

CSCCL 2013

Computer Supported Collaborative Learning

June 15–19, 2013

University of Wisconsin – Madison



To see the world *and* a grain of sand:
Learning across levels of space, time and scale

Conference Proceedings Volume I Full Papers & Symposia

Editors: Nikol Rummel, Manu Kapur, Mitchell Nathan, Sadhana Puntambekar

Sponsored By



Hosted By



To See the World *and* a Grain of Sand:
Learning across Levels of Space, Time, and Scale

CSCCL 2013 Conference Proceedings
Volume 1
Full Papers & Symposia

10th International Conference on Computer-Supported Collaborative Learning
June 15-19, 2013, Madison, WI

Title:

To See the World *and* a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013 Conference Proceedings Volume 1 — Full Papers & Symposia

Editors:

Nikol Rummel, Manu Kapur, Mitchell Nathan, Sadhana Puntambekar

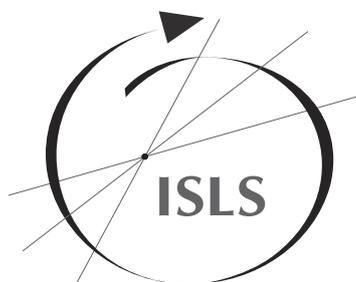
Citation format:

Rummel, N., Kapur, M., Nathan, M., & Puntambekar, S. (Eds.) (2013). *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013 Conference Proceedings Volume 1 — Full Papers & Symposia*. International Society of the Learning Sciences.

ISBN:

© 2013, INTERNATIONAL SOCIETY OF THE LEARNING SCIENCES [ISLS]

Copyright 2013 International Society of the Learning Sciences, Inc. Rights reserved



**International Society of
the Learning Sciences**

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without the 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. <http://www.isls.org/>

Published by:

International Society of the Learning Sciences (ISLS)

Proceedings Printed by:

LuLu <http://www.lulu.com>

Proceedings Distributed by:

LuLu and Amazon

Proceedings Created by:

Dana Gnesdilow, Julia Rutledge, Amanda Evenstone, Matthew Gaydos, Garrett Smith, and Sadhana Puntambekar

Book Cover Design:

Allyson Casey

Sponsorship

CSSL 2013 would like to thank our sponsors:



**International Society of
the Learning Sciences**



School of Education
UNIVERSITY OF WISCONSIN-MADISON



WCER
WISCONSIN CENTER FOR EDUCATION RESEARCH

in?scribe
• simple transcription + subtitling



Steering Committee

Conference Chair

Sadhana Puntambekar, University of Wisconsin, Madison, USA

Program Co-Chairs

Nikol Rummel, Ruhr-Universität, Bochum, Germany

Manu Kapur, National Institute of Education, Singapore

Mitchell Nathan, University of Wisconsin, Madison, USA

International Relations Co-Chairs

Chris Hoadley, New York University, USA

Nancy Law, University of Hong Kong, China

Sten Ludvigsen, InterMedia, University of Oslo, Norway

Preconference Workshops and Tutorials Co-Chairs

Gijsbert Erkens, Utrecht University, Netherlands

Chee-Kit Looi, Nanyang Technological University, Singapore

Demos Co-Chairs

Richard Halverson, University of Wisconsin, Madison, USA

Hans Spada, University of Freiburg, Germany

Daniel Suthers, University of Hawai'i, USA

Doctoral Consortium Co-Chairs

Frank Fischer, Ludwig-Maximilians-Universität München, Germany

Heisawn Jeong, Hallym University, South Korea

Erica Halverson, University of Wisconsin, Madison, USA

Rosemary Luckin, Institute of Education, University of London, England

Early Career Workshop Co-Chairs

Kristine Lund, University of Lyon, France
Carolyn Rose, Carnegie Mellon University, USA

Iris Tabak, Ben Gurion University of the Negev, Israel

Special sessions

Cindy Hmelo-Silver, Rutgers University, USA

Bilge Mutlu, University of Wisconsin, Madison, USA

Consultants

Pierre Dillenbourg, École Polytechnique Fédérale de Lausanne, Switzerland

Susan Goldman, University of Illinois, Chicago, USA

Paul Kirschner, Open University, Netherlands

Tim Koschmann, Southern Illinois University, USA

Naomi Miyake, University of Tokyo, Japan

Peter Reimann, University of Sydney, Australia

Nancy Butler Songer, University of Michigan, USA

Gerry Stahl, Drexel University, USA

Reviewers of CSCL 2013 Conference

The conference chair and program chairs sincerely thank the following scientists who reviewed proposals for the CSCL 2013 conference.

- Abrahamson, Dor, University of California, United States
- Acholonu, Ugochi, Stanford University, United States
- Alterman, Richard, Brandeis University, United States
- Alvarez, Isabel, Autonomous University of Barcelona, Spain
- Andrea, Kienle, University of Applied Sciences, Germany
- Anne, Deiglmayr, Eidgenössische Technische Hochschule Zürich, Switzerland
- Armin, Weinberger, Saarland University, Germany
- Arnseth, Hans Christian, University of Oslo, Norway
- Asensio, Juan I., University of Valladolid, Spain
- Asterhan, Christa, Hebrew University of Jerusalem, Israel
- Avouris, Nikolaos, University of Patras, Greece
- Bader-Natal, Ari, Grockit, Inc., United States
- Baghaei, Nilufar, United Institute of Technology, New Zealand
- Bairral, Marcelo, Federal Rural University of Rio de Janeiro, Brazil
- Baker, Michael, CNRS - Telecom ParisTech, France
- Bauters, Merja, Helsinki Metropolia University of Applied Sciences, Finland
- Belland, Brian, Utah State University, United States
- Bennerstedt, Ulrika, University of Gothenburg, Sweden
- Berland, Matthew, University of Texas San Antonio, United States
- Blake, Canan, Open University, United Kingdom
- Blavier, Adelaide, University of Liège, Belgium
- Bodemer, Daniel, University of Duisburg-Essen, Germany
- Bonsignore, Elizabeth, University of Maryland College Park, United States
- Bouyias, Yannis, Aristotle University of Thessaloniki, Greece
- Brahm, Taiga, University of St. Gallen, Switzerland
- Bratitsis, Tharrenos, University of Western Macedonia, Greece
- Brennan, Karen, Massachusetts Institute of Technology, United States
- Buckingham Shum, Simon, Open University, United Kingdom
- Buder, Juergen, Knowledge Media Research Center, Germany
- Bures, Eva, Bishop's University, Canada
- Cakir, Murat, Middle East Technical University, Turkey
- Carell, Angela, Ruhr - University of Bochum, Germany
- Carletti, Laura, Horizon - University of Nottingham/University of Exeter, United Kingdom
- Chan, Carol, University of Hong Kong, China
- Chan, Margaret, Columbia University, United States
- Chan, Tak-Wai, National Central University, Taiwan
- Chang, Ben, National Chiayi University, Taiwan
- Chapman, Robbin, Massachusetts Institute of Technology, United States
- Charles, Elizabeth, Dawson College/Georgia Institute of Technology, United States
- Chau, Clement, Tufts University, United States
- Chen, Ching-Huei, National Changhua University of Education, Taiwan
- Chen, Chiu-Jung, National Chiayi University, Taiwan
- Chen, Wenli, National Institute of Education, Singapore
- Chen, Gaowei, University of Pittsburgh, United States
- Cherniavsky, John, National Science Foundation, United States
- Ching, Yu-Hui, Boise State University, United States
- Cierniak, Gabriele, Knowledge Media Research Center, Germany
- Clarke-Midura, Jody, Massachusetts Institute of Technology, United States
- Cober, Rebecca, University of Toronto, Canada
- Condamines, Thierry, Université de Picardie Jules Verne, France
- Correia, Ana-Paula, Iowa State University, United States
- Cress, Ulrike, Knowledge Media Research Center, Germany

- D'Angelo, Cynthia, SRI International, United States
 Damsa, Crina, University of Oslo, Norway
 Dascalu, Mihai, University Politehnica of Bucharest, Romania
 Dasgupta, Chandan, University of Illinois Chicago, United States
 de Jong, Frank, University of Applied Sciences and Teacher Education, Netherlands
 de Leng, Bas, Maastricht University, Netherlands
 De Wever, Bram, Ghent University, Belgium
 DeJaegher, Crystal, University of Virginia, United States
 Demetriadis, Stavros, Aristotle University of Thessaloniki, Greece
 Demmans Epp, Carrie, University of Saskatchewan, Canada
 Derry, Sharon, University of Wisconsin Madison, United States
 Di Blas, Nicoletta, Politecnico di Milano, Italy
 Dillenbourg, Pierre, École Polytechnique Fédérale de Lausanne, Switzerland
 Dimitriadis, Yannis, University of Valladolid, Spain
 Ding, Jie, Beijing Normal University, China
 DiSalvo, Betsy, Georgia Institute of Technology, United States
 Dowell, John, University College London, United Kingdom
 Dugstad Wake, Jo, University of Bergen, Norway
 Duh, Henry Been-Lirn, National University of Singapore, Singapore
 Duncan, Sean, Indiana University, United States
 Dyke, Gregory, Ecole Nationale Supérieure des Mines de Saint-Etienne, France
 Eberle, Julia, University of Munich, Germany
 Eliot, Matt, CQUniversity Australia, Australia
 Erkens, Gijsbert, Utrecht University, Netherlands
 Ertl, Bernhard, Universität der Bundeswehr München, Germany
 Evans, Michael, Virginia Tech, United States
 Feldmann, Birgit, University of Hagen, Germany
 Ferreira, Deller, Federal University of Goias, Brazil
 Fesakis, Georgios, University of Aegean, Greece
 Fields, Deborah, University of California Los Angeles, United States
 Filsecker, Michael, Duisburg-Essen University, Germany
 Fischer, Frank, University of Munich, Germany
 Forte, Andrea, Drexel University, United States
 Fujita, Nobuko, University of Toronto, Canada
 Garzotto, Franca, Politecnico di Milano, Italy
 Gegenfurtner, Andreas, University of Munich, Germany
 George, Sebastien, Institut National des Sciences Appliquées de Lyon, France
 Goggins, Sean, Drexel University, United States
 Gogoulou, Agoritsa, University of Athens, Greece
 Gomes, Alex Sandro, Universidade Federal de Pernambuco, Brazil
 Gouli, Evangelia, University of the Aegean/University of Athens, Greece
 Grant, Jamillah, Northcentral University, United States
 Gressick, Julia, Indiana University South Bend, United States
 Grigoriadou, Maria, University of Athens Panepistimiopolis, Greece
 Guribye, Frode, University of Bergen, Norway
 Haake, Joerg, FernUniversität in Hagen, Germany
 Hackbarth, Alan, University of Wisconsin Madison, United States
 Halverson, Erica, University of Wisconsin Madison, United States
 Hassman, Katie, Syracuse University, United States
 Hayama, Tessai, Kanazawa Institute of Technology, Japan
 Hernandez, Juan Carlos, Universidad Nacional Abierta y a Distancia, Colombia
 Herrmann, Thomas, University of Bochum, Germany
 Hesse, Friedrich, Knowledge Media Research Center, Germany
 Hirashima, Tsukasa, Hiroshima University, Japan
 Hmelo-Silver, Cindy, Rutgers University, United States
 Hod, Yotam, University of Haifa, Israel
 Hoidn, Sabine, Harvard University, United States
 Hong, Kian Sam, Universiti Malaysia Sarawak, Malaysia
 Hong, Huang-Yao, National Chengchi University, Taiwan
 Hoppe, H. Ulrich, University of Duisburg-Essen, Germany
 Horn, Michael, Northwestern University, United States
 Horney, Mark, University of Oregon, United States
 Hou, Huei-Tse, National Taiwan University of Science and Technology, Taiwan
 Hsu, Ching-Kun, National University of Tainan, Taiwan
 Hubscher, Roland, Bentley University, United States
 Hakkinen, Paivi, University of Jyväskylä, Finland
 Ioannou, Andri, Cyprus University of Technology, Cyprus
 Jahnke, Isa, Umeå University, Sweden
 Jeong, Heisawn, Hallym University, Republic of Korea
 Jermann, Patrick, Ecole Polytechnique Federale de Lausanne, Switzerland

- Joiner, Richard, University of Bath, United Kingdom
- Jones, Christopher, Liverpool John Moores University, United Kingdom
- Juang Yih-Ruey, Jinwen, University of Science and Technology, China
- Järvelä, Sanna, University of Oulu, Finland
- Kafai, Yasmin, University of California Los Angeles, United States
- Karakostas, Anastasios, Aristotle University of Thessaloniki, Greece
- Karasavvidis, Ilias, University of Thessaly, Greece
- Khine, Myint Swe, Emirates College for Advanced Education, United Arab Emirates
- Kim, Mi Song, Nanyang Technological University, Singapore
- Kim, Kibum, Virginia Tech, United States
- Kim, Beaumie, University of Calgary, Canada
- Kimmerle, Joachim, University of Tuebingen, Germany
- King Chen, Jennifer, University of California, United States
- Kirschner, Paul, Open University of the Netherlands/Utrecht University, Netherlands
- Knight, Katherine, University of Wisconsin Madison, United States
- Knipfer, Kristin, Technische Universitaet Muenchen, Germany
- Kollar, Ingo, University of Munich, Germany
- Koops, Willem, Maastricht University, Germany
- Kopp, Birgitta, University of Munich, Germany
- Koschmann, Timothy, Southern Illinois University, United States
- Koumpis, Adamantios, ALTEC Software S.A., Greece
- Kraemer, Nicole, University Duisburg-Essen, Germany
- Krange, Ingeborg, University of Oslo, Norway
- Krumm, Andrew, University of Michigan, United States
- Kumar, Swapna, University of Florida, United States
- Kwon, Samuel, Concordia University Chicago, United States
- Kyza, Eleni, Cyprus University of Technology, Cyprus
- Laferriere, Therese, Laval University, Canada
- Laffey, James, University of Missouri, United States
- Lafifi, Yacine, Guelma University, Algeria
- Lakkala, Minna, University of Helsinki, Finland
- Lambropoulos, Niki, IntelliGenesis, United States
- Lan, Yu-Ju, National Taiwan Normal University, Taiwan
- Lavoué, Elise, University Lyon, France
- Law, Nancy, University of Hong Kong, China
- Lee, Victor, Utah State University, United States
- Lee, Wincy Wing Sze, University of Hong Kong, China
- Leeder, Chris, University of Michigan Ann Arbor, United States
- Levy, Dalit, Harbin Institute of Technology, China
- Lewis, Armanda, New York University, United States
- Li, Wenjuan, University of Illinois Chicago, United States
- Li, Ken W., Hong Kong Institute of Vocational Education, China
- Li, Xiao, Southwest China University/South China Normal University, China
- Liang, Rose, National Institute of Education, Singapore
- Liao, Chang-Yen, National Central University, Taiwan
- Lin, Hsien-Ta, National Taiwan University, Taiwan
- Lin, Feng, University of Hong Kong, China
- Lin, Chiu Pin, National Hsinchu University of Education, Taiwan
- Lingnau, Andreas, Catholic University Eichstätt-Ingolstadt, Germany
- Liu, Shiyu, University of Minnesota Minneapolis, United States
- Liu, Lei, University of Pennsylvania, United States
- Lonchamp, Jacques, University of Lorraine, France
- Looi, Chee-Kit, Nanyang Technological University, Singapore
- Lu, Jingyan, University of Hong Kong, China
- Luckin, Rosemary, The London Knowledge Lab, United Kingdom
- Ludvigsen, Sten, University of Oslo, Norway
- Lund, Kristine, University of Lyon, France
- Ma, Jasmine, New York University, United States
- Madeira, Cheryl Ann, University of Toronto, Canada
- Magnifico, Alecia Marie, University of Illinois Urbana-Champaign/University of Wisconsin Madison, United States
- Mahardale, Jay, Admiralty Primary School, Singapore
- Manca, Stefania, Institute for Educational Technology, Italy
- Markauskaite, Lina, University of Sydney, Australia
- Martin, Crystle, University of California Irvine, United States
- Martínez Monés, Alejandra, University of Valladolid, Spain
- Matuk, Camillia, University of California Berkeley, United States
- Matuku, Batanayi, University of Cape Town, South Africa

