cambridge University Press.
- National Research Council (2007). *Taking science to school*. Washington, DC: NAP
- National Research Council (2009). *Learning science in informal environments*. Washington, DC: NAP
- Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (October 2011). Organizing research and development at the intersection of learning, implementation and design. *Educational Researcher*, 40(7), 331-337.
- Scalone, G. (2014). Cultivating Agency in Science Argumentation and Sense-making in Elementary Science Classrooms. (Unpublished doctoral dissertation). University of Washington, Seattle, WA.
- Spiro, M. (1982). Collective representations and mental representations in religious symbol systems. In J. Maquet (Ed.), *On symbols in anthropology: Essays in honor of Harry Hoiyer* (pp. 45-72). Malibu: Udena.
- White, B., & Frederiksen, J. (1998). Inquiry, modeling and metcognition: Making science accessible to all students. *Cognition and Instruction*, 16(1), 3-118.

Acknowledgments

This work was supported by the NSF-SLC LIFE Center (#0835854) and NSF-DRK12 (#1019503). We wish to note that the findings and conclusions expressed here are those of the authors and do not necessarily reflect the views of the NSF. In addition, we wish to thank the teachers and students who participated in this research.

Filling in the Gaps: Capturing Social Regulation in an Interactive Tabletop Learning Environment

Abigail Evans and Jacob O. Wobbrock, Information School, University of Washington, Seattle, USA
Email: abievens@uw.edu, wobbrock@uw.edu

Abstract: A study of small groups collaborating at an interactive tabletop was conducted. Group discussions were coded according to the type and quality of social regulation processes used. Episodes of high and low quality social regulation were then matched with the software logs to identify patterns of interaction associated with quality of social regulation. A key finding is that instances of low-quality social regulation were characterized by more than twice as much interaction with the software as high-quality instances.

Introduction

Creating software that can adapt to the needs of learners and create personalized learning environments is an oft-cited goal of researchers and designers of educational technology (Jermann, Mühlenbrock, & Soller, 2002; Lonchamp, 2010; Marcos-García, Martínez-Monés, Dimitriadis, & Anguita-Martínez, 2007; Martínez, Collins, Kay, & Yacef, 2011a). Collaborative learning environments represent a particular challenge in personalization, as they must take into account interactions between multiple learners. Collaborative learning at tabletop computers brings further challenges, as much of the interaction between learners takes place face-to-face, away from the shared interface, and differentiating between individuals can be difficult (Dillenbourg & Evans, 2011).

The process of small group face-to-face collaborative learning is complex and many factors contribute to learning outcomes. The effectiveness of the group learning experience can be influenced by the personalities and motivations of individual students, their background knowledge, the nature of the task, and even their time management skills. Students do not always know how to collaborate in a manner that is productive and conducive to learning for all members (Rogat & Linnenbrink-Garcia, 2011). In the classroom, the teacher will often circulate among groups, but she can only visit one group at a time, meaning that struggling groups may not receive the help they need in a timely manner. Software that can adapt to a group's collaboration needs and abilities in real-time could prove useful in filling in the gaps while the teacher spends time with other groups.

The goal of this research is to uncover what physical interactions with a tabletop computer reveal about the collaborative interactions taking place above the table, which is arguably where the learning actually occurs. Rather than trying to determine if a group is meeting content-based learning objectives, our intention is to understand the collaboration itself to inform future development of tabletop software that can adapt to scaffold effective collaboration. While it may not always be true in practice (Dillenbourg & Evans, 2011; Rogat & Linnenbrink-Garcia, 2011), an implicit assumption in this work is that effective collaboration leads to positive learning outcomes, while ineffective collaboration hinders learning.

Important to this work is the concept of “social regulation,” which refers to “the social processes groups use to regulate their joint work on a task” (Rogat & Linnenbrink-Garcia, 2011). Rogat and Linnenbrink-Garcia (2011) showed social regulation to be very important to small group collaboration. They identified three dimensions of social regulation: *planning* the group's approach to a task, *monitoring* of understanding and progress, and *behavioral engagement*—efforts to get group members to engage with the task. We use this conception of social regulation to guide the analysis of our participants' collaboration.

Related Work

When learners collaborate at an interactive tabletop computer, only interactions with the software are usually captured. The verbal and gestural interactions among students that represent the learning process are not visible to the computer, making it tricky to capture meaningful assessment data or to provide relevant dynamically generated feedback to learners and teachers. A small body of work has begun to explore this problem space.

Martínez *et al.* (2011a; 2011b) created visualizations of learners' verbal contributions and physical interactions with the software during a group concept-mapping activity at a tabletop computer. The visualizations proved useful for determining how equitable the group members were in terms of the quantity of participation, but could not communicate the quality of the collaboration itself. The same authors (2013) later created a dashboard that helped teachers determine which groups needed attention as they worked by enabling the teacher to see how closely each group's concept map matched that of the teacher. Martínez *et al.*'s positive results suggest that this is an effective strategy for activities in which students work towards a particular ideal, or well-defined solution. It would not, however, be effective in the absence of an expert model, as is generally the case in creative work, or in collaborative work in the spirit of knowledge-building (Scardamalia & Bereiter, 1994), in which learners build new knowledge and understanding. Our work seeks to model and adapt to collaborative processes in these more learner-led contexts.

Fleck *et al.* (2009) explored the relationship between students' discussions and their actions on an interactive tabletop. They found that verbal elements of successful collaboration, such as making and accepting suggestions or negotiating, were often complemented by particular actions in the software. This finding suggests that some aspects of collaborative discussion should be detectable through interaction analysis. Our work builds on these findings to identify patterns of action over time that can reveal groups' social regulation processes.

The Study

Eleven adults (7 female, 4 male) worked in small groups (two dyads, a group of three, and a group of four) to analyze and compare two poems, *Birches* by Robert Frost (1969), and *Fern Hill* by Dylan Thomas (2003). We chose poetry analysis for our study because many students find it challenging, and yet it requires little background knowledge, making it an authentic learning activity for most participants. Poetry analysis is interpretative, which also makes it an ideal activity for small group work as students can reach deeper levels of understanding by externalizing their own thoughts and building on the ideas of others. The participants were given 30 minutes to work on the following task adapted from McMahan, Funk, & Day (1998):

Compare and contrast two poems, *Fern Hill* by Dylan Thomas and *Birches* by Robert Frost.

Answer the following questions and support your answers with evidence from the text. As you answer each question, consider how the two poems are similar and how they are different.

1. What is the **theme** (the central idea) of each poem?
2. Who is the **speaker** in each poem? How would you describe the speaker?
3. What **imagery** does each poet use? How do the images relate to each other and contribute to the theme?



Figure 1. A screenshot of the software used in the study

The participants used a Microsoft PixelSense, which has a 40" multitouch screen and can comfortably seat four adults around it. The software (see Fig. 1), created specifically for this study, was intended to be used as a tool to support discussion rather than for creating a result to be turned in. At the end of the activity, the participants were asked to verbally summarize their findings to a researcher as if they were feeding back to the rest of a class after a period of small-group work. This meant that the participants were relatively unconstrained in how they tackled the task and were free to make use of the software in the ways they found most useful.

The text of both poems was presented on-screen and supplemented by audio recordings of readings by the poets. Participants could take notes by clicking buttons to add "notecards" to the screen. Notecards were color-coded according to each element of the poems they were asked to discuss (*e.g.*, red for theme, green for speaker, and yellow for imagery) and participants could add as many notecards as they wanted. Each notecard contained a text field in which participants could type an observation about the text, and an area for collecting supporting evidence by dragging and dropping lines from the poems onto the notecard. Once dragged onto a notecard, individual words could also be highlighted. Participants could annotate the poems directly by drawing on them with their fingers. The software also included a set of optional prompt questions to stimulate discussion. Each tool described so far could be freely moved around the screen. In addition, fixed pieces of "scratch paper" were provided at each corner to enable each participant to take their own notes.

Two video cameras recorded the discussions. Interactions with the tabletop were recorded in log files detailing when and where a touch happened, and what action was performed. The researcher left the room while the participants worked to encourage them to give an authentic summary at the end of the activity.

Data Analysis

To determine how social regulation manifested in each group, each video was coded along the three dimensions of social regulation described by Rogat and Linnenbrink-Garcia (2011), and for high, moderate, and low quality.

The codes are listed in Table 1, but detailed descriptions of each code, as well as high and low quality examples, can be found in the authors' article.

Table 1: Descriptions of social regulation codes, excerpted from Rogat and Linnenbrink-Garcia (2011)

<i>Codes</i>	<i>Description</i>
Social regulation	Group efforts to regulate their conceptual understanding, task work, and engagement.
Planning	Reading and interpreting task directions, designating task assignments, discussing how to go about solving the problems.
Monitoring	Evaluating content understanding, the shared product, assessing progress, or plan for completing the task.
Behavioral engagement	Encouraging an off-task group member to re-engage, reminding a group member to return to task.

In order to draw useful insights from the log files of touch data, individual touches needed to be grouped into sequences that represented purposeful actions—actions that served a particular purpose. For example, imagine a participant wants to read one of the poems, which is currently located closer to another participant on the other side of the screen. He makes one touch to move the poem closer to himself, followed by a second touch to rotate it so it is oriented properly from his perspective. In this case, the two touches together represent a single purposeful action. Three features are needed to make that determination: (1) who carried out the touch (the owner); (2) the object that was touched; and (3) the timing of the touches. Items 2 and 3 are readily available in the logs but, as the PixelSense cannot natively differentiate individual users, for this study, the owner was manually labeled by matching the timestamp of the touch to a frame from the video.

For the first group, touches were grouped into meaningful actions by close inspection of the log file alongside the video. The timing between touches was studied to determine the distinction between sequences of touches that represented purposeful actions, and sequences that should be considered separate actions. Clear rules emerged from this analysis for sequences of touches owned by a single person in terms of object type and length of time between touches (see Table 2).

Table 2: Rules for grouping touch event data into purposeful actions by time between touches

<i>Objects touched are:</i>	Purposeful action sequence	Separate actions
<i>Related</i>	≤ 50 seconds	> 50 seconds
<i>Unrelated</i>	≤ 15 seconds	> 15 seconds

The object types *related* and *unrelated* refer to objects' function in the context of the activity. For example, the time between a touch on the Thomas poem followed by a touch on a line from the Thomas poem that has been dragged to a notecard would be categorized as *related*, while a touch on the Frost poem followed by a touch on the instructions would be *unrelated*. Objects could be related either by poem, *e.g.*, a notecard containing lines from Thomas is related to the Thomas poem; or by task sub-question, *e.g.*, a theme notecard containing lines from Frost is related to the Thomas poem in "Annotate Themes" mode.

For groups 2, 3, and 4, logged touches were automatically grouped according to the rules given in Table 2. The rules were then validated by repeating the video inspection for these groups and, for each category, calculating the percentage of touches that were correctly categorized. The rules proved to be highly accurate, ranging from 87.2% to 100.0% correct.

The screen of the PixelSense is highly sensitive and as a result, the touch logs contained a lot of noise, often due to participants catching on-screen objects with their elbows while reaching for another object. Another source of noise came from multi-touch input—using more than one finger to carry out a single action, such as rotating an object using multiple fingers. Simple heuristics were used to filter out these types of noise: (1) touches overlapping (in time) by the same person on the same object were combined, (2) when there were overlapping touches by the same person on different objects and one touch was close to the edge at which the person was sitting, the edge touch was removed, and (3) when there were overlapping touches on multiple objects located directly on top of each other, the top touch was kept and all others were removed.

The third stage of analysis brought together verbal and physical interactions. The goal of this stage was to find patterns of touches that reflected social regulation processes used by each group, as well as their quality. Therefore, we identified the processes for which there were clear differences among the groups in terms of quality. This substantially narrowed the focus of the subsequent analysis as the groups had more similarities than differences. Timestamps from the video that delineated episodes where a group was engaging in a particular regulatory process were applied to the pre-processed touch data. The episodes were then compared to see whether patterns emerged. Only touches to the shared objects (*i.e.*, everything but individuals' separate "scratch paper") were considered at this stage of the analysis.

Results: Forms and Quality of Social Regulation

All groups spent around ten minutes of the activity silently reading the poems and listening to the recordings before starting discussion. The nature of the task demanded relatively little *planning* and all groups were coded as high quality in this area.

After the initial period of preparation, however, each group took a different approach in their discussions and their use of the software. Group 1 (3 participants) structured their discussion around the task's three sub-questions, comparing the two poems as they considered the themes, speakers, and imagery in order. They took notes using the notecards and copied lines from each poem to support their observations. Group 1 made the most effective use of the allotted time, covering each aspect of the task in depth. Their *monitoring* processes were consistently high quality across all 3 sub-codes.

Group 2 (4 participants) did a great deal of comparing and contrasting of details in the poems. They were the only group to make use of the optional discussion prompts. They annotated the poems directly and wrote notes on their individual virtual notepaper. Group 2 did not follow a clear structure. They covered each sub-question, but not in a linear fashion. They exhibited consistently high-quality *monitoring* of the plan and progress. Although their content monitoring was mostly high quality, there were some low-quality instances, largely due to one group member's occasionally dismissive and unresponsive treatment of others' contributions.

After reading both poems, Group 3 (a dyad) decided to tackle one poem at a time. They started with the Thomas poem and covered each task sub-question in turn with just that poem, taking notes using the notecard feature for the theme and speaker before switching to directly annotating the imagery in the poem. They did not get to the Frost poem until near the end of the activity, eventually running out of time having only considered its theme and not its speaker or imagery. As a result, the group was not able to make many comparisons between the two poems, although their analysis of *Fern Hill* was exceptionally thorough. In terms of social regulation, their discussion was focused on content monitoring, all of which was high quality. In contrast, progress monitoring was largely missing and monitoring of the plan was very limited, though generally high quality.

Group 4 (a dyad) were more focused on organizing the virtual workspace and collecting lines than on actually engaging with the poems. Their discussion of both poems, though valid, was mostly superficial and more of a summary than an analysis. Most instances of content monitoring were of low quality, although they did begin to improve near the end of the discussion. They were easily distracted by superficial concerns, which meant that thoughtful contributions were often unheeded—an indicator of low-quality progress monitoring. They did, however, also engage in high-quality progress monitoring, regularly tracking what remained to be done. Their plan monitoring was generally moderate as, although they frequently sought to clarify the instructions and evaluate their plan, their continual monitoring of the plan hindered their ability to enact it.

Results: Social Regulation in the Touch Data

As described above, the groups differed in quality along the monitoring dimension of social regulation, so the three monitoring sub-codes were used to identify patterns of touch data that reflected quality of monitoring processes. Results were normalized by group size to facilitate comparison.

Monitoring Content

Monitoring content was the most-used process of social regulation, with noticeably longer duration than the other processes. Quality variations in content monitoring revealed distinctive patterns of touches. Low-quality episodes, all from group 4, were characterized by sequences of purposeful touches to unrelated objects. Group 4, which was consistently low quality, averaged 23.5 such sequences, more than double that of the high-quality groups. Group 1 averaged just 2.3, while group 2 averaged 10.5, and group 3 averaged 11.0. During low quality episodes, sequences of touches were also much temporally denser—many touches in quick succession. Overall, group 4 spent more than twice as much time, 476 seconds per person, touching the screen than any other group. The high-quality groups were similar to each other: group 1 = 199s, group 2 = 136s, and group 3 = 190s. Finally, high-quality episodes were characterized by only one person interacting with the screen at a time, moving from one person to another with little overlap, suggesting turn-taking. In groups 1, 2, and 3, overlapping touches by two or more people made up 10.5%, 10.7%, and 9.6% of their total touches to shared objects, respectively. Conversely, low-quality episodes often featured overlapping touches. For group 4, overlapping touches made up 23.6% of their total touches to shared objects.

Monitoring the Plan

The log data contained indicators that plan monitoring was occurring. Participants sometimes touched the instructions on-screen when revisiting the task and re-evaluating their plan and otherwise did not interact with it. However, the variations in quality among the groups did not present distinctive patterns.

Monitoring Progress

Episodes of progress monitoring did not show any distinctive patterns in the touch data, regardless of quality. Progress monitoring episodes were generally brief, predominantly verbal interactions that did not generate much

touch data. While deictic gestures did appear in the video, they were generally above the screen and so were not captured by the log data.

Conclusion

High-quality content monitoring has been shown to be important in the construction of shared meaning; therefore, the ability to automatically detect the quality of content monitoring in a collaborative tabletop environment would be extremely useful. While the clear distinction between the high- and low-quality episodes is a promising result, it is important to note that the low-quality episodes were limited to just one group and therefore could be idiosyncratic to their particular style of interaction. However, the similarities between high-quality episodes across groups suggest that the patterns found might be generalizable.

The prevalence of episodes of content monitoring significantly aided the process of finding obvious patterns. In the context of the poetry analysis activity, there was relatively little need for planning or monitoring planning as the task structure was simple, so the episodes were brief and few. Also, as the task did not require the production of a shared artefact to be submitted for evaluation, there was perhaps less need for progress monitoring. It would be interesting to see if activities with different structures evoke different social regulation behaviors, which might cause more patterns to emerge. Additionally, the adult volunteers participating in this study remained on-task and engaged, so regulatory processes related to behavioral engagement were not employed. We expect this would not be the case in a real classroom environment. We are planning to conduct a follow-up study with middle and high school aged children to investigate these questions, and also to see whether the content monitoring patterns discovered transfer to other domains and activities, along with the rules we used to identify purposeful touch sequences.

References

- Dillenbourg, P., & Evans, M. (2011). Interactive tabletops in education. *International Journal of Computer-Supported Collaborative Learning*, 6(4), 491-514.
- Fleck, R., Rogers, Y., Yuill, N., Marshall, P., Carr, A., Rick, J., & Bonnett, V. (2009). Actions speak loudly with words: Unpacking collaboration around the table. *Proceedings of the International Conference on Interactive Tabletops and Surfaces* (189-196). New York: ACM Press.
- Frost, R. (1969). *Birches. The Poetry of Robert Frost*. New York: Henry Holt and Co.
- Jermann, P., Muhlenbrock, M., & Soller, A. (2002). Designing computational models of collaborative learning interaction. *Proceedings of the Conference on Computer-Supported Collaborative Learning* (730-732). Boulder: International Society of the Learning Sciences.
- Lonchamp, J. (2010). Customizable computer-based interaction analysis for coaching and self-regulation in synchronous CSCL systems. *Journal of Educational Technology & Society*, 13(2), 193-205.
- Marcos-Garcia, J. A., Martinez-Mones, A., Dimitriadis, Y., & Anguita-Martínez, R. (2007). A role-based approach for the support of collaborative learning activities. *e-Service Journal*, 6(1), 40-58.
- Martinez-Maldonado, R., Kay, J., Yacef, K., Edbauer, M.T., Dimitriadis, Y., (2013). MTClassroom and MTDashboard: Supporting analysis of teacher attention in an orchestrated multi-tabletop classroom. *Proceedings of the Conference on Computer-supported Collaborative Learning* (320-327). Madison: International Society of the Learning Sciences.
- Martinez, R., Collins, A., Kay, J., & Yacef, K. (2011a). Who did what? Who said that? Collaid: An environment for capturing traces of collaborative learning at the tabletop. *Proceedings of the International Conference on Interactive Tabletops and Surfaces* (172-181). New York: ACM Press.
- Martinez, R., Kay, J., & Yacef, K. (2011b). Visualisations for longitudinal participation, contribution and progress of a collaborative task at the tabletop: *Proceedings of the Conference on Computer-Supported Collaborative Learning* (25-32). Sydney: International Society of the Learning Sciences.
- McMahan, E., Funk, R., & Day, S. (1998). *The elements of writing about literature and film*. London: Longman.
- Rogat, T. K., & Linnenbrink-Garcia, L. (2011). Socially shared regulation in collaborative groups: An analysis of the interplay between quality of social regulation and group processes. *Cognition and Instruction*, 29(4), 375-415.
- Scardamalia, M. & Bereiter, C. (1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences*, 3(3), 265-283.
- Thomas, D. (2003). Fern Hill. *The Poems of Dylan Thomas*. New York: New Directions.