- Mavrikis, Manolis, London Knowledge Lab, Institute of Education, United Kingdom
- McLaren, Bruce, Carnegie Mellon University, United States/Saarland University, Germany
- Medina, Richard, University of Hawaii Manoa, United States
- Mercie, Emma, Durham University, United Kingdom
- Jacobson, Michael, University of Sydney, Australia
- Mochizuki, Toshio, Senshu University, Japan
- Molinari, Gaelle, Distance Learning University, Switzerland
- Mor, Yishay, Open University, United Kingdom
- Morch, Anders, University of Oslo, Norway
- Mu, Jin, University of Hong Kong, China
- Muukkonen, Hanni, University of Helsinki, Finland
- Najafi, Hedieh, Ontario Institute for Studies in Education, Canada
- Namdar, Bahadir, University of Georgia, United States
- Nicolaou, Christiana, University of Cyprus, Cyprus
- Nnoroozi, Omid, Wageningen University, Netherlands
- Notari, Michele, PHBern University of Teacher Education, Switzerland
- Nwaigwe, Aadaeze, American University of Nigeria, Nigeria
- O'Malley, Claire, University of Nottingham, United Kingdom
- Oshima, Jun, Shizuoka University, Japan
- Papanikolaou, Kyparisia, School of Pedagogical and Technological Education, Greece
- Paul, Brna, University of Leeds, United Kingdom
- Pea, Roy, Stanford University, United States
- Pemberton, Lyn, University of Brighton, United Kingdom
- Pfister, Hans-Rüdiger, Leuphana Universität Lüneburg, Germany
- Pierroux, Palmyre, University of Oslo, Norway
- Piety, Philip, American Institutes for Research, United States
- Pinkwart, Niels, Clausthal University of Technology, Germany
- Prieto, Luis P., University of Valladolid, Spain
- Prilla, Michael, Ruhr University of Bochum, Germany
- Prodan, Augustin, Iuliu Hatieganu University, Romania
- Quintana, Chris, University of Michigan Ann Arbor, United States
- Raes, Annelies, Ghent University, Belgium
- Rebedea, Traian, University Politehnica of Bucharest, Romania
- Reedy, Gabriel, King's College London, United Kingdom
- Reeve, Richard, Queen's University, Canada
- Reimann, Peter, University of Sydney, Australia
- Renninger, K. Ann, Swarthmore College, United States
- Richter, Christoph, University of Kiel, Germany
- Rick, Jochen, Saarland University, Germany
- Rogers, Jim, Utah State University, United States
- Romero, Margarida, University of Toulouse, France
- Roschelle, Jeremy, SRI International, United States
- Rose, Carolyn, Carnegie Mellon University, United States
- Rusman, Ellen, Open University of the Netherlands, Netherlands
- Russell, Donna, University of Missouri Kansas City, United States
- Ryan, Stephanie, University of Illinois Chicago, United States
- Salden, Ron, Madeira Interactive Technologies Institute, Portugal
- Saleh, Asmalina, Indiana University, United States
- Saravanos, Antonios, University of Oxford, United Kingdom \ United Nations, United States
- Savelyeva, Tamara, University of Hong Kong, China
- Baruch, The Hebrew University of Jerusalem, Israel
- Schwendimann, Beat, University of California Berkeley, United States
- Sethi, Ricky, University of California Los Angeles, United States
- Shen, Cindy, National Taiwan Normal University, Taiwan
- Siqin, Tuya, University of Hong Kong, China
- Slof, Bert, University of Groningen, Netherlands
- Smith, Carmen, University of Vermont, United States
- Smith, Garrett, University of Wisconsin Madison, United States
- So, Hyo-Jeong, National Institute of Education/ Nanyang Technological University, Singapore
- Songer, Nancy, University of Michigan, United States
- Sosnovsky, Sergey, German Research Center for Artificial Intelligence, Germany
- Spada, Hans, University of Freiburg, Germany
- Stahl, Gerry, Drexel University, United States
- Stegmann, Karsten, University of Munich/Ludwig-Maximilians University, Germany
- Stoyanov, Slavi, Open University of the Netherlands, Netherlands
- Su, Yen-Ning, National Cheng Kung University, Taiwan
- Subramanian, Shree, University of Wisconsin Madison, United States

- Svihla, Vanessa, University of New Mexico/University of Texas Austin, United States
- Sylvan, Elisabeth, Massachusetts Institute of Technology, United States
- Tabak, Iris, Ben Gurion University, Israel
- Takeuchi, Lori, The Joan Ganz Cooney Center at Sesame Workshop, United States
- Tchounikine, Pierre, University of Grenoble, France
- Teasley, Stephanie, University of Michigan, United States
- Teo, Hon Jie, Virginia Tech, United States
- Teplovs, Chris, Problemshift, Inc., Canada
- Thomas, Jakita, Spelman College, United States
- Thompson, Kate, University of Sydney, Australia
- Tran, Cathy, University of Oslo, Norway \ University of California Irvine, United States
- Trausan-Matu, Stefan, University Politehnica of Bucharest, Romania
- Tsiatsos, Thrasyvoulos, Aristotle University of Thessaloniki, Greece
- Tzagarakis, Manolis, Research Academic Computer Technology Institute, Greece
- Underwood, Jody, Intelligent Automation, Inc., United States
- Underwood, Joshua, University of London, United Kingdom
- van 't Hooft, Mark, Kent State University, United States
- van Aalst, Jan, University of Hong Kong, China
- Van Amelsvoort, Marije, Tilburg University, Netherlands
- van Leeuwen, Anouschka, Utrecht University, Netherlands
- Vanover, Charles, University of South Florida, United States
- Vatrapu, Ravi, Copenhagen Business School, Denmark
- Verdejo, Felisa, Universidad Nacional de Educación a Distancia, Spain
- Voulgari, Argiro, University of Patras, Greece
- Vouras, George, University of Piraeus, Greece
- Vovides, Yianna, Georgetown University, United States
- Vu Minh, Chieu, University of Michigan, United States
- Wang, Su-Chen, National Cheng-Kung University, Taiwan
- Wardrip, Peter, University of Pittsburgh, United States
- Wasson, Barbara, University of Bergen, Norway
- Wee, Juan Dee, National Institute of Education/ Nanyang Technological University, Singapore
- Wegerif, Rupert, University of Exeter, United Kingdom
- Weible, Jennifer L., Pennsylvania State University, United States
- Wessner, Martin, Fraunhofer IESE, Germany
- White, Tobin, University of California Davis, United States
- Wilkerson-Jerde, Michelle, Tufts University, United States
- Winter, Marcus, University of Brighton, United Kingdom
- Wise, Alyssa, Simon Fraser University, Canada
- Wong, Lung Hsiang, National Institute of Education, Singapore
- Wopereis, Iwan, Open University of the Netherlands, Netherlands
- Xhafa, Fatos, Technical University of Catalonia, Spain
- Xie, Ying, Idaho State University, United States
- Yacef, Kalina, University of Sydney, Australia
- Yeh, Ron Chuen, Meiho University, Taiwan
- Yoon, Susan, University of Pennsylvania, United States
- Yuen, Johnny K.L., University of Hong Kong, China
- Yun, Wen, National Institute of Education, Singapore
- Zahn, Carmen, Knowledge Media Research Centre, Germany
- Zhang, Jianwei, University at Albany SUNY, United States
- Zheng, Binbin, University of California Irvine, United States
- Zimmerman, Heather Toomey, Pennsylvania State University, United States
- Zourou, Katerina, University of Luxembourg, Luxembourg

**To See the World and a Grain of Sand:
Learning across Levels of Space, Time, and Scale.
Proceedings of the International CSCL Conference 2013**

Nikol Rummel, Ruhr-Universität, Bochum, Germany, nikol.rummel@rub.de

Manu Kapur, National Institute of Education, Singapore, manu.kapur@nie.edu.sg

Mitchell J. Nathan, University of Wisconsin, Madison, USA, mnathan@wisc.edu

Sadhana Puntambekar, University of Wisconsin, Madison, USA, puntambekar@education.wisc.edu

The 10th International Conference on Computer-Supported Collaborative Learning (CSCL) is to be held at the University of Wisconsin-Madison, USA, from June 15 through 19, 2013 (<http://www.isls.org/cscl2013/>).

The CSCL conference is a multidisciplinary, international meeting sponsored by the International Society for the Learning Sciences (ISLS). The conference is held biennially in the years alternating with the International Conference of the Learning Sciences (ICLS). So far the conference has been held in the USA, Europe, and Asia. This conference is an important venue for CSCL researchers to come together from around the world to meet, report recent research findings and discuss timely and important issues of interest to the community. It draws researchers from psychology (educational, social, developmental, cognitive, linguistic, cultural-historical), the social sciences (anthropology, sociology, communication studies, philosophy of language), and design disciplines (computer and information science, curriculum and didactics), as well as researchers from Artificial Intelligence (AI) and the cognitive sciences.

CSCL interactions in both online and face-to-face contexts occur at multiple levels of time, space, scales of analysis, and scales of group/population structure, to name a few. The title of our conference theme, inspired by (and modified from) William Blake's poem "Auguries of Innocence" reflects this unique aspect of CSCL in which interactions and learning need to be understood, supported and analyzed at multiple levels. We see an attention to the theoretical, methodological and technological issues of addressing research at multiple levels to be one that is highly responsive to current research among the CSCL community as well as developing emerging epistemological and methodological issues that will shape our intellectual efforts well into the future.

The relevance and timeliness of the conference theme is evidenced by workshops and presentations at previous ISLS conferences and by recent publications in relevant journals. For instance, the issue of *analyzing CSCL interactions at multiple levels and with various methodological approaches*, in order to further our understanding of the learning mechanism underlying CSCL, has received a lot of attention in the community over the last decade. At ICLS 2004 Nikol Rummel and Hans Spada organized a symposium entitled "Cracking the Nut – But Which Nutcracker to Use? Diversity in Approaches to Analyzing Collaborative Processes in Technology-Supported Settings." At ICLS 2008, Daniel D. Suthers, Nancy Law, Carolyn P. Rose, Nathan Dwyer held a workshop with the title "Developing a Common Conceptual and Representational Framework for CSCL Interaction Analysis", which was followed by a series of workshops at the recent CSCL and ICLS conferences and culminated in an edited book to appear in June 2013 at Springer: Suthers, D., Lund, K., Rose, C. P., Teplovs, C., & Law, N. (in press). *Productive Multivocality in the Analysis of Group Interactions*. Furthermore, in his introduction to the most recent issue of the International Journal of Computer-Supported Collaborative Learning (ijCSCL, Volume 8, Issue 1), which is dedicated to the topic of "Learning across levels", editor-in-chief Gerry Stahl cites and takes up the CSCL 2013 conference theme to argue that "time has come for CSCL to address the problem of traversing levels of analysis with exacting research" (p. 10).

At CSCL 2013, the conference theme is addressed from different perspectives through three keynote talks by Josep Call (Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany), Kori Inkpen Quinn (Microsoft Research, USA), and Justine Cassell (Carnegie Mellon University, Pittsburgh, USA). The conference theme will further be showcased by an invited plenary session on "Multiple methods in CSCL research", and by various pre-conference workshops and multiple contributions (paper and poster sessions, submitted and invited symposia, panels, and demonstrations of innovative educational technology) throughout the main conference.

We received many high-quality submissions for CSCL 2013. Submission categories included full papers (8 pages, presenting mature work), short papers (4 pages, summarizing work that is still in progress or of smaller scale) and posters (2 pages, sketching work in early stages or novel and promising ideas). Further submission categories were symposia (8 pages, conveying larger ideas or integrating findings around a specific issue), panels (3 pages, coordinating multiple perspectives on a specific, timely topic), demonstrations (3 pages, providing an opportunity to interactively present new tools and technologies for supporting and/or analyzing

collaborative learning), and pre-conference workshops and tutorials (5 pages, proposing collaborative knowledge-building sessions where participants actively work together on a focused issue). Submissions were also invited for a doctoral consortium and an early career workshop.

Each full paper, short paper, or poster proposal was reviewed blind by one or two peer reviewers and one program committee member. Program committee members then summarized the reviews and provided the program co-chairs with a brief assessment. Finally, the program chairs carefully considered the reviews and the meta-review, and in many cases read the submissions themselves before making the final decision. Proposals for symposia and panels were reviewed by two program committee members and by the program co-chairs. Demonstration, workshop and tutorial proposals were reviewed by members of the respective steering committee within the CSCL 2013 organization.

As in previous years, the acceptance rates for full and short papers were competitive: The acceptance rates for full papers and short papers were 36% and 39% respectively. For posters, the acceptance rate was more inclusive (78%) to allow for presentation of work that is in early stages and for productive discussions of novel and promising ideas.

The CSCL 2013 proceedings comprise two volumes: Volume 1 includes full papers and symposia. Volume 2 includes short papers, posters, and panels, demonstrations, as well as abstracts for all community events (keynotes, workshops and tutorials, early career and doctoral workshops, and invited panels and symposia).

Many fields within the physical and social sciences and the design sciences have long grappled with the notion of *supervenience* -- how phenomena at one scale of time or space can influence and be influenced by those at larger and smaller scales. New technologies and methodologies are making theoretical advancements possible, and leading the exciting and growing field of CSCL into frontiers of research and development that stand to contribute to improvements in education, the design of new means for collaborating, and new end-user experiences. Our world is becoming a more connected place because of the ways -- both large and small -- that we interact with technologies, and in so doing, come to interact with one another. As organizers of this conference and editors of this volume, we hope these interactions continue well beyond the bounds of this event or these proceedings, but continue to reshape ourselves and the world.