Acknowledgements

This work was supported in part by the National Science Foundation under grant IIS-0952786. Any opinions, findings, conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect those of the National Science Foundation.

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

Sasha Lotas, University of Washington, Seattle, Washington, sashalotas@gmail.com

Abstract: As adult literacy GED programs are increasingly tasked with serving the educational needs of both young and older adult students in a knowledge-based economy, there is growing interest in how to make these programs more effective and efficient, especially at helping students develop the academic writing skills needed to pass the new common-core aligned GED 2014. However, in the attempt to make these programs more efficiently effective, the promotion of prescriptive GED writing pedagogies, based on an autonomous view of literacy, is likely. The long-term goal of this reflection study is to learn how to thoughtfully develop a framework - based on a functional linguistics approach to grammar that foregrounds the social concept of identity - that can produce alternative GED writing pedagogies, promoting more socially-just effective practices within adult literacy programs and helping to ensure that adult literacy students have access to empowering academic writing curricula.

Issues Addressed

Although adult literacy programs predate public schooling, these programs have historically been marginalized within the educational system; however, as adult literacy programs are increasingly tasked with serving the educational needs of both young and older adult students in a knowledge-based economy, there is growing interest in how to make these programs more effective, especially at helping students develop academic writing skills. Starting next year, adult literacy programs will begin to prepare students for the GED 2014, a revised version of the exam based on Common Core Standards which will measure “career- and college-readiness (GED Testing Service, 2013). Although the exam is undergoing numerous changes, one of the greatest changes is its new focus on academic writing – a change reflecting the increased national focus on improving writing skills “in a global information economy that continually raises the bar for what counts as literacy” (Juzwik et al., 2006, p.453). While the current version of the exam assesses writing only in its Language Arts Writing section through multiple-choice grammar-based questions and a narrative essay prompt (Gillespie, 2007), the 2014 version of the GED will test evidence-based, analytic writing in both the Language Arts section as well within the Science and Social Studies sections (GED Testing Service, 2013).

However, this Common Core Standards makeover with its new emphasis on academic writing does not presume the adoption of more academically challenging and socially just adult literacy pedagogies. Instead, as the GED Testing Service (part of the nonprofit American Council on Education) merges with Pearson (a for-profit educational testing company), it may be likely that more commercialized, script-based materials – based on an autonomous view of literacy – will be adopted by GED instructors, possibly continuing to send cognitive “deficit” messages to GED students and making their “second-chance” at an education unproductively similar to their “first-chance”. This may be especially true in writing instruction, which may be reduced to superficial culturally irrelevant grammar and skills-based instruction.

The goal of this paper is reflect on how to develop an academic writing instructional framework within GED programs that promote tools to encourage meaningful student participation. One possible framework to explore is grounded within a functional linguistics approach to grammar that foregrounds the social concept of identity.

Contextualization

Often, GED preparation classes are overly reliant on remedial educational pedagogy and test-taking strategies, making GED preparation classes less likely to serve as authentic educational pathways for students; this is especially true in GED writing classes which tend to focus on skill and drill decontextualized grammar exercises, conveying to students “a very restricted model of the composing process” (Rose, 1983, p.109). Russell (cited in Gillespie, 1998), for example, in her interviews with GED writing students, found that students’ strong beliefs about the importance of mechanics was possibly interfering with their writing development. When she asked a student to describe what makes someone a good writer, the student answered “Knowing how to punctuate things. And not having to have so many mistakes on a paper and everything just being fine the first time” (p.5).

Because of the restricting possibilities inherent in a decontextualized-grammar-based GED class, some GED instructors choose not to teach grammar, hoping to “transcend the deluge of skills exercises” by focusing on meaning over form (Shafer, 2004). Yet, there are problems with “dethron[ing]” form and “crown[ing]”

meaning (Myers, 2003, p.55). Powerful academic writing depends on a writer's grammatical choices, and not teaching students to define and identify grammatical elements may be another way to constrain GED students, constraining their ability to use academic writing to express their intended meanings as well as to make new meanings (Micciche, 2013). "Meaning does not flow from knowledge of syntax and lexis, but the ability to express meaning [in academic context] does" (Myers, 2003, p.55).

One possible solution to address this grammar-teaching dilemma - on the one hand, relying on grammar instruction and basic skills remediation, "which comes from socially constructed deficiency assumptions about students' work" (Sanchez & Paulson, 2013, p.117), may further marginalize GED students; on the other hand, choosing not to focus on grammar may also further marginalize GED students by not giving them access to and the power to affect the academic Discourse community - is to incorporate a functional linguistics approach to grammar. A focus on grammar that is functional focuses on language in-use. Instead of presenting grammar as a set of rules about what is correct or incorrect, a functional linguistics approach presents grammar as meaning-making (Halliday & Mathiessen, 2004; Schelepppegrell, 2007).

In particular, functional grammar emphasizes syntactic morphological knowledge, the understanding of how a derivational suffix changes the part of speech of a word (Kiefer & Lesaux, 2012, p.521). For example, because academic writing often focuses on abstract concepts, academic writers need to be able to turn verbs into nouns (nominalization). Through the use of morphologically related words (e.g., *colonies*, *colonial*, *colonists*, *colonize*, *colonization*), students are able to maintain thematic coherence across paragraphs as they subtly foreground, background, and nuance concepts in service of their intended message. Rather than completing isolated grammar exercises, students come to understand how they can harness word forms and syntax to communicate their ideas more effectively without "continu[ing] to confine students to the impoverished meaning carried out by the conventional rules of language" (Reither, p.144). Syntactic morphological knowledge also allows students to understand and experience the invention in sentence making. As Shaughnessy (1970) explains: "There is a kind of carpentry in sentence making, various ways of joining or hooking up modifying units to the base sentence" (p.7). Instead of feeling afraid of making mistakes and having "everything [on paper] just being fine the first time, students can experience "how to mess with sentences;" students can become "sensitive to the questions that are embedded in sentences, which, when answered, can produce modifications within the sentence or can expand into paragraphs or an entire essays" (Shaughnessy, 1977, p.4). Ultimately, helping students understand how to syntactically play with words brings them closer to being able to express their intended meanings, while allowing them to experience writing as "knowledge transforming" (Bereiter & Scardamalia, 1987), which, ultimately, makes academic writing worth doing.

Although adopting a functional approach to grammar may help to solve the grammar-dilemma, it also may be problematic because it does not encourage students to reflect on how they negotiate their identities as academic writers. Proponents of a functional linguistic approach acknowledge that students bring a range of linguistic repertoires into the classroom; however, they seem less likely to consider how these linguistic repertoires may influence students' perceptions of themselves, as well as impact students' decisions to participate in literacy practices, especially academic writing.

Academic writing "is not simply a replacement of some practices by others. It is a process . . . which requires deep emotional commitment and involvement on the part of the language user" (Cook-Gumperz, 1993, p.338). Lillis (2001) explains that students' sense of identity significantly influences "what writers (don't) write and (don't) wish to write in academia" (p.48). Students have many different selves - selves with sometimes conflicting voices - which, if not addressed, may interfere with their understanding and completion of writing assignments (Lea, 2004). The question of identity seems particularly important when it comes to vocabulary and grammar instruction, for accepting another "set of grammatical paradigms, rhetorical practices, and usage conventions" has identity and ideological implications (Cook-Gumperz, 1993, p.338). As Gee (1996) explains, language use is connected to power "often in the service of privileging certain types of literacies and certain types of people" (p.46). Purcell-Gates (1993), in her case-study of Jennie, a white woman attending adult literacy classes in Appalachia, illustrates what may occur when a student's primary discourse is marginalized from the dominant Discourse (Gee, 1996). Despite many years of adult literacy grammar instruction, Jennie continued failing year after year "trying to memorize rules, trying to memorize terms like adverb and pronoun" because "none of these words, these rules, these linguistic terms were hers . . . and thus she could not succeed because her words were never acknowledged and affirmed, never allowed" (as cited in Gillespie, 1998).

Theoretical and Methodology Approaches

Because of the importance of considering identity and literacy, identity should be foregrounded within a functional linguistic approach to academic writing in a GED program. The diagram below is a start of what this framework may look like. It depicts how when a word and its forms are taught within an academic writing class to help students eventually build an academic essay, that word and its forms also have an impact on a student's language, which impacts his/her discourse, ultimately impacting his/her identity.