Contents: Volume 1

Full Papers

Peer Scaffold in Math Problem Solving <i>Rotem Abdu</i>	2
Intensification of Group Knowledge Exchange with Academically Productive Talk Agents <i>David Adamson, Colin Ashe, Hyeju Jang, David Yaron, Carolyn P. Rosé</i>	10
Variation in Other-Regulation and the Implications for Competence Negotiation <i>Karlynn R. Adams-Wiggins, Toni Kempler Rogat</i>	18
The Blogosphere as Representational Space <i>Richard Alterman, Bjorn Levi Gunnarsson</i>	26
Measuring 'Framing' Differences of Single-Mouse and Tangible Inputs on Patterns of Collaborative Learning <i>Luis Andrade, Joshua Danish, Yanín Moreno, Lenin Pérez</i>	34
Measuring Social Identity Development in Epistemic Games <i>Golnaz Arastoopour, David Williamson Shaffer</i>	42
Experiences as Resources for Sense Making: Health Education Students' I-positioning in an Online Science Philosophy Course <i>Maarit Arvaja</i>	49
Supporting School Group Visits to Fine Arts Museums in the 21st Century: A CSCL Concept for a Multi-Touch Table Based Video Tool <i>Moritz Borchers, Philipp Mock, Carmen Zahn, Jörg Edelmann, Friedrich W. Hesse</i>	57
Navigating through Controversial Online Discussions: The Influence of Visualized Ratings <i>Jürgen Buder, Christina Schwind, Anja Rudat, Daniel Bodemer</i>	65
Constructive Use of Authoritative Sources Among Collaborative Knowledge Builders in a Social Science Classroom <i>Fei-Ching Chen, Chih-Hsuan Chang, Cheng-Yu Yang</i>	73
Making Collective Progress Visible for Sustained Knowledge Building <i>Mei-Hwa Chen, Jianwei Zhang, Jiyeon Lee</i>	81

Designing Reference Points in Animated Classroom Stories to Support Teacher Learners' Online Discussions <i>Vu Minh Chieu, Patricio Herbst</i>	89
Identifying Gender Differences in CSCL Chat Conversations <i>Costin-Gabriel Chiru, Traian Rebedea, Stefan Trausan-Matu</i>	97
The Impact of CSCL Beyond the Online Environment <i>Sherice N. Clarke, Gaowei Chen, Catherine Stainton, Sandra Katz, James G. Greeno, Lauren B. Resnick, Gregory Dyke, Iris Howley, David Adamson, Carolyn Penstein Rosé</i>	105
When Face-to-Face Fails: Opportunities for Social Media to Foster Collaborative Learning <i>Tamara Clegg, Jason C. Yip, June Ahn, Elizabeth Bonsignore, Michael Gubbels, Becky Lewittes, Emily Rhodes</i>	113
Aggregating Students' Observations in Support of Community Knowledge and Discourse <i>Rebecca Cober, Colin McCann, Tom Moher, Jim Slotta</i>	121
Making Use of Collective Knowledge - A Cognitive Approach <i>Ulrike Cress</i>	129
The Benefits and Limitations of Distributing a Tangible Interface in a Classroom <i>Sébastien Cuendet, Pierre Dillenbourg</i>	137
Cohesion-based Analysis of CSCL Conversations: Holistic and Individual Perspectives <i>Mihai Dascalu, Stefan Trausan-Matu, Philippe Dessus</i>	145
Going Deep: Supporting Collaborative Exploration of Evolution in Natural History Museums <i>Pryce Davis, Michael Horn, Laurel Schrementi, Florian Block, Brenda Phillips, E. Margaret Evans, Judy Diamond, Chia Shen</i>	153
Real-Time Collaboration for Web-Based Labs <i>Luis de la Torre, Ruben Heradio, Sebastian Dormido, Carlos Jara</i>	161
Fostering Learning and Collaboration in a Scientific Community - Evidence from an Experiment Using RFID Devices to Measure Collaborative Processes <i>Julia Eberle, Karsten Stegmann, Kris Lund, Alain Barrat, Michael Sailer, Frank Fischer</i>	169
Supporting Active Wiki-based Collaboration <i>Adam Eck, Leen-Kiat Soh, Chad Brassil</i>	176

Inhibiting Undesirable Effects of Mutual Trust in Net-Based Collaborative Groups <i>Tanja Engelmann, Richard Kolodziej, Michail Kozlov</i>	184
Constructing and Deconstructing Materially-Anchored Conceptual Blends in an Augmented Reality Collaborative Learning Environment <i>Noel Enyedy, Joshua Danish, David DeLiema</i>	192
Understanding Collaborative Practices in the Scratch Online Community: Patterns of Participation Among Youth Designers <i>Deborah Fields, Michael Giang, Yasmin Kafai</i>	200
Incentives in Educational Games: A Multilevel Analysis of Their Impact on Elementary Students' Engagement and Learning <i>Michael Filsecker, Daniel Thomas Hickey</i>	208
The Joint Action Theory in Didactics: A Case Study in Videoconferencing at Primary School <i>Brigitte Gruson, Gérard Sensevy</i>	216
Inter-Personal Browsing: Supporting Cooperative Web Searching by Face-to-Face Sharing of Browser Pages <i>Tomoko Hashida, Koki Nomura, Makoto Iida, Takeshi Naemura</i>	224
Learner-Support Agents for Collaborative Interaction: A Study on Affect and Communication Channels <i>Yugo Hayashi</i>	232
An Adapted Group Psychotherapy Framework for Teaching and Learning About CSCL <i>Yotam Hod, Dani Ben-Zvi</i>	240
The Sequential Analysis, Modeling and Visualization of Collaborative Causal Mapping Processes and Effects on Causal Understanding <i>Allan Jeong, Woon Jee Lee</i>	248
When Instruction Supports Collaboration, But Does Not Lead to Learning - The Case of Classroom and Small Group Scripts in the CSCL Classroom <i>Ingo Kollar, Christof Wecker, Sybille Langer, Frank Fischer</i>	256
Interface Tangibility and Gesture in Mediating Individual Agency Within Group Spatial Problem Solving with an Ecosystem Simulation <i>Helen Kwah, Leilah Lyons, Dixie Ching, Adam Eck, Leen-Kiat Soh, Chad Brassil</i>	264

Teacher Framing, Classroom Collaboration Scripts, and Help-Seeking and Help-Giving Behaviors <i>Eleni Kyza, Yiannis Georgiou, Demetra Hadjichambi, Andreas Hadjichambis</i>	272
Using Gartner's Hype Cycle as a Basis to Analyze Research on the Educational Use of Ubiquitous Computing <i>Jari Laru, Sanna Järvelä</i>	280
Repurposing Everyday Technologies for Math and Science Inquiry <i>Sarah Lewis, Wendy Ju</i>	288
Delaying Instruction Alone Doesn't Work: Comparing and Contrasting Student Solutions is Necessary for Learning from Problem-Solving Prior to Instruction <i>Katharina Loibl, Nikol Rummel</i>	296
Exploring Evolutionary Concepts with Immersive Simulations <i>Michelle Lui, James D. Slotta</i>	304
Designing for Group Math Discourse <i>Rachel M. Magee, Christopher M. Mascaro, Gerry Stahl</i>	312
MTClassroom and MTDashboard: Supporting Analysis of Teacher Attention in an Orchestrated Multi-Tabletop Classroom <i>Roberto Martinez-Maldonado, Judy Kay, Kalina Yacef, Marie-Theresa Edbauer, Yannis Dimitriadis</i>	320
Juxtaposing Practice: Uptake as Modal Transposition <i>Richard Medina, Daniel Suthers</i>	328
Emotion Feedback During Computer-mediated Collaboration: Effects on Self-Reported Emotions and Perceived Interaction <i>Gaëlle Molinari, Guillaume Chanel, Mireille Bétrancourt, Thierry Pun, Christelle Bozelle</i>	336
Knowledge Organization with Multiple External Representations in an Argumentation Based Computer Supported Collaborative Learning Environment <i>Bahadir Namdar, Ji Shen</i>	344
Epistemic Trajectories: Mentoring in a Game Design Practicum <i>Padraig Nash, David Williamson Shaffer</i>	352
Gameplay as Assessment: Analyzing Event-Stream Player Data and Learning Using GBA (A Game-Based Assessment Model) <i>V. Elizabeth Owen, R. Benjamin Shapiro, Richard Halverson</i>	360

The Role of Identities in the Process of Knowledge Construction in CSCL Settings <i>Murat Oztok</i>	368
Effects of an Interculturally Enriched CSCL Script on Students' Attitudes and Performance <i>Vitaliy Popov, Harm J.A. Biemans, Martin Mulder</i>	375
Fostering CSCL Adoption: An Approach to Professional Development Focused on Orchestration <i>Luis P. Prieto, Sara Villagr�-Sobrino, Yannis Dimitriadis, Juan I. Asensio-P�rez, Iv�n M. Jorr�n-Abell�n</i>	383
The Effect of Formative Feedback on Vocabulary Use and Distribution of Vocabulary Knowledge in a Grade Two Knowledge Building Class <i>Monica Resendes, Bodong Chen, Alisa Acosta, Marlene Scardamalia</i>	391
Youth Roles and Leadership in an Online Creative Community <i>Ricarose Roque, Natalie Rusk, Amos Blanton</i>	399
Using Eye-Tracking Technology to Support Visual Coordination in Collaborative Problem-Solving Groups <i>Bertrand Schneider, Roy Pea</i>	406
'Co-Alienation' Mediated By Common Representations in Synchronous E-Discussions <i>Baruch B. Schwarz, Yifat Ben-David Kolikant, Maria Mishenkina</i>	414
Ethics for Design-Based Research on Online Social Networks <i>R. Benjamin Shapiro, Pilar N. Ossorio</i>	422
Understanding Collaborative Program Comprehension: Interlacing Gaze and Dialogues <i>Kshitij Sharma, Patrick Jermann, Marc-Antoine N�ssli, Pierre Dillenbourg</i>	430
Effects of Robots' Revoicing on Preparation for Future Learning <i>Hajime Shirouzu, Naomi Miyake</i>	438
Examining Dynamics of Implementing Flexible Group Discourse in a Principle-Based CSCL Environment <i>Tuya Siqin, Jan van Aalst, Samuel Kai Wah Chu</i>	446
Resources for Connecting Levels of Learning <i>Gerry Stahl, Diler �ner</i>	454

Learning with Collaborative Inquiry: A Science Learning Environment for Secondary School Students <i>Daner Sun, Chee-Kit Looi, Evelyn Teo</i>	462
Experts Learn More (than Newcomers): An Exploratory Study of Argumentation in an Online Help Forum <i>Hon Jie Teo, Aditya Johri</i>	470
Identification of Patterns of Tool Use and Sketching Practices in a Learning By Design Task <i>Kate Thompson, David Ashe, Dewa Wardak, Pippa Yeoman, Martin Parisio</i>	478
Using Automated and Fine-Grained Analysis of Pronoun Use as Indicators of Progress in an Online Collaborative Project <i>Kate Thompson, Shannon Kennedy-Clark, Nick Kelly, Penny Wheeler</i>	486
Phases of Design: Following Idea Development and Patterns of Collaborative Discussion in a Learning By Design Project <i>Kate Thompson, David Ashe, Pippa Yeoman, Martin Parisio</i>	494
Individualistic Appropriation as a Primary Mechanism of Collaborative Conceptual Change: A Case Study <i>Michael Tscholl, John Dowell</i>	502
Experiences of a Newbie Helper in a Free Open Online Mathematics Help Forum Community <i>Carla van de Sande</i>	510
Multidimensional Teacher Behavior in CSCL <i>Anouschka van Leeuwen, Jeroen Janssen, Gijsbert Erkens, Mieke Brekelmans</i>	518
Learning to Argue in Mathematics: Effects of Heuristic Worked Examples and CSCL Scripts on Transactive Argumentation <i>Freydis Vogel, Elisabeth Reichersdorfer, Ingo Kollar, Stefan Ufer, Kristina Reiss, Frank Fischer</i>	526
Relationships between Listening and Speaking in Online Discussions: An Empirical Investigation <i>Alyssa Friend Wise, Simone Nicole Hausknecht, Yuting Zhao</i>	534
Influence of Epistemological Beliefs and Goal Orientation on Learning Performance in CSCL <i>Kui Xie, Kun Huang</i>	542

Symposia

<p>Embedding Participatory Design into Designs for Learning: An Untapped Inter-disciplinary Resource? <i>Elizabeth Bonsignore, June Ahn, Tamara Clegg, Mona Leigh Guha, Juan Pablo Hourcade, Jason C. Yip, Allison Druin</i></p>	549
<p>Mass Collaboration - An Emerging Field for CSCL Research <i>Ulrike Cress</i></p>	557
<p>Scripting and Orchestration: Recent Theoretical Advances <i>Frank Fischer, Jim Slotta, Pierre Dillenbourg, Pierre Tchounikine, Ingo Kollar, Christof Wecker, Karsten Stegmann, Clark Chinn</i></p>	564
<p>Are CSCL and Learning Sciences Research Relevant to Large-Scale Educational Reform? <i>Nancy Law, Naomi Miyake, Chee-Kit Looi, Riina Vuorikari, Yves Punie, Marcia Linn</i></p>	572
<p>Designing to Improve Biology Understanding Complex Systems in High School Classrooms: No Simple Matter! <i>Susan Yoon, Eric Klopfer, Josh Sheldon, Ilana Schoenfeld, Daniel Wendel, Joyce Wang, Hal Scheintaub, David Reider</i></p>	580

Volume 1

Full Papers

Peer scaffold in math problem solving

Rotem Abdu; the Hebrew University of Jerusalem, Mount Scopus, Israel; Rotem_abdu@yahoo.com

Abstract: One of the most important issues that are dealt with in CSCL environments is self, and collaborative, regulated learning; independent of the support of teachers. In the first part of this paper, I will bring forward an innovative pedagogical approach for collaborative learning of math problem solving, accompanied by appropriate software (1) *Metafora*'s planning tool: a visual based planning and reflecting space for socio-meta-cognitive elicitation of collaborative learning processes, and, (2) *Geogebra*: a math application for the creation of dynamic Geometric figures in Cartesian domain. In the second part of this paper I will illustrate a learning scenario within the context of a collaborative math problem solving scenario. Then, I will highlight a behavior of collaborative learning, in which one team member, S1, makes progress with solving the problem, and goes back to help his peer, S2. S1 scaffolds his peer's work by (1) Reporting what he, S1, did on the shared planning-reflecting space (2) Monitoring his peer's error (3) Explicating this error to his peer (4) scaffolding his peer's construction of a Geogebra model, without giving him the whole answer. This observation serves as an important progress in the attempts of modern educators, and education design-researchers, to share some of the responsibility of the learning processes with students.

Supporting collaborative *Planning & Reflecting* in Math Problem Solving

Collaboration is considered as a central means for individual progress in modern society (Perret-Clermont, 2011, Wheelan, 1999). This serve as a good reason to progress *collaborative learning*, claiming that this mixture of individual and interactive activities can trigger learning mechanisms (Dillenbourg, 1999). In his attempt to illustrate a theoretical framework for such learning, Wegerif, (2006; 2011) explains that successful collaborative problem solving depends on the extent to which the solvers talk together and open up a reflective *Shared-Space*, which allows the emergence of ideas. When opening such a shared space, one's monitoring of his own cognition affects his peers' monitoring of their cognitions (Efklides, 2006).

But what would happen if the collaborative monitoring and regulation of learning is an *explicit* process that takes place as part of the collaborative solution? In the case of math problem solving this is a rather critical question. Heavy research strand (e.g.; Veenman and Spaans, 2005) shows the importance of metacognitive behaviors in the process of math problem solving, bringing forward the importance of being aware of the learning processes *before* (Weinberger, 2011; Rummel and Spada, 2005), *while* (Abdu and Schwarz, 2012; Schoenfeld, 1985) and/or *after* (Hamilton, Lesh, Lester & Yoon, 2007) performing it. My research team and I hypothesize that a tool that will afford the creation of a dialogic shared space between students can support emergence peer monitoring and regulation of collaborative learning process.