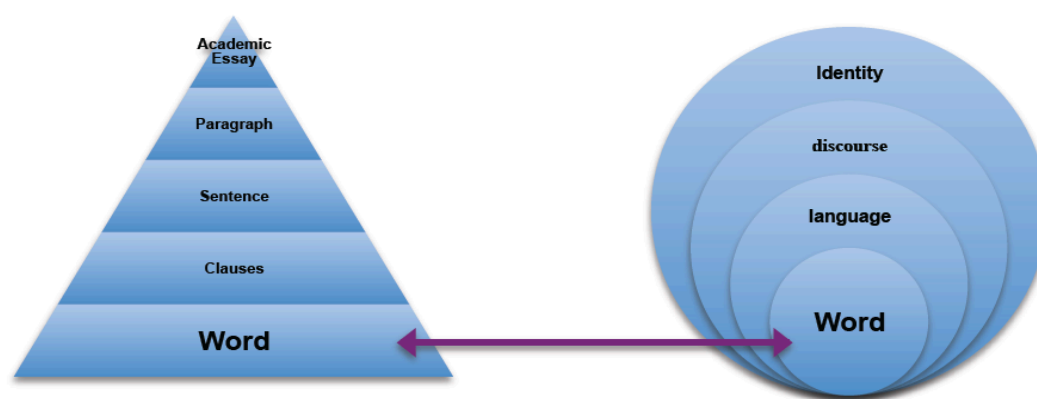


Figure 1. Reverberating words in GED 2014 writing instruction

Based on this emergent framework, questions can be developed to help more effectively develop academic instructional writing tools. For example:

- Based on the idea of a “reverberating” learned academic word, what might a curriculum intervention look like? What tools might be developed and why? How might these tools be studied within the theoretical framework?

The Credibility and Usefulness of the Study for Raising New Questions and Prompting New Possibilities

With the introduction of the new and privatized GED 2014, it is likely that adult literacy education will be further commodified and focused on measurable skills-acquisition, continuing to link literacy to the needs of a postindustrial economy; in addition, it is also likely that adult literacy programs will remain underfunded and continue to lack trained, professional instructors (Gillespie, 2007). Because of this, adult literacy students and their needs as learners may be further backgrounded within some GED programs. The long-term goal of this reflection study is to learn how to thoughtfully develop a framework that can produce student-centered instructional tools that promote effective practices with adult literacy programs, helping to ensure that students have access to empowering academic writing curricula.

The Relevance of the Theme for the Conference

In exploring why and how to develop an academic writing instructional framework grounded in a functional linguistics approach to grammar that foregrounds the social concept of identity, this reflection acknowledges that learning involves participating in different practices, encouraging all levels of educational programs – especially adult literacy programming - to respect the becoming of their students in these practices.

References

- Bereiter, C & Scardamalia, M. (1987). Two models of the composing process. In *The psychology of written composition*. Hillsdale, NJ: Erlbaum.
- Cook-Gumpres, J. (1993). Dilemmas of identity: Oral and written literacies in the making of a basic writing students. *Anthropology & Education Quarterly*, 24(4),336-356
- Delpit, L. (2001). The politics of teaching literate discourse. In Cushman, C., Kintgen, E.R., Knoll, B.M., and Rose, M. (Eds.), *Literacy: A Critical Sourcebook*. Boston, MA: Bedford.
- Demetron, G. (2005). *Conflicting paradigms in adult literacy education: In quest of a US democratic politics of literacy*. Mahwah, NJ: Erlbaum.
- Flowers, L. & Hayes, J.R. (1980). The Cognition of discovery: Defining a rhetorical problem. In Landmark Essays on Writing Process, Perl, ed. (1994), Hermagoras Press: Davis, CA
- GED Testing Service (2011). 2011 Annual Statistic Report on the GED test. Retrieved March 20, 2013 from http://www.iccb.state.il.us/pdf/ged/GEDTS_Report_2011.pdf
- Gee, J.P. (1996). *Social linguistics and literacies*. Philadelphia, PA: Falmer.
- Gillespie, M.K. (1998). Using research on writing: Implications for adult literacy education. Retrieved from <http://www.ncsall.net>.

- Gillespie, M.K. (2007). The forgotten R: Why adult educators should care about writing instruction. In A. Belzer (ed.) *Toward defining and improving quality in adult basic education* (pp.87-106) Mahwah, NJ: Lawrence Erlbaum Associates.
- Halliday, M.A.K. & Math
- Hull, G. & Rose, M. (1989). Rethinking remediation: Toward a social-cognitive understanding of problematic reading and writing. *Written Communication*, 6(2), 139-154.
- Hull, G., Rose, M., Fraser, K.L., & Castellano, M. (1991). Remediation as a social construct: Perspectives from an analysis of classroom discourse. *College Composition and Communication*, 42(3), 299-328.
- Ivanic, R. (2004). Discourses of writing and learning to write. *Language and Education*, 18(3), 220-245.
- Juzwik, M.M., Curcic, S., Wolbers, K., Moxley, K.D., Dimling, L.M., & Shankland (2006). Writing into the 21st century: An overview of research on writing, 1999-2004. *Written Communication*, 23(4), 451-476.
- Lesgold, A.M. & Welch-Ross, (2012). (Eds.). *Improving adult literacy instruction: Options for practice and research*. National Academy of Sciences. Washington, D.C.: The National Academies Press.
- Lillis, T.M. (2001). *Student writing: Access, regulation, desire*. New York: Routledge
- Lillis, T.M. (2003). Student writing as 'academic literacies': Drawing on bakhtin to move from critique to design, *Language and Education*. 17(3), 192-207.
- Lytle, S. (2001). Living literacies: Rethinking development in adulthood. In Cushman, C., Kintgen, E.R., Knoll, B.M., and Rose, M. (Eds.), *Literacy: A Critical Sourcebook*. Boston, MA: Bedford.
- Micciche, L.R. (2013). Making a case for rhetorical grammar. In *Teaching Developmental Writing*. New York, NY: Bedford Press.
- Rose, M. (1983). Remedial writing courses: A critique and a proposal. *College English*, 45(2), 109-127.
- Rose, M. (1985). The language of exclusion: Writing instruction at the university. *College English*, 47(4), 341-351.
- Rose, M. (1988). Narrowing the mind and page: Remedial writers and cognitive reductionism. *College Composition and Communication*, 39(3), 267-302.
- Rose, M. (1989). *Lives on the boundary: A moving account of the struggles and achievements of America's educationally unprepared*. New York: Penguin Books.
- Rose, M. (1996). *Possible Lives*. New York: Penguin Books.
- Rose, M. (2010). The Smithsonian of basic skills. Retrieved from www.mikeroose.com
- Schleppegrell, M. J. (2007). At Last: The Meaning in Grammar. *Research in the Teaching of English*, 42, 1, 121-128.
- Schleppegrell, M. J. (March 01, 2012). Academic Language in Teaching and Learning. Introduction to the Special Issue. *The Elementary School Journal*, 112, 3, 409-418
- Shaughnessy, M.P. (1977). *Errors and expectations*. New York: Oxford University Press.
- Stuckey, E. (1991). *The violence of literacy*. Portsmouth, New Hampshire: Boynton/Cook Publishers.
- Williams, J.D. (2003). *Preparing to teach writing: Research, theory, and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Influence of Public Design Critiques on Fifth Graders Collaborative Engineering Design Work

Michelle E. Jordan, Arizona State University, Mary Lou Fulton Teachers College, Farmer Education Building
1050 S Forest Mall, Tempe, AZ 85287-1811, Email: Michelle.e.jordan@asu.edu

Abstract: Understanding how young students learn to engage in collaborative design practices entails understanding social interaction processes that occur beyond collaborative groups. The purpose of this study was to understand how talk generated during whole-class public design critique sessions influenced collaborative groups in subsequent small-group work sessions. Analysis focused on data from one fifth-grade class in which students were challenged to collaboratively design, build, and program robots. Video-recorded and transcribed whole-class interactions from three design critique sessions across the second, third, and fourth day of a 14-day robotic engineering design project were examined in order to categorize the types of comments made by the teacher and students relative to the nascent design solutions of three focal groups. Collaborative discourse from subsequent small-group work sessions was then examined in order to understand what ideas students took up, as well as how they took them up and to what effect.

Introduction

Grounded in socially oriented theories of learning that emphasize the role of social interaction (Bransford, Brown, & Cocking, 1999; Greeno, 2006; Rogoff, 1990), learning scientists frequently attend to important dimensions of interactional patterns among group members for increasing the potential of learning from collaborative experiences. Fewer scholars have investigated how social interaction in the broader classroom context (i.e., with classmates and teachers) influences small-group collaborative project experiences and outcomes (but see Greiffenhagen, 2012; Roth, 1995). Such investigations, motivated by a learning sciences' focus on authentic learning in complex contexts, needs to include research on ill-structured problems and design projects because such experiences entail the complex authentic tasks that learning and becoming in practice entail. Thus, the study described here focuses on one important practice identified in the ICLS 2014 conference theme, *engaging in design*, by exploring learning opportunities afforded through design critique processes related to fifth graders' collaborative engineering design projects. Drawing on data from a 14-session challenge to collaboratively design, build, and program robots, analysis focused on social interaction that occurred around three public design critique sessions in which students presented their engineering projects-in-progress and received formative feedback from their teacher and classmates. Two research questions guided analysis:

- What was the nature of talk elicited from the teacher and students during the design critique sessions?
- How did students take up ideas from design critique sessions in subsequent small-group work sessions?

The aim of the study was to understand how public design critique processes influence the design activity of early adolescents learning to engage in collaborative engineering design in order to increase knowledge of how classroom-level social interaction influences learning from collaborative design projects.