Metafora

The Metafora system is a software platform that encompasses a suite of tools used to support and encourage the development of "Learning to learn together" (L2L2) skills, through domain-specific activities in science and math. It is currently funded by the European Union. The idea is to have students collaboratively work in periods of 2-3 weeks, in order to give solution to a given challenge. During this solution process the students will plan, solve and reflect upon the learning and solving process in order L2L2. Their teacher and the software will scaffold their learning process. Students mutually engage in achieving the solution to a challenge through developing communication, strategic thinking and problem solving skills.

The System

The main tools that were developed and integrated into Metafora, for this end, include a virtual space for mapping of argumentative discussions (LASAD), a set of microworlds for simulating phenomena in science and math, and a planning tool. These tools are all interconnected, and monitored by an Analysis Component – an artificial intelligence component which is planned to take some moderating-load from the teacher. I will now elaborate about a couple of tools that were used in the current study: The planning/reflecting tool and Geogebra.

The *planning tool*: a shared space with which groups of students collaboratively, and autonomously (thus, in different computers), construct plans and reflections upon their work. This is being done by a creation of a constantly revised map; with the use of a set of icons we call "Visual Language Cards": a closed set of graphical ontology (See figure 1). This ontology is based on models of inquiry-based learning (e.g. Tamir, 2006), and of problem-solving (e.g. Polya, 1945). The ontology organizes the collaborative problem solving: Finding hypotheses, simulation, discussion, etc. The visual language also represents scientific/mathematical moves: understanding the problem, reflect, simulate, etc.



Figure 1. Examples of two visual language elements: Stages of the problem solving and processes undertaken during these stages.

The second tool that was used in this study is *Geogebra*: A math application that affords the creation of dynamic Geometric figures in a cartesian domain. Geogebra serves as a rich ground for learning math (e.g. Gergelitsove and Holan, 2012; Stahl, 2009), and we therefore decided to facilitate it within the overall Metafora system.

The Pedagogical Approach

The main idea of our approach is that bringing students to perform explicit discourse upon the meaning of actions/cards, within the context of the solution process of a specific challenge, elicits the learning of communication, metacognitive and problem solving skills (Rogoff, 1990). We bring students to engage with computer supported collaborative problem solving in mathematics, while explicating their learning processes with the help of the planning tool.

At first we had a rather rigid idea of a tool that affords students with planning their work ahead, and to some extent commit to that plan. After some observations upon students' work with the planning tool we realized two things: First, students were often reluctant to plan in advance, before they make sense of the problem, or, "Understand the problem". Second, we found out (Abdu and Schwarz, 2012) that teachers that use the planning tool, first needed to make sense of the problem and then they were able to reflect upon their solution to that point. When they reached the "current" point they were able to plan ahead their collaborative solution. These two complementary findings led us to the understanding that asking students to plan ahead their collaborative solution should be done mainly after they made sense of the challenge. Therefore, the planning tool becomes also a planning tool, in which the solvers create a model of their own learning process (Hamilton, Lester, Lesh, & Yoon, 2006).

Other observations made by the Metafora pedagogical team led us to identify four key skills that are necessary for any process in which students are learning together, and on a higher level, L2L2. These skills are: *Distributed leadership*, *Mutual engagement*, *Peer assessment*, and *Group reflection on the learning process*. In the conclusions I will show the emergence of these skills. In particular, I will show a behavior that was identified as *Peer Scaffold* by our research team, as a result of the existence of these skills.

The City challenge

The class, the course

Sixteen 8th grade male math-competent students from a religious school, all from mid-high class families in Jerusalem, participated in this study. The students met once a week in a computer class and participated in 8

month course of computer-supported collaborative math problem solving. The teacher in charge was an experienced math teacher, and teachers' tutor, that gave the course as a part of her master's thesis program, at the Hebrew University of Jerusalem.

We created a course with fourteen units that appear as a succession of activities. We adopted the approach of a design research (Cobb, 2001; Collins, Joseph & Bielaczyc, 2004) in which the learning environment is assessed and refined throughout a course.

The course had three phases. **Collaborative learning establishment:** Two double lessons in which the students took part in paper and pencil collaborative problem solving, involving teacher's orchestration. The purpose of this stage was to practice the students with group work and attracting them into the course. **Learning heuristics, strategies, and how to use the computerized tools:** Eight double lessons, in which groups of students solved 1-3 problems in every lesson. Each problem focused on a specific heuristic or strategy that was learned in the context of one or more problems. Students also learned to use different computerized tools: The Planning tool, Geogebra, and other micro-worlds. The teacher focused on the following heuristics and strategies: Planning, Reflecting, "Thinking outside the box", abduction (backward strategies), introducing proper notations, Creating a model, Allocating tasks, Generalizing, Checking a simpler case, Hypothesizing, Checking hypothesis, Trial and error and looking for patterns. **Solving challenges:** Fifteen weeks in which, groups of students were given challenges with longer time frames, in order to collaboratively solve relatively complex problems.

To foster the acculturation to problem solving and collaboration we adapted well known activities to problem solving challenges that have multiple solutions and/or multiple paths to a solution. Part of the content of the problems is directly related to school curriculum. Some other challenges were open-ended, thus affording the elaboration and the application of strategies to solve the challenge (Wee & Looi, 2009). One of them will be brought here: the "City" challenge.

The course was subdivided to fourteen Learning units with different durations: from 45 (one lesson) to 180 minutes (3 double lessons). However, the scenario of a challenge was quite stable: At the beginning the teacher presented a challenge to the class. The students then initiated their work through explorations. In some key milestones of their solution process we asked them to use the planning tool. And at least the couple that we filmed complied and even seldom used it when not asked. Three kinds of reflections were implemented: (1) *within group reflection* that was done throughout the process, when needed (2) a whole class reflection, with the teacher as a leader that gave an overview of different solutions and solution paths. (3) a reflection that was done by a group of students, upon their collaborative problem solving and solution processes, in front of the whole class.

Design Principles of the challenge

The *City challenge*, posed and supported by the teacher, is a relatively complex problem in math that was designed to be solved over three double lessons. In this challenge, students need to find a point that is equidistant from 7 general points in a 2D space. Math fan readers are invited to take a break and come up with the answer for this geometrical place. To the ones that are not big fans of math, I will tell this- there isn't such a point except for one particular case: When the 7 points are located on one circle. However, the ideas that can develop from such inquiry are vast, if done properly. In order to solve such a challenge, students need to find a simpler case of the problem (See table 1) and gradually find the solution to more complex cases. Through such a solution of a challenge, the students need to apply concepts such as "median", "medians' intersecting point" and "a perpendicular bisector". Most of these concepts were a part of the "ordinary" curriculum for these students and some –such as the perpendicular bisector- were rather new. Geogebra serves here as a facilitator for the construction of these shapes and by that- supports the meaning making of these concepts (Stahl, 2009).

A challenge that spans over three double lessons and solved by groups of competent students needs to be hard enough for them, so they will not be able to answer it immediately, but it should be within grasp and attractive enough from their point of view. Thus, they will not give up easily. We carefully planned our scaffold for the solution, while letting the students the freedom to explore directions to the solution and construct their own Geogebra models and planning-reflecting maps. One of the ways to achieve these goals is to plant a *Cognitive conflict*. At the end of step 2 (See table 1), students reach a conclusion about the equilateral triangle: The equidistant point is in the meeting point of the angles bisectors'/medians meeting point/heights. Obviously,

in the case of an equilateral triangle it does not matter, since the three loci are at the same point. In preliminary observations it seems that students come to this conclusion rather fast. However, when they tried this solution in the case of a general triangle (step three, see table 1) they found out that this does not work.

Solving the City challenge

In the city challenge students are asked to place an energy system in an imaginary city, in the center of seven other institutions. Since the challenge was too complicated to be solved in one step, the teacher guided her students to solve it in four steps as can be seen in table 1 below.

Table 1: the four steps of the City challenge:

<p>Step One: Place the energy center in the center of seven institutions, all of which are important to the city and all are packed with people. What is the conclusion of the conference committee?</p> <p>Step two: Simpler case- 3 institutions located in the structure of an equilateral triangle. Where to place the point?</p> <p>Step Three: Three institutions located in the structure of any triangle. Where to place the point? Try to formulate a final conclusion. Is there an equally distant point to the three vertices of a triangle? If so where it is located?</p> <p>Step four: Give a general answer: Where should we put the energy center?</p>
--

We follow two students S1 and S2 while they solve step three - locating the energy center in the middle of a general triangle. This is the second lesson out of three that were dedicated for this challenge. The two came to a solid conclusion in the previous lesson, that the point that is equidistant from the three vertices of an equilateral triangle is the medians' intersection. When they moved to step three, they tried a couple of conjectures. First, they checked the medians' intersection and found out that it is not the desired point. Second, S1 found the solution for the case of a right angled triangle: The equidistant point is the midpoint of its hypotenuse (Thus segments x , y and z are equal in figure 2). However, this solution was only constructed by S1, while S2 is only watching him.

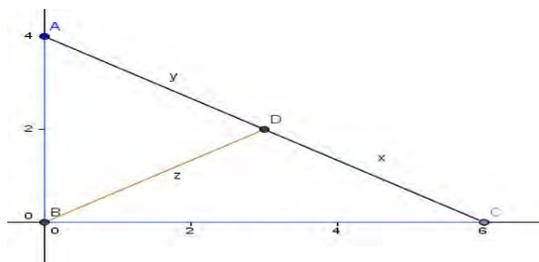


Figure 2: The point that is equidistant from the three vertices of a right angled triangle is the midpoint of the hypotenuse (point D)

In the beginning of the second lesson, the students were asked to reflect upon their work, and plan their solution to the challenge with the planning tool, before they continue to solve the challenge.

Peer scaffold: construction of a Geogebra model for the case of right angled triangle

Between various instances in which we observed peer-scaffold behaviors, the most rigorous episode is the current, in which we show that although S1 and S2 established that their next step will be- checking the intersection of medians; when we looked at their next actions we observed that while S1 regarded it as a reflective move, S2 constructed an intersection of the medians, with Geogebra. This is despite the fact that both constructed this model in the previous lesson, and found out that this point is not equidistant to all vertices. Later, S1 progresses to reflecting upon his solution for the right-angled triangle, and only when he finishes, he scaffolds S2's construction of building such a solution with Geogebra.

The Episode

It starts as S2 inspects the case of the medians' intersection, with the help of Geogebra; while S1 avoided S2's work and reports on the success he had in the previous lesson: Discovering that the equidistant point in the case of a right angled triangle is the midpoint of its hypotenuse. We join them as S2 measures the lengths from the vertices to the medians meeting points in his Geogebra model and S1 looks at the planning tool.

1. S1: *"Let's reflect again on the process...OK, S2, now I am going to write the story of a lifetime"*. S1 writes in the "reflect on process" card: *"With pure genius of his honorable S1 ..."*
2. S2 [Refers to his measurements]: *"And as we thought: a mistake"*
3. S1 [continues to write in the card]: *"we decided that we will (We=I will) solve for a right angled triangle. I tried to check the median to the hypotenuse, and I made it"*
4. S2: *"S1, there are medians, bisectors what else are there? Amm...perpendicular... how do you create a perpendicular? ...Amm, I don't think it's going to work..."*
5. S1 [Keeps writing]"... (After whole 80 minutes!)"

After he is done creating a median's intersection with Geogebra, S2 is ready to work on the case of a right angled triangle that was reported by his friend in the planning-reflecting map. S1 is in a different place: Since he constructed the triangle on the first lesson, he just reported in Planning-reflecting tool that "we decided that we will (We=I will) solve for a right angled triangle. I tried to check the median to the hypotenuse, and I made it". S1 knows the solution for a right angled triangle, and he made it clear so everybody (The teacher, the video camera) will know. Now, S1 is willing to scaffold S2's process of building a right angled triangle, with GeoGebra. First, S2 makes an attempt to create a right angled triangle, but the right angle is not accurate, since S2 did not define it properly (See figure 3).

6. S2: *"So, in the case of right angled triangle, what is the answer, remind me?"*
 7. S1: *"You insult me, didn't you hear my answer? [Points to the map] You can read it here. "*
 8. S2: *"I don't feel like reading"*
 9. S1: *"Read! You do it all the time, you can do it now..."*
 10. S2: *"Ah, it is the median, no?"*
 11. S1 [looks at the camera]: *"Shush! Don't tell everyone"*
 12. S2: *"Which one?"*
 13. S1: *"Of the hypotenuse"*
 14. S2: *"Of the hypotenuse..."*
- Now S2 creates a hypotenuse to the right angled triangle he constructed while S1 looks at his work (Figure 3)

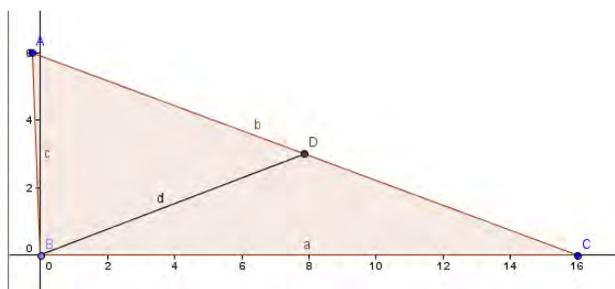


Figure 3: S2's building of a median to the "hypotenuse" of his erroneous right triangle model

15. S2: *"Does it work?"*
16. S1 [Leans towards S2's computer]: *"You see?! But the problem is...how did you build this? S2, I am sorry that I need to break the news, but this is not a right angled triangle."*
17. S2 [points with the mouse cursor on the shape]: *"Look"*
18. S1 [Points with his finger to the values of the segments, in the variables section of GeoGebra]: *"Look here... It is not right angled....."* [Goes back to his computer and grabs the mouse]
19. S2: *"Why...so how am I creating a right angled [triangle]?"*
20. S1: *"You build a right angled triangle."*
21. S2: *"And how can I know it is a right angled triangle?"*
22. S1: *"You build it on a tangent and then build [segments] to sectors"*
23. S2: *"But this is what I am trying to say"*
24. S1: *"And build another point"*
25. S2: *"But it looks like a right angled"*

26. S1: "...it looks like it, but it is not...Do you want me to show you how to do it? Happily...move!"[Grabs the computer mouse] "This is a much easier method."
27. S2: "Erase everything"
28. S1 erases the Geogebra screen.
29. S1: " OK, now we take random points, pay attention [puts a point in (0,0)]"
30. S2: "No, don't put it there ..."
31. S1: "Zero, Zero"
32. S2: "mmm..."
33. S1 [Puts a point in (0,4), puts a point in (6,0)]: "Zero four"
34. S2: "Is this the only way?"
35. S1 [Gives S2 the computer mouse]: "No, but it worked, because I simply built it on the vertices of the...X and Y"
36. S2 Receives the computer mouse from S1 and creates segments AB, BC and AC.
37. S1: "Now you got it!"
38. S2 [Builds a median to section BC]: "OK, now...you looked at the line to the hypotenuse ...And now lines" [Looks at the values of the segments: $x=y=z$ in figure 2]
39. S1: "See?!"
40. S2: "Now what?"
41. S1 [Points to the screen]: *Now you can check and see that AD BD and DC are equal.*
42. S2: OK

Now S2 sees that the midpoint of the hypotenuse in a right angled triangle is equidistant from the three vertices. Later, he will also explain this point to his teacher.

Discussion

The pedagogical setting, instructions and the affordance of the software led S1 and S2 to an interesting situation. In the beginning of the episode, the two worked in parallel- while S2 revisited the work that was done, S1 reports about this work in the planning tool, this presumably gives heads up to S2. In addition, the rather immodest choice of words "His honorable S1" (Lines 1, 3 and 5) implies that S1 wanted to report on the stages taken to that point, in order to talk about his achievement. But he had S2 as his peer, and S2 wanted to make sure he understands what happened in the last session, before they move forward. A possible explanation for that is that they wanted to be aligned with each other since the challenge was given to groups, rather than individuals. S2 asks for help from his peer. This ignites a process in which S1 scaffolds S2's progress.

Although I mentioned the term "scaffold", earlier in this paper, I was yet to define it. I will now do so, and the definition will stay fresh in the reader's mind in the next paragraph. Scaffolding is a kind of mediation in teaching, first termed by J.S. Bruner (Puntambekar & Hubscher, 2005). The main idea is that through *Scaffolding* a caregiver gradually transfers the responsibility of learning through a learner's ZPD. He does not give him the answer to the problem nor lets him figure it out all by himself. We identify four main behaviors that are associated with scaffolding: 1. Modeling- the caregiver shows the child how he performs an assignment. 2. Ongoing diagnosis- The caregiver monitors and regulates the child's actions. 3. Calibrated support- Achieved by recruiting the child's interest, reducing the degrees of freedom by simplifying the task, maintaining direction, highlighting the critical task features, controlling frustration, and demonstrating ideal solution paths, and, 4. Fading out- gradual transfer of responsibilities to the learning- from the caregiver to the child (Puntambekar & Hubscher, 2005).

Modeling the solution S1 gave to the case of an equilateral triangle was done by him in the previous lesson, and now he refers to S2 with the planning and reflecting tool as their tool for communication (line 9). It is safe to assume that S1 does not want to give S2 the solution "right away", since he wants S2 to read the "conclusions" that he wrote into the "reflect on process" card (lines 7 and 9). When S2 does not want to use this line of communication and takes a guess (line 10), S1 completes the insight, orally, (line 13). This, leads S1 to choose a different path, as he starts **monitoring** S2's work (Line 16) and commenting on it: "...how did you build this?" providing him with feedback "S2, I am sorry that I need to break the news, but this is not a right angled triangle." Moreover, S1 refers S2 to evidences they could both see - the values of the segments, in the variables section of GeoGebra- (line 18) "Look here... It is not right angled..." When S2 asks him " ... how am I

creating a right angled [triangle]?" S1 applies **calibrated support** as he reduces the degrees of freedom with the creation of a scaffold with Geogebra, while verbalizing his actions (lines 29 to 33). He puts three vertices of a right angled triangle are in points (0, 0), (6,0) and (4,0), but does not complete the model. Now the three vertices lie on the axes X and Y (see figure 2) and S2 can build the three segments (line 36). Then, based on a prompt by S1, S2 accurately checks if the point that is equidistant from the three vertices of a right angled triangle is the midpoint of the hypotenuse. S1 encourages his friend when he sees that S2 created a right angled triangle based on his scaffold, by saying "now you got it" when S2 creates the right angled triangle (line 37).

So, what are the characteristics of this particular interaction between S1 and S2 that allowed such an interaction to take place? I claim that lots of it has to do with the emergence of L2L2 behaviors, that progress this collaborative learning scenario: (1) We observed a *leadership move*, when S1 monitors S2's faulty construction of right angled triangle (line 16) and goes to sit next to his computer (Line 18), and then S2 asks for his help (line 19): "...so how am I creating a right angled [triangle]?" (2) The two are *Mutually engaged* although they are still quite far from the solution, (3) We can see *peer assessment* when, for example, S1 monitors S2's faulty construction. However, in this particular episode we observed that (4) *Group reflection* was shown only from S1's side, when he implicitly complains about his partner, as he writes in the "reflect on process" card: "we decided that we will (*We=I will*) solve for a right angled triangle, (line 3).

Conclusions

In this study I bring an example for collaborative learning between two students, over a challenging math problem. The learning environment and tools afford their collaborative problem solving, dynamic simulations of their mental models and discuss their collaborative work on a shared discussion space. This is being done with the aid of L2L2 skills. Through peer assessment that was done by one student, after his reflection upon his own solution, an error was discovered. This led him to scaffold his peer's creation of a model of a right angled triangle with a median to it hypotenuse, in Geogebra. He scaffolds his peer's work by (1) Reporting what he did in the previous lesson (2) Monitoring his peer's error (3) Explicating this error to his peer (4) scaffolding his peer's construction of his Geogebra model. In this spirit, as S1 supports S2's work, he even gives S2 some moral support by stating his success. However, this is also a socio-metacognitive move that is being done by S1.

The idea of having peers learning together and supporting each other's learning is very appealing for educators: Students that do not give to each other only the bottom line, but guide their peers throughout the learning. We can see how such environment gives some responsibilities of the teacher in the hands of one of the students. For this end, I offer two operational outcomes that can be derived from this instance. The first stems from the assumption that the act of reflection prompted the two towards a process in which they needed to level. This leads me to look up for more peer-regulating learning phenomena that emerge as a result of the use of the planning tool. The second comes as support to the importance of developing appropriate affordances for reflective moves, upon learning scenarios (Suthers, 2003).

There are several limitations, though, to this paper. It presents a study in progress, but it is far from sufficient. The described scenario is an example taken from over 25 hours of learning and instruction in an environment that elicits many interesting learning scenarios, such as peer scaffold. But as this learning episode stands on its own, questions of validity and reliability should come up. The reliability question was addressed in two meetings that were taken in which this episode was presented to my co-researchers in Metafora project, which maintained the same opinion as mine. Addressing the validity question is more complex, and I will have to identify more instances that will help me to define this kind of phenomena in a wide perspective, and eventually help in implementing it as a part of my coding system.

Endnotes

Metafora project is co-funded by the European Union under the Information and Communication Technologies (ICT) theme of the 7th Framework Programme for R&D (FP7), Contract No. 257872. We thank our colleagues in the project for the fruitful discussions and cooperation to support L2L2.

References

Abdu, R. & Schwarz, B.B.,(2012). "Metafora" and the Fostering of Collaborative Mathematical Problem Solving. *Constructionism 2012: Theory, Practice and Impact*. Athens, Greece 21-25 of August, 2012.

- Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of Learning Sciences*, 2(2), 141-178.
- Cobb, P. Stephan, M., McClain, K., & Gravemeijer, K., (2001). Participating in Classroom Mathematical Practices. *Journal of the Learning Sciences*, 10(1&2), 113-163.
- Collins, A., Joseph, D., & Bielaczyc, K., (2004). Design Research: Theoretical and Methodological Issues. *The Journal of the Learning Sciences*, 13(1), 15-42.
- Gergelitsová, G., Holan, T. (2012). An Automatic Evaluation of Construction Geometry Assignments. *Proceedings of the EC-TEL 2012 conference: 21st century learning for 21st skills*, Saabrucken, Germany, 18-21 of September, 2012.
- Dillenbourg P., (1999). What do you mean by collaborative learning? In Dillenbourg, P., (Ed) *Collaborative-learning: Cognitive and Computational Approaches* (pp. 1-19). Oxford: Elsevier.
- Dillon, J & Osborne, J.F., (2007) Science Education in Europe: Report to the Nuffield Foundation London:King's College London European Commission. *Europe needs More Scientists: Report by the High Level Group on Increasing Human Resources for Science and Technology*. Brussels. European Commission.
- Efklides, E., (2006). Metacognition and affect- What can metacognitive experiences tell us about the learning process? *Educational Research Review* 1(1), 3-14.
- Hamilton, E., Lesh, R., Lester, F., & Yoon, C., (2007). "The Use of Reflection Tools to Build Personal Models of Problem-Solving". In Lesh, R., Hamilton, E., and Kaput, J. (eds.), *Foundations for the Future in Mathematics Education* (pp. 347-365). Lawrence Erlbaum Associates, Mahwah, New Jersey.
- Perret-Clermont, A.N, (2011). Social relationships and thinking spaces for growth. *Proceedings of the symposium "Human and Societal Development: The role of Social Relationships"*. Cyprus, May 9th 2011.
- Polya, G., (1945). How to Solve It: A New Aspect of Mathematical Method, *1d ed. Princeton, NJ: Princeton*.
- Puntambekar, S., & Hübscher, R. (2005). Tools for Scaffolding Students in a Complex Learning Environment: What Have We Gained and What Have We Missed? *Educational Psychologist*, 40(1), 1-12.
- Rummel, N., & Spada, H., (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences*, 14(2), 201-241.
- Rogoff, B., (1990) Apprenticeship in Thinking: Cognitive Development in Social Context. *Oxford press*.
- Schoenfeld, A.H., (1985). Mathematical problem solving. *New York: Academic Press*.
- Stahl, G., (2009). *Studying virtual mathematics teams*. New York, NY: Springer.
- Suthers, D. D., (2003). Representational guidance for collaborative inquiry. In J. E. B. Andriessen, M. Baker, & D. D. Suthers, (eds.), *Arguing to learn. Confronting cognitions in computer-supported collaborative learning environments*, (pp. 27-46). Dordrecht: Kluwer.
- Tamir, D., (2006). Inquiry, in science teaching, and its reflection on biology teaching in Israel. *learning by inquiry: An ongoing challenge" edited by Anat Zohar, A.,* (pp. 16-56). Magnes press.
- Wee, J. D., & Looi, C. K., (2009). A Model For Analyzing mathematics Knowledge Building in VMT. In Stahl, G. (Ed.), *Studying Virtual mathematics Teams* (pp. 471-794). Springer-Verlag.
- Veenman, M.V.J. & Spaans, M.A., (2005). Relation between intellectual and metacognitive skills: age and task difference. *Learning and Individual Differences* 15, 159-176.
- Wegerif, R., (2006). Towards a dialogic understanding of the relationship between teaching thinking and CSCL. *International Journal of Computer Supported Collaborative Learning*, 1(1), 143-157.
- Wegerif, R., (2011). Towards a dialogic theory of how children learn to think. *Thinking Skills and Creativity* 6(3), 179-190.
- Weinberger, A., (2011). Principles of transactive Computer- Supported Collaboration Scripts. *Nordic Journal of Digital Literacy*, 6, 189-202.
- Wheelan, S. A., (1999). Creating effective teams: A guide for members and leaders. *Thousand Oaks, CA: Sage*.

Acknowledgments

I would like to thank Reuma DeGroot and Baruch Schwarz for their assistance and mentoring with the process of bringing this paper to life. And thanks to my dear friends, Leor Holzer and Ron Pallivathikal, for collaborating with me on the reflection upon this paper.

Intensification of Group Knowledge Exchange with Academically Productive Talk Agents

David Adamson, Colin Ashe, Hyeju Jang, David Yaron, Carolyn P. Rosé
Carnegie Mellon University, Pittsburgh PA
Email: {dadamson, cashe, hyejuj, yaron, cprose}@cmu.edu

Abstract: In recent years, intelligent conversational agents have been used with some level of effectiveness as dynamic support for collaborative learning in online chat. The classroom discourse community offers insights from analysis of effective classroom discussion facilitation practices that might productively inspire the design of such facilitator agents. In this paper, we evaluate one such conversational agent-as-facilitator design, drawn from the literature on what has been termed Academically Productive Talk. Specifically we evaluate the effect of a facilitation strategy referred to as Agree/Disagree, where students are prompted to evaluate the assertions of a partner student. In a simple two condition study, we evaluate the effect of this facilitation strategy in comparison with an otherwise identical condition where this facilitation strategy is absent. The results demonstrate a marginal positive effect on learning (effect size .55 standard deviations) and a significant intensification effect on the collaborative discourse.

Introduction

The literature on scripted support for Computer Supported Collaborative Learning describes scripts as a set of scaffolds and interventions that structure and facilitate student interaction, at both the macro-level of the collaborative activity and at the micro-level of individual actions (Dillenbourg, 2008). In particular, an instructor's role is to orchestrate multiple scripts (Fischer & Dillenbourg, 2006) to provide comprehensive, suitable support for the students throughout the collaborative learning experience. Recently, work building on this body of research has explored the role of dynamically scripted support for CSCL in the form of conversational agents, which have been shown to be successful in promoting student learning and conversation in collaborative discussion environments (Kumar et al., 2007; Chaudhuri et al., 2009; Dyke et al., 2012).

Additionally, analyses of expert teacher talk (Chapin et al., 2003) have revealed a set of discursive instructional practices, suitable for facilitating collaborative knowledge-building. The Academically Productive Talk framework (Michaels et al., 2007) describes a collection of discussion-facilitating moves a teacher can employ to promote rich student-centered conversation and collaboration. This framework can serve as an operationalization of effective group facilitation techniques which, combined with results and experiences from the CSCL scripting and conversational agents communities, lays the groundwork for automatic agent-based facilitation of small group online chat. Recent studies have made important advances in this area, and have identified limitations in agent design and behavior that must still be overcome. The contribution of this paper is to describe a successful new conversational agent behavior based on the principles of Academically Productive Talk, whose use leads to demonstrable gains in conceptually-rich student conversation and shows promising results for student learning.

In the remainder of the paper we first briefly review the literature on Academically Productive Talk and how it motivates design of intelligent conversational agent based support for collaborative learning as a form of dynamic microscripting. Next we describe our experimental design and methodology for process analysis. Then we describe our results and offer some interpretation. We conclude with a discussion of some limitations of this work and our current research directions.

Theoretical Background

The work presented here builds upon prior work from two disciplines: the discursive instructional framework of *Academically Productive Talk*, and the extensive body of CSCL research on supporting collaboration through scripting and conversational agents.

Academically Productive Talk

Academically Productive Talk has grown out of frameworks that emphasize the importance of social interaction in the development of mental processes. Michaels, O'Connor and Resnick (Michaels et al., 2007) describe a number of core moves that discussion facilitators can employ to foster effective student-centered classroom discussion. A selection of these moves are presented in Table 1.

Table 1. Selected Accountable Talk Moves

Academically Productive Talk Move	Example
Revoicing a student's statement	"So, let me see if I've got your thinking right. You're saying XXX?" (with time for students to accept or reject the teacher's formulation)
Asking students to <i>restate</i> someone else's reasoning	"Can you repeat what she just said, in your own words?"
Asking students to apply their own reasoning to someone else's reasoning	"Do you <i>agree or disagree</i> , and why?"

The teacher's facilitation plays a key role in encouraging transactive conversational behavior between students, but, importantly, does not lead to a teacher-centered discussion. Instead, the teacher uses Academically Productive Talk to hold students accountable for their own knowledge and reasoning, and to remind them to hold themselves and each other accountable likewise. In studies where teachers used approaches like Academically Productive Talk, students have shown steep changes in achievement on standardized math scores, transfer to reading test scores, and retention of transfer for up to 3 years (Bill et al., 1992; Chapin et al., 2004). In another recent study, urban high-school teachers were trained in Academically Productive Talk practices. During the same period, the teachers' students participated in computer-supported collaborative learning activities that promoted Academically Productive Talk. Over the course of the study and especially following the interventions, the amount of Academically Productive Talk moves performed in the classroom was shown to increase (Clarke et al., this volume)

Script-Based Support for Collaborative Learning

The CSCL community shares many of the same values related to desired conversational practices in student group discussions. To support the growth of student discussion skills, we can design environments with affordances that play the same role as the teacher-as-discussion-facilitator.