Theoretical Framework

Learning scientists and educational researchers increasingly investigate the use of engineering design projects with children and young adolescents as understanding of the affordances of design for learning increases (e.g., Kafai & Ching, 2004) and as K12 engineering education gains prominent attention (e.g., NRC, 2009, 2011). Helping learners engage in engineering design favors project-based instruction as design tasks are open in their problem specification and solutions (Dorst, 2003). Constructionist theorists emphasize that learning is more effective when children make their own projects, construct their own ideas, and design their own solutions to problems (Bers, 2008, Papert, 1980). Furthermore, design activities have been shown to be more effective than direct instruction for helping students learn complex science concepts (Hmelo, Holton, & Kolodner, 2000) and helping students see the connections between science concepts and solutions to real world problems (Kolodner et al., 2003). Engineering design projects not only help students improve their ability to build things, but also understand what the building of things entails (Sadler, Coyle, & Schwartz, 2000). Additionally, the ill-structured nature of such projects can help students learn to cope with the uncertainty inherent in design endeavors (Jordan & McDaniel, in press). Finally, engineering design projects may be particularly effective for engaging young girls and other under-represented students in STEM learning (Laursen, Liston, Thiry, & Graf, 2007). These are compelling reasons for K-12 educators to create and implement engineering design projects. Yet, there is still much to understand about the social interaction processes that facilitate learning from such projects.

More broadly, design tasks entail the creation of an idea or physical artifact through thinking and manipulating tools and/or materials. Design practices cut across disciplines, and design literacy is seen by some

scholars and practitioners as a crucial need of the 21st century (Pendleton-Jullian & Brown, 2011). Design practices have most often been investigated in fields such as architecture, engineering, and computer programming, and in the post-secondary education literature pertaining to those fields. Design tasks are underdetermined because a design cannot be fully defined by the problem statement. Designers must rely on creativity to generate a unique product. However, they are not free to invent anything; problem constraints partly determine a problem. Engaging in design involves iteratively redefining the problem and choosing among multiple paths that can be taken to infinite, unpredictable solutions (Jonassen, 2000).

Social interaction plays critical roles in design endeavors. Design is situated action that takes place over time through social interaction and reflection (Fleming, 1998). Communication among engineering team members is widely recognized as a fundamental aspect of design processes (Darling & Dannels, 2008) and learning to design (Jordan & Babrow, 2013). Engineering design also requires the ability to communicate effectively beyond the design team through online and face-to-face mediums (Jordan, 2014; Otto & Wood, 2000). Most notably for the purpose of this study, *design critique sessions* (DCS) are important opportunities for social interaction to facilitate learning. Used extensively in settings such as art schools, design studios, and engineering labs, DCS usually consist of public formative feedback related to an individual or group's ongoing designed project (Sawyer, 2012). The importance of DCS lays in that the teaching of design depends on communication among design students and teachers. Social interaction during DCS consists of cyclical processes of "telling and listening" and "demonstrating and imitating" (Demirbas, 2001). It draws attention to efficient and effective solutions as well as to design inconsistencies while facilitating knowledge exchange related to procedural aspects of design (Uluoglu, 2000). Ideally, students develop arguments about the validity of design solutions, defend their solutions, develop ideas and refine their projects (Schaffer, 2003). However, critical feedback can be difficult for students to accept and make use of in subsequent drafts of design work.

Previous analysis focused on discourse in the DCS, identifying shifts in social interaction patterns that signaled learning across the three design critique sessions. These included shifts (a) from talk focused on helping students deal with negative feedback to helping students consider needs of the customer, (b) from talk in which classmates talked only to the teacher and the presenters to talk in which classmates talked to presenters and to each other, the use of only sketches in presentations of design solutions to use of built structures and websites as discursive tools, (c) from reliance on self-generated design ideas to accepting and valuing remixes of other's creations incorporated into designs (See Kafai, Fields, & Burke, 2010), and (g) from reporting design ideas to co-constructing design ideas. Additionally, across the DCS, there was an increase in authority taken up by students, and a focus on structures and functions gave way to integrated talk about structures, behaviors, and functions (Jordan, in press). The current study builds on these findings by conducting a more fine-grained analysis of discourse related to three focal groups' design solutions during DCS and by expanding analysis to investigate the influence of those sessions on subsequent collaborative interactions.

Methods

The setting for this study was a regular fifth-grade class in a suburban school district in the southwest US. The 24 ethnically and academically diverse students in the class included 15 girls and 9 boys. Their teacher, Ms. Stevens (all names are pseudonyms), had 20 years of experience and was recognized in the district for her expertise in STEM instruction. Engineering experiences were part of the regular curriculum in this class. Ms. Stevens implemented instruction primarily around engineering challenges utilizing Lego robotic Mindstorms kits. Students engaged in three collaborative engineering challenges across the school-year, working in three-to-four-member groups and changing membership for each project. Projects 1 and 2 were well-structured tasks pre-determined by the teacher to help students learn collaborative engineering knowledge and practices. Project 3 (the focus of this study) was an ill-structured design task: to design, build and program a robot to address an environmental problem identified by each group. Six four-member groups were assigned by the teacher for this project. Three of those groups, each of whose membership was diverse in gender, ethnicity, and academic achievement, were selected for focal analysis. The Water Washer group created a remote-controlled paddle boat to clean lake pollution. The Claw Grabber group designed a wheeled-vehicle with a claw operated via touch-sensors to pick up trash for people with mobility impairments. Finally, the Recycling Rover designed a vehicle with a trailer that operated via a light sensor, transporting recyclable materials down colored paths.

During Project 3, the teacher organized instruction around 14 instructional days across one month, during which small-group work sessions were punctuated by whole-class mini-lessons (just-in-time topics) and teacher check-in meetings. The analysis described here focuses on data associated with the first four days of the project. Day 1 began with each group brainstorming possible projects. On Day 2, after settling on and sketching out their initial designs during a 40-minute work session, the class was called together by Ms. Stevens and DCS1 convened. Each group in turn presented their project progress and received critical feedback and design ideas from their teacher and the rest of the class. Members of a presenting group clustered at the front of the room around a document camera and a computer used to project artifacts on a Smartboard. Physical artifacts came into play as students displayed sketches, structures, and websites to explain and defend their design

decisions. The teacher then led an evaluative discussion, offering questions, critiques and suggestions, and eliciting them from students. Two more DCS were convened over the next two days, interspersed between work sessions during which groups continued to refine and prototype their design ideas.

Data consist of (a) video recordings and transcripts of each DCS, (b) transcripts of work sessions pre and post DCS for three focal groups, (c) expanded observational field notes, (d) transcripts of semi-structured interviews with students and the teacher, and (e) artifacts associated with the project (i.e., photos of structures, written reflections, project wiki pages, computer programs). The first step of analysis was to create content logs of video recordings of DCS and all small-group work sessions of the three focal groups. Next, transcripts of all DCS were created and checked to verify accuracy. I drew on structure-behavior-function (SBF) theory (Hmelo-Silver & Pfeffer, 2004) to categorize the nature of talk generated during the DCS sessions for each group. *Structure* refers to the physical nature of the engineered system that students were designing, including how parts were connected (e.g., “Is that attached to the transporter?”); *behavior* refers to mechanisms that allow the function to be carried out (e.g., “You can program it to read different things on one program, too”); and *function* refers to purpose of the system or a subsystem (e.g., “That’s a large-scale recycling; it’s not just about keeping my house dust free”). Turn-by-turn analysis was conducted to identify the number of comments related to each category that were elicited by the teacher and by students, including statements, ideas, and questions.

The next step of analysis focused on understanding how the talk in DCS influenced subsequent small-group interaction. This step was informed by and sociolinguistic analysis of discourse (Erickson, 2004) and interaction analysis (Jordan & Henderson, 1995) in that is focused on identifying how events were structured, how artifacts and tools were used, and how participants were making sense of each other’s actions as they related to collaborative design processes. In keeping with the goals of these analytical frameworks, I examined social interaction in the focal collaborative groups prior and subsequent to DCS to identify regularities in how participants within and between groups utilized social and material resources available in the context in which they were operating (e.g., ideas generated in DCS). Making extensive memos and annotations, I examined ways in which the three focal groups took up opportunities that were afforded in the three DCS for learning and making progress with their robotic engineering design project. Finally, I identified progress on each group’s project pre- and post-critique design sessions. Throughout analysis, trustworthiness was addressed through prolonged engagement, analysis of multiple data sources to triangulate emerging interpretations, and multiple viewing of recording with peer reviewers.

Findings

Comments Elicited During Design Critiques

Table 1 shows the number of comments offered by the teacher and by students during the design critique sessions as they related to structures, behaviors, and functions of the three focal groups’ design solutions. It thus provides an overview of what was discussed as well as an indication of how discourse about design solutions changed over time. Overall, more talk focused on structures (total of 78 ideas) than behaviors (total 53) or functions (total 44). Teacher comments were more evenly spread across the three categories (48:34:39), while more than half of the student comments pertained to structures (30:19:5). This finding is consistent with past research that novices tend to focus on perceptually available system components (i.e., structures), whereas experts cohesively integrate structural, functional, and behavioral elements (Hmelo-Silver & Pfeffer, 2004). Across the three DCS, the number of SBF comments generated in each session dropped for all three groups. Also, the proportion of teacher-generated ideas decreased over time, relative to student-generated ideas, in all three groups. For the Water Washer and Claw Grabber groups, more comments were elicited from students than from the teacher in DCS3.