The most popular approach to providing such affordances in the past decade has been that of script-based collaboration (Dillenbourg, 2002). A script may provide structure at a macro-level, perhaps dividing a collaborative task into roles for the participants to fulfill, or might scaffold a participant's contributions at a micro-level, with prompts to encourage a particular mode of argumentation. Such scripts are typically implemented statically, providing the same support in all cases. This is the work we review in this section. In the next section we describe a dynamic form of scripting that is capable of responding to changes in the state of the environment or discussion to deliver an appropriate level of support at opportune times.

A script may describe any of a wide range of features of collaborative activities, including its tasks, timing, the distribution of roles, and the methods and patterns of interaction between the participants. Scripts can be classified as either macro-scripts or micro-scripts (Dillenbourg, 2008). Macro-scripts are pedagogical models that describe coarse-grained features of a collaborative setting, that sequence and structure each phase of a group's activities to foster learning and social interaction. Micro-scripts, in contrast, are models of dialogue and argumentation that are embedded in the environment, and are intended to be adopted and progressively internalized by the participants. Scripts can be more or less coercive, from strict "follow me" style prompts to subtle suggestions of behavior implicit in the activity's structure. Stricter scripts can work to reduce the gap between expected and observed student behavior, producing a more uniform appearance of discussion, but run the risk over-scripting (Dillenbourg, 2002), where the application of inappropriate or unneeded supports have a detrimental effect on collaboration and learning.

Dynamic Script Based Support With Conversational Agents

Early approaches to scripting have been static, offering the same script or supports for every group in every context. Such non-adaptive approaches can lead to over-scripting, or to the interference between multiple scripts (Weinberger et al., 2007). More dynamic approaches can trigger scripted support in response to the automatic analysis of participant activity (Rosé et al., 2008). This analysis can occur at a macro-level, following the state of the activity as a whole, or it could be based on the micro-level classification of individual user contributions. The collaborative tutoring agents described by (Kumar & Rosé, 2011) were among the first to implement dynamic scripting in a CSCL environment. Scripting such as this offers the potential for minimal interventions to be used more precisely and to greater effect, with greater likelihood of students internalizing the support's intended interaction patterns. Further, the benefits of fading support over time (Wecker & Fischer, 2007) could be more fully realized, as the frequency of intervention could be tuned to the students' demonstrated competence. Indeed, conversational agents have been shown to be more effective when their interaction with students is in response to student initiative (Chaudhuri et al., 2009).

Participants in a collaborative session, including the facilitator, aren't simply focused on the task – they are involved in numerous simultaneous processes including social bonding, idea formation, argumentation, time management, and off-task activity. Just as human teachers orchestrate elements of collaborative learning in their classrooms, a conversational agent-as-facilitator must manage several differently-scoped supports and behaviors

concurrently. Recent work has produced software architectures for conversational agents (Kumar & Rosé, 2011; Adamson et al., 2012) that can implement such orchestration within CSCL environments.

Agents for Academically Productive Talk

Prior work with conversational agents and Academically Productive Talk has directed students to respond to each other with an array of Academically Productive Talk moves, in response to surface-level features of their contributions, with mixed results, prompting a redesign wherein the agent offered “Revoice” prompts that paraphrased student contributions when they were identified as conceptually-rich and relevant to the task (Dyke et al., 2012). Such an agent was shown to have a positive effect on learning and on conceptual richness of later student contributions. Criticism of the agents used in these studies (Stahl, 2013) suggests that the student experience could be improved by more finely targeting its interventions such that they are more responsive to (and not disruptive of) the flow of collaboration, and by minimizing the verbosity of each agent contribution.

We present a conversational agent behavior based on the “Agree-Disagree” Academically Productive Talk move as a dynamic support within a scripted CSCL environment, addressing some the limitations found in earlier work. In our implementation, the conversational agent acts as an instructor and facilitator, and presents a series of group exercises in ConcertChat, a discussion environment with a shared whiteboard (Mühlpfordt & Wessner, 2005). This environment is illustrated in Figure 1. As the group discusses each exercise, the agent monitors the chat for student assertions that could be followed up by a check for agreement or understanding. After such a candidate is identified, the agent waits to see if the students address the assertion on their own – if not, the agent offers a prompt to focus the group on the student’s contribution.

The screenshot shows a software interface for a CSCL environment. On the left is a whiteboard with a table and text. On the right is a chat window with messages from three students (S01, S02, S03) and a tutor named Quinn.

Whiteboard:

name	Formula	Boiling Point (K)
methyl vinyl ether	C ₃ H ₆ O	278.65
propanol	C ₃ H ₈ O	?
methyl ethyl ketone	C ₄ H ₈ O	352.79

6. Where does the boiling point of propanol (propyl alcohol) fall in relation to those of methyl vinyl ether and methyl ethyl ketone?

Discuss the following:

- what type of molecules these are
- what factors might be relevant to their boiling points
- whether those factors argue in favor of higher or lower boiling points.

When you've discussed these things with your group and come to a consensus, everyone should record their answer on their paper and then type "Ready"

Chat:

EDT: oh wait

S03 10/24/12 7:40:09 PM EDT: does that mean higher dipole moment

S01 10/24/12 7:40:09 PM EDT: that

S02 10/24/12 7:40:12 PM EDT: I think that the area is bigger

S02 10/24/12 7:40:18 PM EDT: So therefore the BP is bigger

Quinn (Tutor) 10/24/12 7:40:22 PM EDT: Can you weigh in on what **S03** said?

S02 10/24/12 7:40:26 PM EDT: So it's middle

S01 10/24/12 7:40:28 PM EDT: yeah but propanol is kind of polar

S01 10/24/12 7:40:29 PM EDT: like

S03 10/24/12 7:40:36 PM EDT: yeah propanol's really polarr

Figure 1. Screen shot of the CSCL environment where a group of 3 students is working together, supported by a tutor agent named Quinn, who participates with them in the chat.

Detecting Academically Productive Talk Candidates

In order to identify task-relevant conceptual assertions, we worked with domain experts and instructors to develop a “gold standard” list of statements that captured important concepts and misconceptions for the unit of study. Such statements were drawn from both the experts’ knowledge and expectations and from transcripts of an unsupported dry-run of the task. Using a “bag of synonyms” cosine similarity measure (Mihalcea et al., 2006), which essentially measures overlap in word usage, student assertions which are within a certain threshold of similarity to the gold statements are identified as *agree-disagree candidates* that could be evaluated by the group. This is the same detection technique used by the earlier Revoicing agent behavior (Dyke, et al. 2012), although as the agent does not need to produce an accurate paraphrase from the matched statements, a lower threshold can be used. This results in the detection of a greater number of candidate statements, and more

opportunities for support than the Revoicing agent could afford. Statements that match only the stricter threshold are also tracked – these *revoicable* assertions serve as a conservative indicator of conceptual, on-target contributions by each student. In earlier studies, the number of revoicable assertions was found to significantly correlate with learning. In this study, we expected the Agree-Disagree agent to intensify the contribution of this type of valued contribution by students.

Responding to Candidates

When a candidate statement is identified, the agent waits for the other students in the group to respond to it. If another student responds with an evaluation of their peer’s contribution (along the lines of “I agree” or “I think you’re wrong”), but doesn’t support the evaluation with an explanation, the agent will encourage this second student to provide one. If a student instead follows up with another APT candidate statement, the agent does nothing, leaving the floor open for productive student discussion to continue uninterrupted, reducing the risk of over-scripting their collaboration. If the other students do not respond with either an evaluation or a contentful followup, the agent prompts them to comment on the candidate statement – for example, “What do you think about Student’s idea? Do you agree or disagree?” This process is illustrated in Figure 2.

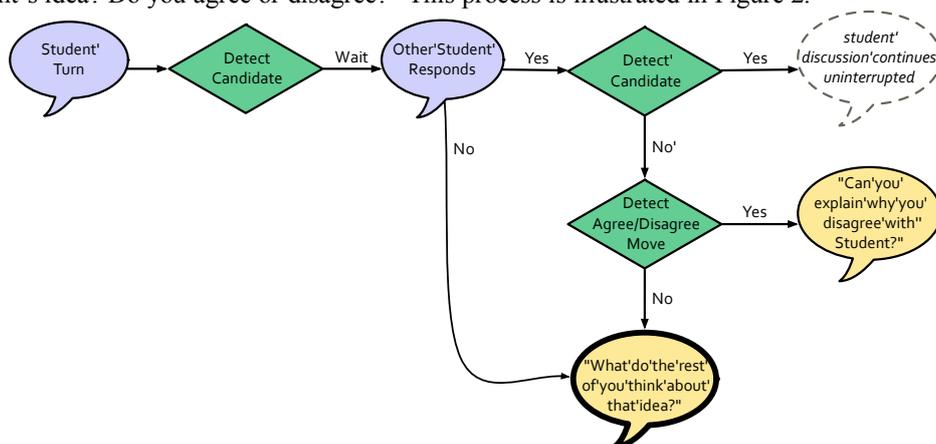


Figure 2. the Agree-Disagree Agent’s response to student statements

The excerpt shown in Table 2 is drawn from the study described in the next section. Times are given in seconds from the beginning of the excerpt, and the columns “Agree/Disagree Candidate” and “Student Evaluation” are the automatically detected labels the agent uses to motivate its facilitation moves. This exchange is typical of a group interaction in this environment – S07’s contribution at 17 seconds, although a candidate for evaluation, is not acted upon by the agent because S08’s followup preempts it. After 15 seconds following S08’s statement without any sort of uptake by the group, the agent prompts the group to agree or disagree with it. S09 offers a challenge, which leads to an extended back-and-forth between all three students.

Table 2: Selected interaction with the Agree-Disagree agent

Time	Author	Text	Agree/Disagree Candidate	Student Evaluation
00:00	S07	ok lots of things to do...		
00:13	S07	first one		
00:17	S07	surface area is higher	✓	
00:20	S09	arrow up?		
00:22	S08	ok, boiling pt will go up and vdW will go up for all of them consecutively... right?	✓	
00:37	TUTOR	Do you concur with S08? Why, or why not?		
00:41	S09	hmm not necessarily		✓
00:47	S07	area goes up for each	✓	
00:50	S09	would it?		
00:51	S09	im not sure		✓
00:56	S08	yea for sure area goes up	✓	
01:10	S07	dipole increases first one		

While this approach goes far in providing productive prompts at appropriate points, students can still be thrown off by these interventions. In Table 3, the agent does not identify the ongoing exchange as relevant to the discussion, and thus does not suppress its prompt for evaluating S08’s earlier statement. This causes

confusion for S08, who is unclear about which of their messages the agent is referring to. Occasional missteps such as this do not appear to utterly derail the group and, the agent is generally accepted as a facilitator and its prompts are taken as opportunities for reflection.

Table 3: Infelicitous interaction with the Agree-Disagree agent

Time	Author	Text	Agree/Disagree Candidate	Student Evaluation
01:10	S07	dipole increases first one		
01:13	S08	dipole moment is based on the whole thing though, and it's tetrahedral... agh	✓	
01:14	S07	then its symmetric?		
01:16	S08	shapes are hard		
01:19	S07	so decrease		
01:24	S07	and then increase		
01:27	TUTOR	What do you think about S08's idea? Do you agree or disagree?		
01:27	S07	and then decrease?		
01:29	S08	wait what?		
01:49	S08	TUTOR, if it ends in a question mark its probs not an idea		✓
01:54	S07	CF4 is symmetric so dipole would be 0?	✓	

Other Agent Behaviors

We employ the Bazaar agent framework (Adamson et al., 2012) to dynamically orchestrate the full set of agent behaviors, prioritizing and regulating the proposed contributions from each of the agent's components proposed contributions, so as to avoid interference between components, and to work with the flow of the group's conversation. In addition to the agree-disagree behavior described above, the agent executes a flexibly-timed macro-script to present a series of instructional materials and exercises on the group's shared whiteboard. This script begins when a sufficient number of students have joined the group chat. When all students indicate that they are ready to proceed to the next phase of the task, the agent clears the whiteboard and presents the material for the next problem. The agent also implements a set of social support moves, providing responses to student behavioral cues in order to promote group bonding and task-oriented positivity. Such support has been shown to correlate with gains in student learning and perception of the agent (Kumar et al., 2010, Ai et al., 2011).

Method

To investigate the efficacy of the Agree-Disagree agent as a way to promote student interaction and critical thinking, we situated our study within a first-year undergraduate chemistry course.

Participants

The participants in our study were first-year undergraduate students studying intermolecular forces. Students were randomly assigned to groups of 3 or 4, and then groups were randomly assigned to conditions. The balance of 3 and 4 person groups was even between conditions, and there was no effect of team size on any of our dependent measures. All students in the course were required to participate in the online exercise for course credit, but they had the option of not consenting for their data to be included in our research. Thus, we only report results for consenting students. Altogether, our analysis includes data from 18 students from 6 different groups, which is 9 students and 3 groups in each condition. We employ multi-level modeling techniques in our analyses of results in order to account for the statistical dependencies between data from students in the same group.

Task

The collaborative task focused on intermolecular forces and their influence on the boiling points of liquids. For each problem in the activity (illustrated in Figure 1), students were asked to predict whether a given substance would have a higher or lower boiling point than two of its relatives, explaining their reasoning about the set of molecules in terms of their structure and the forces at play. Each problem of this sort was followed up by revealing the actual boiling point of the mystery molecule, and asking students to revisit their predictions and explanations in light of the new data. A liquid's boiling point can be influenced simultaneously by a number of different intermolecular forces, each of which arises as a consequence of the molecules' particular structural attributes. Correctly identifying the pertinent structural features of molecules and reasoning about how they will

affect the liquid's boiling point is a non-trivial and multi-faceted task. Because multiple types of intermolecular forces influence liquids' boiling points, we used the Jigsaw technique (Aronson, Blaney, Stephan, Sikes, & Snapp, 1978), assigning students within each group to read individually about one of three forces that contribute to a molecule's boiling point. In cases where a four-person group was formed, the fourth student received the same training material as the first student. This division also provided intrinsic motivation for collaboration, as the task could not be completed without knowledge from each of the student experts.

Experimental Design

Our experimental design was a simple 2-condition between-subjects design where teams were assigned randomly either to the Agree-Disagree condition or the Control condition. Both conditions were identical except for inclusion of the Agree-Disagree facilitation move by the agent. Thus, both conditions benefitted both from macro-level and micro-level script based support. In the Agree-Disagree condition, whenever the agent was not engaged in a directed dialog, it was receptive to opportunities to dynamically support the conversation by requesting students to evaluate whether they agreed or disagreed with assertions that were made in the chat, as discussed above.

Pre/Post Tests

Pre and Post tests were used to measure learning during the collaborative exercise. We used two isomorphic versions of the test (Version A and Version B) and counter-balanced their assignment such that half of the students received A as a pretest and B as a posttest, while the other half of students received B as pretest and A as posttest. There was no significant difference between scores on A and B.

Process Analysis

The goal of the Agree/Disagree agent was to engage students in a more intensive exchange of explanations (revoicable assertions), to raise the level of critical thinking. Thus, in addition to a Pre/Post test measure of learning, a process analysis is also important for evaluating our hypothesis. Variables related to the elicited conversational behavior may then be examined in order to test whether they served a mediating or moderating effect on learning. In order to accomplish this, the chat logs were segmented into 2 minute intervals such that one observation was extracted per student for each interval. In each observation, we counted the number of revoicable assertions contributed by the student, the number of revoicable assertions contributed by other group members, the number of Agree-Disagree prompts targeted at the student in the previous time slice, and the number of Agree-Disagree prompts targeted at other students in the group in the previous time slice.

We can evaluate the effect of condition on the correlation within time slices between occurrences of revoicable assertions of a student with those of the other students in the same group. We used a multi-level model to analyze the results in order to account for group effects. We expect to see that the correlation is significantly higher in the condition with the Agree/Disagree agent. Specifically, we used what is referred to as a *random intercept and slope model*, which allows estimating a separate latent trajectory for a student's behavior in relation to that of their partner students within time slices. In this model, each student trajectory is characterized by a regression with latent slope and intercept, relative to a slope and intercept per group, which are in turn relative to the global model's slope and intercept. To do this analysis, we used the Generalized Linear Latent and Mixed Models (GLLAMM) (Rabe-Hesketh, Skrondal, & Pickles, 2004) add-on to STATA (Rabe-Hesketh & Skrondal, 2012). The dependent measure was number of revoicable assertions by the student within the time slice. The independent variable was the number of revoicable assertions contributed by the other students in the group within the same time slice. The condition variable was added as a fixed effect, and as an interaction term with the independent variable. A significant interaction between condition and independent variable in this case would indicate a significant difference in correlation between a student's contribution of revoicable assertions and that of their partner students.

Results/Analysis

Our hypothesis was that the introduction of the Agree/Disagree agent would intensify the interaction between students, which might increase critical thinking, and subsequently increase learning. Our analysis offers qualified support for the hypothesis.

First we evaluated the effect of condition on learning. For this analysis, we tested for any significant difference in pretest scores between conditions using an ANOVA with pretest as a dependent variable, Condition as an independent variable, and Group as a random variable nested within condition in order to account for the non-independence between data collected from students who worked in the same group. There was no significant or marginal effect of Group on pretest scores, confirming that students were distributed with sufficient randomness between groups. There was no significant or marginal difference between conditions on pretest score, though there was a trend for students in the Agree/Disagree condition to have lower pretest scores.

Table 4: Summary of Results

	Agree/Disagree Condition	Control Condition
Pretest	6.14 (3.7)	6.73 (3.1)
Posttest	9.99 (2.72)	8.81 (3.3)
Revoicable Assertions	7.89 (3.8)	7.33 (4.3)

Thus to evaluate the effect of condition on learning, we used an ANCOVA with posttest as the dependent variable, pretest as a covariate, Condition as an independent variable, and Group nested within condition as a random variable. In this analysis, there was a marginal effect of Condition on learning ($F(1,11) = 1.82, p < .1$, effect size .55 standard deviations), such that students in the Agree/Disagree condition learned more. The effect was moderate.

Next we examined the intensifying effect of the intervention on the interaction between students. We evaluated this by looking for evidence that the Agree/Disagree prompts increased the extent to which students constructed knowledge together, at least in pockets of intensive knowledge exchange. As can be seen in the conversation excerpts above, students contribute a variety of types of contributions, not all of which are revoicable assertions. However, when they are engaged in intensive exchange of ideas with one another, we find regions of the conversation with denser concentrations of revoicable assertions, because when one student offers his perspective, others tend to follow up with their own. If the discussion is divided into time slices, we can distinguish transcripts that contain regions of dense group knowledge construction from those where students present their ideas intermittently without precipitating intensive group knowledge construction. We do this by looking at the correlation of the count of a student's revoicable assertions with the count of revoicable assertions from other students in the group, within each time slice. In the first case, we expect that there will be many time slices where there are revoicable assertions from both the student and the other students in the group, whereas in the second condition, we don't expect to see this occur frequently.

The analysis using the random intercept and slope model described in the Methods section showed the pattern that we expected. There was no significant difference in intercept between conditions, confirming that, as we suspect from Table 4, there was no difference in absolute number of revoicable assertions between conditions. However, this is not problematic since the number of revoicable assertions was found to have a moderating but not mediating effect on learning. Specifically, when the revoicable assertions variable was added to the ANCOVA evaluating the effect of Condition on learning as an additional covariate, it had a significant positive correlation with posttest score that increased the percent of posttest variance explained from 69% to 83% but did not reduce the effect of Condition on learning. Thus, we must conclude that the effect of condition on learning is not explainable by this simple summative measure.

More importantly, there was no significant correlation between the number of revoicable assertions of a student and that of his partner students in the control condition where there was not an Agree/Disagree agent. However, there was a significant interaction between the condition variable and the number of revoicable assertions contributed by partner students ($R = .14, z = 2.03, p < .05$), indicating that in the Agree/Disagree condition, there was a significant positive correlation between the number of revoicable assertions contributed by a student and that contributed by partner students. Thus, we do see evidence that the intervention had the effect of precipitating pockets of intensive discussion.

We then evaluated the extent to which this effect was explained by the local presence of Agree/Disagree prompts. Surprisingly, a student contributes significantly more revoicable assertions in time slices following ones wherein the agent prompted the other students to agree or disagree with that student ($F(1,847) = 4.9, p < .05$, effect size .35 standard deviations) but not when the agent asked the group to agree or disagree with a different student (no significant effect). And time slices with revoicable assertions from both students were not primarily the same ones that contained prompts for Agree/Disagree. This suggests that the primary, or at least first, effect of the prompt may in fact be to elicit followup explanations from a student rather than to elicit feedback from the other students. Seen in conjunction with the correlation analysis above, it is possible that the prompts more often first elicited followup explanation from the student who contributed the initial agree-disagree candidate, and in response to this elaboration, the other students were drawn in and responded in turn. Thus, we see a subtle ripple effect of the intervention that is not easily quantified, even in the analysis of intensification above.

Discussion

We have described and demonstrated the effectiveness of a new conversational agent behavior in a college chemistry context. Advances in its design that address sensitivity to the flow and content of student conversation differentiate this agent from similar agents in earlier work, allowing the facilitative behavior to be minimally intrusive while still actively promoting rich student-centered discussion. Future work with larger samples should provide clarification and amplification of the positive learning trend seen here. We look forward to future studies where conversational agents successfully orchestrate multiple strategies drawn from Academically Productive Talk and other instructional discourse frameworks, to provide many-dimensional support for group

collaboration and productive discussion. Such agents may be critical in fostering effective conversation in the rapidly growing domain of distributed-learning university courses.

References

- Adamson, D., & Rosé, C. (2012). Coordinating multi-dimensional support in collaborative conversational agents. In *Proceedings of Intelligent Tutoring Systems* (pp. 346-351). Springer Berlin/Heidelberg.
- Aronson, E., Blaney, N., Stephan, C., Sikes, J., & Snapp, M. (1978). *The jigsaw classroom*. Sage Publications.
- Bill, V. L., Leer, M. N., Reams, L. E., & Resnick, L. B. (1992). From cupcakes to equations: The structure of discourse in a primary mathematics classroom. *Verbum*, 1, 2, 63-85.
- Chapin, S., O'Connor, C., & Anderson, N. (2003). *Classroom Discussions. Using Math Talk to Help Students Learn, Grades 1, 6*.
- Chaudhuri, S., Kumar, R., Howley, I., & Rosé, C. P. (2009, July). Engaging collaborative learners with helping agents. In *Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling* (pp. 365-372). IOS Press.
- Clarke, S., Chen, G., Stainton, C., Katz, S., Greeno, J.G., Resnick, L.B., Howley, I., Adamson, D. and Rosé, C.P. (2013). The Impact of CSCL beyond the Online Environment. *Proceedings of the 10th International Conference on Computer Supported Collaborative Learning*, July 2013.
- Dillenbourg, P. (2002). Over-scripting CSCL : The risks of blending collaborative learning with instructional design . *Three worlds of CSCL - Can we support CSCL?* pp. 61–91 .
- Dillenbourg, P., Hong, F. (2008) The mechanics of CSCL macro scripts. *International Journal of Computer-Supported Collaborative Learning* 3(1), 5–23.
- Dyke, G., Howley, I., Adamson, D., Rosé, C.P. (2012). Towards Academically Productive Talk Supported by Conversational Agents. *Proceedings of Intelligent Tutoring Systems 2012*.
- Howley, I., Adamson, D., Dyke, G., Mayfield, E., Beuth, J., Rosé, C.P. (2012). Group Composition and Intelligent Dialogue Tutors for Impacting Students Academic Self-Efficacy. *Proceedings of Intelligent Tutoring Systems 2012*.
- Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning* 2(2-3), 211–224.
- Kollar, I., Fischer, F., Hesse, F.W. (2006). Collaborative scripts - a conceptual analysis. *Educational Psychology Review* 18(2), 159–185.
- Kumar, R., Rosé, C. P., Wang, Y. C., Joshi, M., and Robinson, A. (2007). Tutorial dialogue as adaptive collaborative learning support. *Proceedings of Artificial Intelligence in Education*.
- Kumar, R., Rosé, C.P. (2011). Architecture for Building Conversational Agents that Support Collaborative Learning. *IEEE Transactions on Learning Technologies* 4(1), 21-34.
- Mihalcea, R., Corley, C., and Strapparava, C. (2006). Corpus-based and knowledge-based measures of text semantic similarity, *Proceedings of the National Conference on Artificial Intelligence (AAAI 2006)*, Boston, Massachusetts, pp. 775-780.
- Mühlpfordt, M., Wessner, M. (2005). Explicit referencing in chat supports collaborative learning, *Proceedings of Computer Support for Collaborative Learning (CSCL)*.
- Rabe-Hesketh, S. & Skrondal, A. (2012). *Multilevel and Longitudinal Modeling Using Stata*, Stata Press.
- Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). *GLLAMM Manual*. University of California, Berkely. U. C. Berkeley Division of Biostatistics Working Paper Series, Paper 160.
- Rosé, C., Wang, Y. C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., & Fischer, F. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, 3(3), 237-271.
- Stahl, G. (2013). Interaction analysis of a biology chat. In D. Suthers, K. Lund, C. P. Rosé & N. Law (Eds.), *Productive multivocality in the analysis of group interactions*. New York, NY: Springer. Web: <http://GerryStahl.net/pub/multivocal.pdf>.
- Wecker, C., Fischer, F. (2007). Fading scripts in computer-supported collaborative learning: The role of distributed monitoring. *Proceedings of the 8th International conference on Computer Supported Collaborative Learning*, pp. 764–772 .
- Weinberger, A., Stegmann, K., Fischer, F., Mandl, H. (2007). Scripting argumentative knowledge construction in computer-supported learning environments. In *Scripting Computer-Supported Collaborative Learning*, CSCL Book Series Volume 6, Chapter 6, pp191–211 .

Acknowledgements

This work was supported in part by NSF grant SBE 0836012 to the Pittsburgh Science of Learning Center.

Variation in Other-Regulation and the Implications for Competence Negotiation

Karlyn R. Adams-Wiggins, Toni Kempler Rogat, Rutgers University, Graduate School of Education,
10 Seminary Place, New Brunswick, NJ 08901-1183
Email: karlyn.adams.wiggins@rutgers.edu, toni.kempler.rogat@gse.rutgers.edu

Abstract: To succeed, groups need skills to jointly regulate their shared task work. The current study examines variation in other-regulation, or efforts by one student to regulate their group's work. We consider the relationship of directive and facilitative forms of other-regulation with efforts to negotiate competence, given that directive other-regulators may raise doubts about relative ability. Three groups of four 7th grade students were observed while working on two collaborative activities during an inquiry-based science unit. Results suggest the nature and quality of facilitative and directive other-regulation varies, with directive regulators focused on controlling the task product in ways that excluded others' attempts to contribute. In response, teammates worked to renegotiate their positions of competence within the group to ensure their ideas were considered for integration. The focus on relative competence promoted by directive other-regulation may diminish a focus on group learning, given the social nature of joint activity.

Successful teamwork is increasingly necessary for learning in and beyond school (Strijbos, Kirschner & Martens, 2004). To succeed, groups need skills for jointly coordinating and regulating work on a shared task product. Recent research has expanded prior emphases on individual self-regulated learning to consider the contextualized nature of students' experiences during group work, with the ultimate aim of understanding group's regulation of behavior, learning, and understanding during shared activity (Volet, Vauras & Salonen, 2009). Social regulation research has focused on who is regulating within the group indicating a range from other-regulation or coregulation, in which one student temporarily predominates the group's interactions, to socially shared regulation, whereby multiple group members jointly regulate group activity (Vauras, Iiskala, Kajamies, Kinnunen, & Lehtinen, 2003). While other-regulation is typically conceptualized as a group member temporarily guiding others' understanding (Hadwin & Oshige, 2011), there is some evidence that other-regulation may be stable once a group leader is established (Li, et al., 2007) and more directive in conducting the regulatory processes for the group (Rogat & Linnenbrink-Garcia, 2011). Rogat and Linnenbrink-Garcia's (2011) results characterized *directive other-regulation* as one group member's efforts at determining the next step of the task, detailing exactly what group members should do, and maintaining control of monitoring and task contributions. These findings indicate that there may be a broader spectrum of other-regulatory behaviors that go beyond facilitation to include more directive forms. Beyond more clearly understanding the nature of other-regulation in collaborative groups, it is also critical to understand the implications for group process. Research indicates some evidence that, unlike prior conceptualization of other-regulation, directive forms can provoke conflict in response to group members' regulation being ignored and having limited means for contributing to the task (Rogat & Linnenbrink-Garcia, 2011). Those subjected to other-regulation may come to experience their participation as peripheral and in conflict with the other-regulator, particularly when their regulatory and content contributions are not integrated. This lies in contrast to what we expect groups engaging in a guiding, or *facilitative*, form of other-regulation to experience: disagreement instead could promote co-construction of meaning and task improvement (Darnon, et al., 2006).

The current study has two goals. First, we examine other-regulation within collaborative groups to more richly understand the regulatory processes employed by directive and facilitative other-regulators. Towards this end, we designate each group's other-regulator by examining frequency and types of regulatory strategies employed by all members of the group, using distinctions made in prior research, such as regulation of content understanding, use of disciplinary norms, task process, group process, and behavior (e.g. Hogan, Nastasi & Pressley, 1998; Rogat & Linnenbrink-Garcia, 2011; Salonen, Vauras & Efklides, 2005). We also qualitatively examine these social regulatory processes to more richly characterize the nature of the employed other-regulation. Understanding variation in other-regulation is critical since, with directive forms, regulation of the group may be low quality and limit equitable involvement in ways that convey information about group members' competence.