Table 1: Ideas elicited across three design critique sessions

	Water Washer Group			Claw Grabber Group			Recycling Rover Group*	
	DC1	DC2	DC3	DC1	DC2	DC3	DC1	DC3
Teacher-generated ideas	Structures	10	16	1	Structures	4	1	1
	Behaviors	8	6	4	Behaviors	5	5	0
	Functions	23	1	1	Functions	6	1	0
Student-generated ideas	DC1	DC2	DC3	DC1	DC2	DC3	DC1	DC3
	Structures	2	8	11	Structures	2	0	1
	Behaviors	3	4	1	Behaviors	1	3	4
	Functions	3	0	0	Functions	1	0	0

*note that the DCS sessions were only convened for the Recycling Rover group two times

The Water Washer group garnered the largest number of SBF comments in all three DCS. Their early struggles to design a robot that met the specification to address an environmental problem were evident in the first DCS, during which the teacher presented 28 ideas related to the function of this group's design ("a robot to heat and cool and clean your pool"). That the group successfully re-tooled the function of their robot subsequent to DCS1 is reflected in a shift in DCS2 and DCS3 toward comments about structural aspects of their design.

The Claw Grabber group garnered the lowest number of SBF comments from the teacher and from their classmates. Their total of 35 ideas was far less than the 110 garnered by the Water Washer group and even less than the 40 garnered by the Recycling Rover Group in just two DCS. Perhaps the teacher and students alike shied away from critiquing this group in which one member, Adriana, became visibly upset after the teacher negatively critiqued the group's original design for a robotic tire pressure reader. Although the teacher ended this group's DCS1 with, "I'm going to have you all work on that a little bit more," the group ultimately dropped the idea. Following DCS1, they went back to the drawing board and brainstormed an initial conception of their Claw Grabber robot. This idea was ratified by the teacher in a small-group meeting prior to DCS2, which might be why the teacher contributed fewer SBF comments during this group's DCS2. The majority of comments offered to this group during DCS2 were about behaviors as the teacher questioned how the group could make their robot turn when they had only three motors (a design constraint), and the class struggled to understand how a person with mobility impairments could operate the remote control. Many comments during DCS2 expressed doubt about how the group planned to use a sound sensor as part of their propulsion system, an idea the group later abandoned when the programming proved too difficult.

How Ideas Were Taken Up

Ideas from the design critiques were taken up by all three focal groups following each DCS in which a group presented, though not always in expected or productive ways. Across the three DCS, groups seemed to increase their ability to take up teacher and peer feedback from DCS. Additionally, the largely positive affect with which they interpreted their DCS experiences was reflected in subsequent interviews. To illustrate, the following excerpt shows the integrated response of the Water Washer group to a suggestion made by a classmate about a structure associated with their design solution (i.e., the shape of their paddles).

Excerpt 1 (Water Washer group post-DCS3)

- Ida: Adam raised his hand and said how about make them oval shapes, so that's what I think we should do.
- Bobby: He said it should be oval.
- Dante: That's not oval.
- Interviewer: What did you guys think when Ms. Stevens asked the class if they had any suggestions?
- Ida: I thought there would be a lot more people giving suggestions.
- Roy: I like Adam's//
- Dante: //I didn't think there would be a whole lot of suggestions though.
- Roy: Me neither, but I liked that they were giving us ideas. But I think the tire might be a little heavy.
- Dante: Yeah, because I think it'll spin and then this would leak, push it through the water faster and have more power so it'll go faster. And then I started to agree with my group because tires they're too thick and they make too much splash and it might not move through the water.
- Roy: Yeah, I think it was a good idea, but I think we should stick to our sticks, but//
- Ida: //It's a good idea, but still, that's more weight adding up, though.
- Roy: Yeah.
- Ida: It's just more weight adding up. It was a good idea though.
- Bobby: Yes it was.
- Roy: It was a very good idea.

Conclusions and Implications

This study is one step in a line of scholarly inquiry investigating the influence of whole-class social interaction in learning to engage in collaborative design projects. Findings contribute to understanding how talk elicited in DCS influence the design activity of early adolescents learning to engage in collaborative engineering design by categorizing the nature of talk in DCS by the teacher and students relative to the nascent design solutions of three focal groups, and by examining the subsequent discourse in small-group work sessions to see what ideas students took up as well as how they took them up and to what effect. The ultimate value of the study is its potential to impact learning-teaching practices related to engaging in collaborative engineering design processes. For instance, results could guide strategies for (a) structuring public critique processes to enhance students' learning to design and learning from design, and (b) helping students make use of critique in subsequent collaborative sessions. Findings may also illuminate learning-teaching processes in art settings, composition classes, or design studios - contexts in which public design critique is a long-standing practice.

References

- Bers, M. U. (2008). *Blocks to robots: Learning with technology in the early childhood classroom*. New York: Teachers College Press.
- Bransford, J. D., Brown, A., & Cocking, R. (1999). *How people learn: Brain-mind experience and school*. Washington, DC: National Academy Press.
- Darling, A. L., & Dannels, D. P. (2008). A report on the role of oral communication in the workplace. *Communication Education*, 52, 1-16.
- Demirbas, O. (2001). *The relation of learning styles and performance scores of the students in interior architecture education*. Unpublished PhD dissertation, Bilkent University, Ankara.
- Dorst, K. (2003). The Problem of design problems. *Design thinking research symposium, Sydney*, 17(19-11).
- Erickson, F. (2004). *Talk and social theory*. Malden, MA: Polity Press.
- Fleming, D. (1998). Design talk: Constructing the object in studio conversations. *Design Issues*, 14(2), 41-62.
- Greeno, J. G. (2006) Learning in activity. In K. Sawyer (Ed.) *Handbook of the Learning Sciences* (pp. 79-96). Oxford: Cambridge.
- Greiffenhagen, C. (2012). Making rounds: The routine work of the teacher during collaborative learning with computers. *International Journal of Computer-supported Collaborative Learning*, 7(1), 11-42.
- Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to learn about complex systems. *Journal of the Learning Sciences*, 9 (3), 247-298.
- Hmelo-Silver, C. E. & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, 28, 127-138.
- Jonassen, D. H. (2000). Towards a design theory of problem solving. *Educational Technology, Research and Development*, 48(4), 63-85.
- Jordan, B. & Henderson, A. (1995). Interaction Analysis: Foundation and Practice. *Journal of the Learning Sciences*, 4, 39-103.
- Jordan, M. E. (in press). Exploring how public design critique processes shape fifth graders' peer interaction during collaborative engineering projects. *ASEE proceedings*, Indianapolis, IN.
- Jordan, M. E. (2014). Interweaving the digital and physical worlds in collaborative project-based learning experiences. In D. J. Loveless, B. Griffith, M. Berci, E. Ortlieb, & P. Sullivan (Eds.), *Academic knowledge construction and multimodal curriculum development* (pp. 266-284). Hershey, PA: IGI.
- Jordan, M. E. & Babrow, A. S. (2013). Communication in creative collaborations: The challenges of uncertainty and desire related to task, identity, and relational goals. *Communication Education*, 62(2), 105-126.
- Jordan, M. E. & McDaniel, R. R. (in press). Managing uncertainty during collaborative problem solving in elementary school teams: The role of peer influence in robotics engineering activity. *Journal of the Learning Sciences*.
- Kafai, Y. B., & Ching, C. C. (2004). Children as instructional designers: Apprenticing, questioning, and evaluating in the Learning Science by Design project. In N. M. Seel & S. Dijkstra (Eds.), *Curriculum, plans and processes of instructional design* (pp. 115-130). Mahwah, NJ: Lawrence Erlbaum.
- 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.
- Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., & Holbrook, J. (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(4), 495-547.
- Laursen, S., Liston, C., Thiry, H., & Graf, J. (2007). What good is a scientist in the classroom? *CBE-Life Sciences Education*, 6(1), 49-64.
- National Research Council. (2009). *Engineering in K-12 education: Understanding the status and improving the prospects*. Washington, DC: The National Academies Press.
- Otto, K. N. & Wood, K. L. (2000). *Product design techniques in reverse engineering and new product development*. Upper Saddle River, NJ: Prentice Hall.
- Pendleton-Julian, A. & Brown, J. S. (2011). *Design unbound: Evolving design literacy pathways of efficacy*. Draft manuscript, Charleston, SC.
- Papert, S. (1980). *Mindstorms: children, computers, and powerful ideas*. New York: Basic Books.
- Rogoff, B. (1990). *Apprenticeship in thinking*. New York: Oxford University Press.
- Roth, W.-M. (1995). Inventors, copycats, and everyone else: The emergence of shared resources and practices as defining aspects of classroom communities. *Science Education*, 79, 475-502.
- Sawyer, K. (2012). Learning how to create: Toward a learning sciences of art and design. *Proceedings of the International Conference of the Learning Sciences* (pp. 33-40). Sydney, Australia.
- Shaffer, D. W. (2003). *Portrait of the Oxford design studio: An ethnography of design pedagogy* (Working paper No. 2003-11). Madison, WI: Wisconsin Center for Education Research.
- Uluoglu, B. (2000). Design knowledge communicated in studio critiques. *Design Studies*, 21(1), 33-58.

Tensions in a Multi-Tiered Research-Practice Partnership

Samuel Severance
School of Education
University of Colorado Boulder
Boulder, CO 80309-0249, USA
samuel.severance@colorado.edu

Heather Leary
Institute of Cognitive Science
University of Colorado Boulder
Boulder, CO 80309-0344, USA
heather.leary@colorado.edu

Raymond Johnson
School of Education
University of Colorado Boulder
Boulder, CO 80309-0249, USA
raymond.johnson@colorado.edu

Abstract: Bridging the research-practice gap remains an important focus of much learning research. An important challenge is for researchers to organize design to leverage both the expertise of practitioners and that of researchers in an equitable arrangement. This study examines such an arrangement by analyzing the joint work of a design-based research-practice partnership. It focuses on the *design tensions* (Tatar, 2007) associated with coordinating the needs of participants from three different tiers of design—involving teachers, researchers, and district administrators—related to the content of designs and to the mechanisms for bringing content to scale within the district. This study argues for the value of a common vision and design methodology to enable design tensions at multiple levels to become generative influences on design.

Major Issues Addressed

Given the seemingly intractable nature of some of the problems facing education today, it seems reasonable to approach these problems in a different manner in hopes of developing a novel solution. One such approach includes developing expertise horizontally across multiple dimensions rather than vertically within a single dimension for individuals and communities (Engeström, Engeström, & Karkkainen, 1995; Gutiérrez, Hunter, & Azurbiaga, 2009). In such an arrangement, those who possess the ability to leverage knowledge from different domains and viewpoints will more likely develop a novel solution than those who approach the problem with vertical or isolated expertise (Engeström et al., 1995).

Leveraging expertise is a central challenge in design. Design research is premised on the idea that multiple forms of expertise are needed to develop powerful, usable, and practical innovations to improve learning (Collins, Brown, & Bielaczyc, 2004). Teachers are often participants in design research, most often as implementers of designs (Ormel et al., 2012). Sometimes, their involvement is structured to be more equitable, as in co-design (Penuel, Roschelle, & Shechtman, 2007).

Research-practice partnerships (RPPs) are long-term, mutualistic collaborations between practitioners and researchers that are intentionally organized to investigate problems of practice and solutions for improving district outcomes (Coburn, Penuel, & Geil, 2013). RPPs are aimed at increasing the relevance and usefulness of research for practitioners to support instructional improvement at scale (National Research Council, 2003). Some RPPs are *design research partnerships*, that is, they are organized with the aim of designing innovations that can have an impact on entire educational systems, such as school districts.

This study uses a *design tensions framework* (Tatar, 2007) to analyze how the organization of a design research partnership successfully leverages horizontal arrangements of expertise in order to solve complex problems in a way that accounts for multiple perspectives. Studying RPPs can provide practical insights into how the research-practice gap between traditionally insular research universities and traditionally isolated school practitioners can be made narrower through joint activity. Such efforts, however, rarely unfold without complications. This study seeks to highlight the effect of tensions on the design process within the RPP, the benefit of a shared vision among participants, and routines for identifying alternate perspectives on design in partnership work. Specifically, this study examines how tensions of scaling can still become issues of productive design rather than disrupting design. As such, its primary aim is to contribute to our understanding of the practices of design in the learning sciences, particularly how these practices can be organized to effect improvement at scale.

Contextualization

During the past two years, researchers from the University of Colorado Boulder and a large research-non-profit in conjunction with leaders and teachers from a large urban school district in the northeastern United States have engaged in joint work as part of a RPP funded by the National Science Foundation through a multi-year grant. The overall aim of this partnership seeks to support the instructional goals of the district through the design and implementation of a digital platform called EdTrex. EdTrex will allow for the customization of the district's curriculum in 9th grade Algebra, enhancing it with high-quality mathematical tasks aligned to the Common Core State Standards for Mathematics (CCSS-M) while simultaneously increasing teachers' understanding of the CCSS-M through professional development as part of the co-design process.

The RPP under study here has an interwoven “multi-tiered” structure. Figure 1 presents a schematic of the overlapping organization of design activity within the RPP. One tier, dubbed the “leadership tier,” involves regular meetings between district leaders, researchers, and designers to plan activities for the year in regards to the customization of the district’s Algebra curriculum through the EdTrex platform as well as determine avenues for eventually bringing the EdTrex project to scale across the entire district. Another tier involves regular meetings between the project leadership tier and a Teacher Design Team (TDT) comprised of other practitioners in the district in order to engage them in co-design activities for EdTrex. Yet another group meets regularly to discuss the research and engineering aspects of the project. Researchers and designers from the university and research non-profit participate in these sessions.

Despite the potential affordances of direct collaboration between education researchers and education practitioners, productive and sustained examples of such deeply collaborative partnerships remain rare. RPPs provide a collaborative structure with which to implement in practice a means of addressing this lack of coordination of expertise between schools and academia. A need for understanding how such RPPs engage in joint work—notably the nature of breakdowns and tensions that arise and affect the efficacy of collaboration as well as possible avenues for their amelioration—still remains. For those in the learning sciences community pursuing similar collaborations around bridging the research-practice gap, this study offers fundamental issues of collaborative design research to consider. Specifically, this study seeks to add to our understanding of design tensions framework methodology through its analysis of tensions in a multi-tiered, tripartite RPP participating in joint design work.

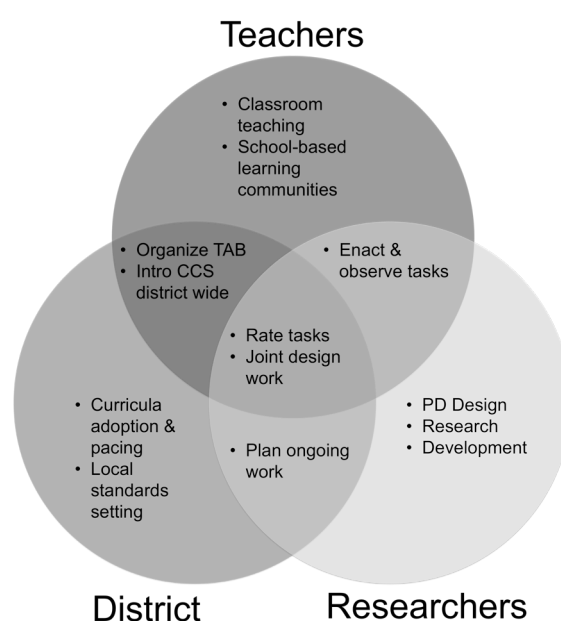


Figure 1. Overlapping activity structure of the multi-tiered EdTrex RPP.

Theoretical Framework

Focusing on the joint design work undertaken by the three tiers of the RPP, particularly the breakdowns and tensions that occur within the collaboration, this study utilized a *design tensions framework* (Tatar, 2007) to guide data analysis. A design tensions framework, fundamentally, posits that “design exists because of the tension between what *is* and what *ought to be*” (Tatar, 2007, p.415). Tensions become the source of design decisions or choice points; design becomes no longer about solving problems but about goal balancing and seeking optimal compromises within the system (Tatar, 2007). Such an approach allows for more design possibilities because it transcends the rigid solution choices that typify a design space and instead seeks unbounded, crosscutting design possibilities to create a solution space of greater flexibility (Tatar, 2007). Within this study, a design tensions framework affords maneuverability and guidance in understanding and addressing design tensions at multiple levels. In addition, making design tensions explicit facilitates conversations among participants about diverse perspectives on design and whether these different perspectives are being adequately taken up in deliberations.

Design tensions can occur at four different levels. The highest level of tension, *vision*, describes the state of an incongruity between “what is and what ought to be” (Tatar, 2007, p.417), which underscores the fact that design is a goal-driven, value-laden enterprise. The tension found on the next level, *approach*, centers on bringing the values that make up the vision of what “ought to be” into the current state of reality (Tatar, 2007).

Here, much recursive design work occurs since participants have the volition to choose and alter approaches to address design tensions (Tatar, 2007). Below *approach*, we find the *project tensions* associated with the actual decisions of implementation to enact the approach that either fall under the designers' influence or have become a point of contention between participants (Tatar, 2007). Finally, the last design tension, designated "*as created*" situations (Tatar, 2007 p.418), describes the fallout from any course of action taken. In other words, though an action may resolve one tension, new tensions will inevitably result from taking the aforementioned action.

Methodological Approaches

A descriptive case study of dynamics within a single RPP, this study applied an ethnographic approach to data collection and focused on the interactions among and between members of the research practice partnership. Participants included 11 TDT members, 3 district administrators, and 6 researchers. The authors of the study collected data from the vantage point of a participant observer during all collaborative design work. Sources of data for interactions between and among the groups include observations of 21 regular teleconference and face-to-face meetings, analysis of correspondence that spanned 10 months between groups via e-mail, analysis of 15 semi-structured interviews with TDT members, and analysis of various artifacts related to both the meetings themselves, such as handouts and agendas, and artifacts related to the district aims of the joint collaboration.

Substantiation

Throughout the joint work of the RPP, the participants engaged with all four levels of design tensions put forth by Tatar (2007). Table 1 provides a summary of the different levels of design tensions encountered. How the group engaged with these design tensions greatly determined the course and effectiveness of the design process. This study argues for the value of a common vision and design methodology in order to enable design tensions to become generative influences on design. Below, how each level of design tensions influenced the design process within the RPP is discussed.

Values

Within this collaborative endeavor, the vast majority of participants shared a common understanding of the design tension of *vision*, of "what is and what ought to be," in regards to the overall telos of the EdTrex project (Tatar, 2007). With few exceptions, participants from all levels of the multi-tier RPP felt a desire to revise the current 9th grade Algebra curriculum. As summarized by a researcher speaking at an early meeting to fellow participants, this task called on RPP members to "digitally enhance their curriculum," to "increase the richness of mathematical tasks" and "increase student engagement." A district administrator at the same meeting echoed this enthusiasm by referring to the EdTrex platform as "the place to go...not just for us but for all of the city." Generally, the 11 TDT members expressed enthusiasm with doing curricular work during interviews with only one practitioner remarking "to be completely honest the curriculum development type of thing is actually the aspect of teaching that I'm least interested in." This alignment of vision, a shared sense through all levels and groups in the RPP of what "ought to be," allowed increased cohesiveness in the direction of design and provided a common touchstone for participants.

Approach

In terms of the design tensions regarding *approach*, this area too saw much alignment. Though the leadership tier of the RPP, comprised of the district administrators and researchers, devised the overall framework of how participants went about co-design activities within meetings, etc., teachers nonetheless endorsed the general approach with one teacher even exclaiming during a meeting, "I'm technically a designer now too, aren't I?" As outlined in the theory of the design tensions framework, this alignment of approach should not prove surprising (Tatar, 2007). Since the values portion of the design tension *approach* matches with the "ought" portion of the higher-level *vision* design tension, and having already established a congruency amongst participants in the RPP in that regard, one would expect similar attitudes from participants regarding the approaches applied.

Project Tensions

Though numerous *project tensions* did arise, we will examine only one in detail here (see Table 1 for a list of other project tensions). Project tensions proved both the most problematic in terms of affecting the cohesiveness of the RPP but also the most useful in terms of framing design choices. With one of the agreed-upon goals of the project stipulating that the EdTrex platform scale across the district, much of the design attended to this effort. As mentioned previously, design is a value-laden enterprise and participants apply those values in conjunction with their expertise when making design choices.

Though participants seemingly all valued the idea of bringing the EdTrex to scale, the decisions of implementation to achieve this goal surfaced different, equally valid concerns. The district administrator group of the RPP adopted a more conservative approach to scaling by gating researcher access to teachers whereas

researchers desired a more aggressive approach of direct access to teachers at their school site to support scaling and EdTrex. The district administrators' reasoning seems to stem from their expertise of each school's unique situation, demonstrated through their impressive knowledge of the personnel at each site on more than one occasion, and a likely concern for teacher autonomy as reflected in their rebuffing of an offer from the researchers to assist with a new teacher training, saying in an e-mail the teachers "are fine with handling this introduction to EdTrex for the new Algebra 1 teachers themselves."

In keeping with the tenets of a design tensions framework, participants navigated this tension around scaling by attempting to balance the goals of participants to reach an optimal compromise (Tatar, 2007). Recognizing a need of teachers for supporting the implementation of tasks along certain dimensions, particularly language supports for emerging bilingual students, the leadership tier of the RPP realized that the EdTrex could be the "delivery vehicle" to provide teachers across the district with the instructional supports they need. Proposing that TDT teachers could broker the distribution of materials, one district administrator stated during a meeting, "[W]e'll focus on where we have a TDT member to do that and then where we don't have a TDT member..." "Try to get one," completed a researcher. Scaling the EdTrex more quickly now seemed more desirable. Even though researchers did not gain direct access to teachers, they still seemed agreeable to the compromise since in the words of one researcher during the same meeting, "More TDT teachers would help to spread the work load," and, according to another researcher "in some of these schools it's kind of fuzzy to see how things operate, but as we learn more about it then hopefully that will help our scaling issue." Current plans are for the RPP to co-design professional development task support modules that TDT members can share with their colleagues at their sites.

"As Created" Situations

As predicted by the theory of the design tensions framework, the enactment of approaches to alleviate design tensions in turn created new tensions falling into the category of "*as created*" situations (Tatar, 2007). Similar to the previous section regarding *project tensions*, numerous "*as created*" situation tensions surfaced in analysis of the data. The most prominent will receive discussion here (see Table 1 for a list of other "as created" situation design tensions). From interview and survey data, clearly, the approach conceived of by the leadership tier of the RPP placed a large time burden on members of the TDT. When asked, "What challenges did you face in participating in EdTrex?" fully seven of the nine teachers who answered this survey item referred to "time" being an issue. Specifically, one teacher responded, "Finding the time to analyze and go through the tasks on top of my current load of work." Undoubtedly, attempts to remedy this situation will lead to the creation of a new design tension. Perhaps, however, an optimal compromise agreeable to all participants of the RPP can be achieved (Tatar, 2007).

Table 1: Design tensions within the EdTrex RPP

Vision	Is: Current Algebra curriculum does not align to the CCSS-M and has tasks of low quality	Ought: Revised Algebra curriculum has high-quality mathematical tasks aligned to the CCSS-M; wide adoption; sustainable
Approach	Project Drivers: Co-design of digital platform, selection and rating of tasks, implementation of tasks in classroom	Values: Rigorous curriculum, equitable design partnership, high utility of tool
Project Tensions	Tension 1	Gated Access to Practitioners vs. Direct Access to Practitioners
	Tension 2	Practice Tasks vs. High Cognitive Demand Tasks
	Tension 3	Teacher Pace on Scope & Sequence vs. District Guidelines
"As created" Situations	High demands on teacher time; teachers disagreeing with EdTrex content and leaving RPP; too few tasks in some areas of EdTrex	

Conclusion

This study examined the design process of a multi-tiered, tripartite RPP by collecting data using a participant observer ethnographic approach and analyzing data through the theoretical lens of a design tensions framework (Tatar, 2007). Central to the contribution of this study is an enhanced understanding of the practice of design, particularly in relation to the unique multi-tiered structure of the RPP. This study highlighted the importance of

participants at all levels of a multi-tiered RPP sharing a common understanding of the *vision* design tension. As mentioned previously, a common *vision* within this framework posits that the *approach* design tension shall also demonstrate congruency amongst participants. Within this multi-tiered RPP, such alignment allowed participants to leverage their expertise synergistically in designing the EdTrex platform and avoided devolving into directionless design. Additionally, having a shared understanding amongst participants at the highest levels of the design tensions hierarchy supported the navigation of design tensions found at the *project tensions* level as participants engaged in balancing the goals of members of the RPP towards an optimal compromise.

Additionally, analysis in this study suggests potential mechanisms for learning scientists and design researchers to consider in terms of how the structure of a multi-tiered, tripartite RPP can support effective co-design. In the interweaving multi-tiered participation model examined, the intentional allowance for interaction between different participant groups within different tiers in the RPP increased the likelihood of surfacing more design tensions and revealed areas for the RPP to jointly attend to across multiple participant groups. The design tensions served as opportunities for participants to make potentially productive design decisions since the tensions themselves represented needs to be addressed in the design of a co-designed object. In sum, this suggests that fostering interaction between participant groups within different tiers can lead to more unearthing of design tensions (particularly at the “as created” situations level), which in turn leads to more possible opportunities to engage in potentially productive design decisions across multiple levels. Therefore, the possibility that multi-tiered participation structures are more generative in terms of pushing design than other structures seems a plausible avenue worthy of further exploration.

Lastly, while bringing together groups of participants with differing expertise into a horizontal arrangement, as found in a multi-tiered RPP, may increase the likelihood of a lack of cohesiveness in shared understandings of design tensions, it can also allow for more maneuverability in reaching compromises amongst groups. As illustrated in the project tension around scaling in this study, with more participants—each with differing expertise, roles, and needs within a system—more capacity for negotiation of an optimal arrangement develops as particular groups’ interests can overlap productively in a complementary manner with others’ interests in ways they could not have with less varied participant groups.

Relevance to Conference Theme

In a fundamental manner, this study demonstrates the conference theme of “learning and becoming in practice.” Indeed, a research practice partnership (RPP) epitomizes the notion of applying what is learned—either at the level of the learning researcher or at the level of the learning practitioner but more likely all of the above—towards the creation of a co-designed object, the express purpose of which is to positively impact the realm of practice at multiple levels. The inclusive nature of a multi-tiered RPP serves as an exemplar of how the practice of learning research can intersect with the practice of others, bettering all involved.

References

- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, 13(1), 15-42.
- Coburn, C. E., Penuel, W. R., & Geil, K. (2013). *Research-practice partnerships at the district level: A new strategy for leveraging research for educational improvement*. Berkeley, CA and Boulder, CO: University of California and University of Colorado.
- Engeström, Y., Engeström, R., & Karkkainen, M. (1995). Polycontextuality and boundary crossing in expert cognition: Learning and problem solving in complex work activities. *Learning and Instruction*, 5(4), 319–336.
- Gutiérrez, K., Hunter, J., & Arzubíaga, A. (2009). Re-mediating the university: Learning through sociocritical Literacies. *Pedagogies: An International Journal*, 4, 1-23.
- National Research Council. (2003). *Strategic education research partnership*. Washington, DC: National Research Council.
- Ormel, B. J. B., Roblin, N. N. P., McKenney, S. E., Voogt, J. M., & Pieters, J. M. (2012). Research-practice interactions as reported in recent design studies: Still promising, still hazy. *Educational Technology Research and Development*, 60(6), 967-986.
- Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). The WHIRL co-design process: Participant experiences. *Research and Practice in Technology Enhanced Learning*, 2(1), 51-74.
- Tatar, D. (2007). The Design tensions framework. *Human-Computer Interaction*, 22, 413-451.

Acknowledgments

This work was supported by a grant from the National Science Foundation (Award #1147590). The opinions expressed here are those of the authors and do not necessarily reflect those of the NSF. The authors are grateful to teachers and administrators in our partner district for their support and contributions to this research and development effort.

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

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, Tsafir, 1362
 Goldin, Simona, 1505
 Goldman, Shelley, 1228
 Goldman, Susan R., 1541, 1571
 Goldwasser, Lloyd, 689
 Gomez, Kimberley, 1671, 1700
 Gomez Zaccarelli, Florencia, 1724
 Gong, Bowye, 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, Nosakhare, 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
 Holding, Brandon, 1527
 Heredia, Sara, 1727
 Herrenkohl, Leslie Rupert, 1152, 1247
 Hertel, Jonathan D., 1587
 Hetmanek, Andreas, 1189
 Hever, Rakheli, 1342
 Heyd-Metzuyanin, 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, Tuuke, 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
 Mejia-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
 Mohny, 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, Jène, 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
 Sottolare, 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, Brea, 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