The second goal of the research is to investigate the relationship between variation in other-regulation and group process. Specifically, we explore how other-regulation impacts efforts by group members to negotiate their own and others' positions of competence. In the context of directive other-regulation, group members may experience constrained opportunities to participate, and thus have fewer opportunities to engage as competent contributors. Group members being repeatedly ignored and having their contributions treated as irrelevant when

in conflict with the directive other-regulator's ideas may introduce doubts about relative competence (Darnon, et al., 2006). In response to relative ability being made salient and questioning the competency of one's own contributions, group members may engage in competence negotiation. In this work, we conceptualize competence negotiation by drawing on the literature on social comparison and systems of competence to investigate how individual students work to reposition themselves and their peers within the system of competence (Altermatt, et al., 2002; Darnon, et al., 2006; Gresalfi et al., 2009). Facilitative other-regulators may correspondingly work to ensure everyone ideas are solicited, fostering inclusion and respect for everyone's contributions, thus reducing the need to negotiate relative competence. Toward this end, we examine how individuals within groups work to ensure that their ideas are not ignored and are considered for integration. We expect group members with directive other-regulators to assert their own competence by *self-advocating* or *self-presenting*. We also explore whether there are moves to reposition others as more or less competent by promoting or criticizing other's contributions (i.e., *other-positive* or *other-negative*). Fundamentally, it is critical to consider implications of other-regulation since social comparison can have negative consequences for group functioning. If directive other-regulation promotes a focus on relative competence it may ultimately discourage a shared focus on learning from the task (Ames, 1992) and disengagement (Nolen, 2007), with implications for diminished conceptual understanding during collaborative tasks given the social nature of joint activity (Barron, 2000).

Method

Three 4-person groups of 7th grade students were observed during two inquiry-based science tasks focused on cell organelles and the development of reasoning skills. Groups were purposefully selected to afford exploration of variation in other-regulation. We selected groups that appeared to vary in the degree of balance of participation among group members, without including extreme cases (Patton, 1990). Two observations per group were selected that lasted at least five minutes, had minimal off-task behavior, and involved a collaborative task (excluding pair work and teacher-led tasks).

Elaborated running records were prepared from video-taped observations to contain information about body language and gestures. Next, we coded the records for social regulation. Sub-codes then were applied to differentiate regulatory types (see Table 1), and each instance was designated as taken up, ignored, or rejected with or without rationale (Barron, 2000). Frequencies and percentages were calculated for participation, regulatory moves, and responses to regulation. Other-regulators were identified by their frequent regulatory contributions relative to their group members as well as their broad use of types of social regulation (i.e. regulated more areas than most group members). Subsequent qualitative analysis of the regulation employed by these other-regulators informed our designations of the type of other-regulation as facilitative or directive. We also coded for attempts by individuals to negotiate their position of competence or that of others (see Table 1). Individuals within the group can attempt to convey that they are capable contributors via self-presentation (i.e., self-positive) or self-advocating (see Barron, 2000). Competence negotiation can also be targeted toward one's group members in efforts to promote and advocate for other's contributions (i.e., other-advocate and other-positive) or by diminishing the competence of one's teammate (i.e., other-negative). Other types of competence-relevant language which may contribute to negotiation include self-deprecation (self-negative) and group-targeted statements (group-positive, group-negative). Further, groups who focus on relative ability may shift between discussing ability within the group to discussing between-group ability comparisons (i.e., group-positive and group-negative) (Kempner & Linnenbrink, 2004). Both explicit (e.g., I am smart) and implicit (e.g. refusing or soliciting help from a particular group member) evaluative statements were considered evidence of competence negotiation given findings that older children typically rely on subtle forms of social comparison (Altermatt et al., 2002). After coding the observations, reliability was established and disagreements were resolved to yield final codes.

Table 1: Description of Codes

Codes	Description
<i>Regulation Types</i>	
Content	Regulatory moves focus on the group's understanding or use of content
Disciplinary	Focus on ensuring group's adherence to norms of disciplinary practice
Task	Regulation specific to task components, directions, procedure, and enacting task
Group Process	Focus on coordinating group interactions and turn order
Behavioral	Specific to re-engaging off-task group members and sustaining on-task behavior
<i>Competence Positioning</i>	
Self-positive	Positive self-targeted statements that aren't inclusive of group (e.g. "I'm right")
Self-negative	Self-deprecating comments that aren't inclusive of group (e.g. "I'm so stupid")

Other-positive	Positive comments regarding a group member's capabilities (e.g. "That's a great idea")
Other-negative	Negative deprecating comments targeting a group member (e.g. "You're wrong")
Group-positive	Positive statements about own group's competence (e.g. "We're good at this")
Group-negative	Negative statements about own group's capabilities (e.g. "We are behind")
Self-advocate	Pushing for one's own perspective to be heard by the group by repeating claims or re-asserting claims not previously taken up
Other-advocate	Promoting contributions of a groupmate (e.g. intervening when someone is excluded or ignored, asking to return to a previously dismissed idea)

Results

We began by exploring differences in other-regulation among the groups by examining whether the amount of talk contributed by students to discussion was equitable using percentage of total turns taken by each group member (Hogan, et al., 1999). Results confirmed the variation in other-regulation in line with our purposeful sampling of groups (see method and Table 2). In Group 1, all four students participated relatively equally during group discussion, while Groups 2 and 3 showed more imbalanced contributions among group members with both groups having a single member who participated more frequently than others. Group 2 was differentiated by having one group member who evidenced very limited participation, while Group 3's remaining group members showed more equal participation. To identify the other-regulators, we also examined frequencies of participation, total regulation and counts of regulation types employed by members of the group (see Table 2). Below we characterize the nature of each group's other-regulation given these frequency data and a qualitative analysis of the how the other-regulator engaged in regulatory processes for the group.

Table 2: Frequencies of regulation types by individuals within groups and participation

		Content		Disciplinary		Task		Group Process		Behavioral		Total Regulation		Participa-tion - % of total turns
		No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	%
Group 1	Allison	1	9	8	36	5	22	20	42	2	14	36	31	28.3
	Bob	2	18	1	5	1	4	9	19	2	14	15	13	24
	Cindy	4	36	3	14	3	13	7	15	0	0	17	14	20.5
	Donna	4	36	10	45	14	61	12	25	10	71	50	42	27.1
	<i>Totals</i>	<i>11</i>	<i>--</i>	<i>22</i>	<i>--</i>	<i>23</i>	<i>--</i>	<i>48</i>	<i>--</i>	<i>14</i>	<i>--</i>	<i>118</i>	<i>--</i>	
Group 2	Amy	4	18	2	29	22	46	4	31	3	38	35	36	27.9
	Billy	0	0	0	0	1	2	1	8	0	0	2	2	1.9
	Carla	9	41	2	29	12	25	5	39	5	63	33	34	42.3
	David	9	41	3	43	13	27	3	23	0	0	28	29	27.9
	<i>Totals</i>	<i>22</i>	<i>--</i>	<i>7</i>	<i>--</i>	<i>48</i>	<i>--</i>	<i>13</i>	<i>--</i>	<i>8</i>	<i>--</i>	<i>98</i>	<i>--</i>	
Group 3	Adam	8	24	2	100	4	29	0	0	0	0	14	25	23
	Bridget	0	0	0	0	1	7	0	0	0	0	1	2	13.1
	Carol	11	33	0	0	7	50	2	29	0	0	20	35	41.8
	Dylan	15	44	0	0	2	14	5	71	0	0	22	39	22.1
	<i>Totals</i>	<i>34</i>	<i>--</i>	<i>2</i>	<i>--</i>	<i>14</i>	<i>--</i>	<i>7</i>	<i>--</i>	<i>0</i>	<i>--</i>	<i>57</i>	<i>--</i>	

Note: This table presents the frequencies and percentages of social regulation types that occurred during the three groups' discussions. Percentages assist group comparisons since observation length and on-task engagement varied between groups. In addition, other-regulators for each group are designated using bold font. Finally, total number of turns for each group was - Group 1: 258 turns; Group 2: 215 turns; Group 3: 122 turns.

Variation in Other-Regulation

Donna was identified as Group 1's other-regulator. She engaged more frequently and broadly than her groupmates in all forms of social regulation, with the exception of content regulation in which her frequencies were typical for her group. Donna's other-regulation can be characterized as facilitative: she regularly posed task regulation questions to evoke widespread participation (e.g. "This evidence, okay well how does it relate to the movement and energy models?"), brought the group back to task (e.g. asking "What'd you guys write this as?" during off-task behavior), and regulated group process in ways that prevented exclusion of Bob's minority perspective on Task 1 (e.g. "Now we're gonna add 'How would you rate this, Bob?"). Donna also kept the group accountable to the class' jointly created disciplinary criteria by referencing the norms during discussion

(e.g. “The most important thing of good evidence is if it’s relevant. If it’s irrelevant, it’s not good.”). Allison also made regulatory moves frequently in Group 1, but she most heavily regulated group process and disciplinary practice, with her group process regulation primarily focused on turn order. Bob and Cindy both regulated less often overall, but also tended to regulate group process. Overall, Donna’s facilitative other-regulation generally prompted all group members to participate and thus fostered equal access to knowledge construction, widespread inclusion and integration of ideas.

Group 2’s participation was skewed and involved directive other-regulation by Carla. Carla’s regulation was focused on task and content, with many moves involving directive statements that offered little room for response from groupmates (e.g. “Get a calculator”). Further, Carla’s directive regulation was low quality, with few rationales and a focus on simply besting others. David and Carla battled for control over whose regulation and positions would be taken up in the group. In one example where the task required discussing evidence quality, Carla argued for Evidence F and David for Evidence C. David and Carla battled over which evidence was best in extended and heated fifteen turn exchange. Carla said “I don’t like it. F. F!” and David yelled back “C!” Further, David routinely made Carla the target of his regulation, responding to her directive approach (e.g. “Add it up!”). Of the remaining group members, Billy largely did not contribute to the group and his views were not solicited. Amy mostly made moves to regulate the task, such as checking what steps the group would take, but was largely ignored. Amy used disciplinary regulation to intervene during Carla and David’s arguments, but she was ignored every time. Thus, Group 2’s directive other-regulation was low quality given limited elaboration, highly critical monitoring, as well as limited opportunities for all group members to contribute.

Group 3’s other-regulation can be characterized as directive, although the participation was more balanced than Group 2’s. While Dylan regulated at a similar frequency to Carol, it was Carol who used regulation to ensure that her response was incorporated in the final product and whose ideas were not ignored. Characteristic of Carol’s other-regulation, she engaged in some high quality content and task regulation, with her regulation including rationales as feedback. However, these instances of high quality regulation often involved a single partner with whom she chose to engage, while excluding others. For example, while arguing with Dylan about how to interpret the evidence during Task 1 (i.e., content and disciplinary regulation), Carol did not make efforts to include Bridget or Adam using group process regulation. Also, Adam’s many attempts to regulate were ignored despite the high quality of his regulation during Task 1. In one example, Adam provided disciplinary monitoring by indicating that the group was relying on inference, rather than focusing on the evidence. Overall, Group 3’s pattern suggested that Carol dominated the group’s regulation in ways that primarily involved ignoring group members’ regulatory contributions: her directive other-regulation involved excluding ideas that contradicted hers and refusing to incorporate monitoring offered by groupmates.

In summary, our results point to two forms of other-regulation which can be differentiated in nature and quality. For Group 1, facilitative other-regulation involved inclusion of everyone’s contributions through the regulation of group process, behavior, and task in ways that afforded co-construction of knowledge and coordinated work on the shared product. In contrast, the other-regulation in Group 2 and 3 can be characterized as directive. Group 2 exhibited directive regulation leading to reduced access to participation opportunities and informational feedback, with regulation that could be characterized as harsh (i.e., battling, ignoring). Group 3’s other-regulation involved a directive other-regulator working to control the group’s final product. While at times her content and task regulation provided feedback, she simultaneously failed to be inclusive of other group member’s contributions.

Competence Negotiation

In this section we explore whether differences in other-regulation co-occurred with attempts to negotiate competence. More specifically, given Darnon and colleagues’ (2006) findings that conflict can promote a focus on relative competence, directive forms of other-regulation may lead group members to experience threats to their competence. In response, students may resort to promoting their competence via self-presentation or self-advocating, and in some cases, by putting down teammates.

We turn now to an analysis of competence positioning moves in these groups. Competence positioning differed between groups in its frequency and function (Table 3). While Group 1 and Group 3 had a similar tally of positioning, Group 2 engaged in three times as many competence positioning moves relative to the other groups. Further qualitative analysis suggested that this was representative of high salience of relative competence in the group’s interactions. We engaged in qualitative analysis to consider how negotiating competence functioned within each group. In the following sections, we first discuss how messages regarding relative ability were made salient via competence messages within the group context in ways that provided background for negotiating competence, followed by discussing individuals’ efforts made to present themselves as competent within the group.

Table 3: Frequency of Competence Positioning Moves by Group

Competence Positioning Moves	Group 1	Group 2	Group 3
Self-positive	0	11	5
Self-negative	0	1	0
Other-positive	2	0	1
Other-negative	1	25	9
Group-positive	0	0	0
Group-negative	0	2	0
Self-advocate	18	40	13
Other-advocate	2	1	0
Framing Task Competence	1	0	0
Totals	24	80	27

Group-level Efforts to Negotiate Competence

Group 1 can be differentiated by competence positioning moves that were positive and affirmed the competence of all group members. Four of their moves evidenced bolstering the competence of fellow group members rather than oneself (i.e., 2 other-positive, 2 other-advocate). Specifically, the facilitator, Donna, helped to ensure that everyone's views were valued during group discussion. For instance, Donna other-advocated for Bob to ensure his perspective was heard and prevent his view from being prematurely rejected. Here, Donna requested that Cindy stop interrupting Bob because the group would not know his contribution unless he had the opportunity to voice it. Also, on two separate occasions Bob and Donna acknowledged each other's correct contributions (i.e., other-positive) after having voiced opposing views. For example, Donna chuckled while voicing agreement with Bob's statement that their prior argument justified to the conclusion that his claim was plausible. Next, Allison offered the sole other-negative comment to Bob in jest: "No, you're wrong this time. You said 3." Despite the underlying poke at Bob's competence, this comment can be interpreted in the context of friendly competition as was indicated by Allison's follow-up statement to Donna recognizing the validity of Bob's contribution: "Look, him 1 [point], us 1. We're tied." Bob followed this exchange with an other-positive comment, "Because even though you guys are right it doesn't relate to the model, it's still good evidence." In sum, Donna's other-regulation ensured that everyone's views were treated as competent and valued. This seemed to foster positive valuing of ideas, complimenting contributions, and even other-advocating for the inclusion and respect of the minority perspective within the group.

We observed negotiation that involved dismissiveness and criticism of group members' contributions which devalued members' competence in the two groups with directive other-regulators. Group 2's positioning involved a focus on mistakes paired with mostly explicit relative competence messages. There were 25 instances of other-negative comments; some of the most striking included Carla singling out David with salient references to relative competence in his class standing saying "You're the only one in this class who likes C" and repeating "Don't ever grow up to be a scientist." Even when engaged in a hypermedia task that included a team game, Carla and David repeatedly pointed out each other's errors. For example, Carla asked "How are you in honors literacy?" when David was slow to read from the screen. David repeatedly pointed out to the group that Carla was to blame for the computer's malfunction (e.g. "You broke it! You broke it!"; "Carla broke the computer!"). These criticisms are representative of Carla and David's interactions with each other: the two regularly putdown and criticized one another's competence, with some criticisms extending beyond the immediate group context to each other's more global competence (e.g. literacy, use of technology, career choices). Use of negative competence messages allowed David and Carla to restrict access to opportunities for others to be competent contributors. Notably, the highly salient relative ability comments escalated among other group members. Amy made two group-negative statements that served to compare their group with the other groups (e.g. "Hurry up, I think we're the only group still doing this"; group-negative). The salience of competence produced putdowns of groupmates and even between-group social comparison. In sum, Group 2 stood out as being a highly competitive context rife with competence threats: there were high stakes for perceived incompetence as group members were very publicly and harshly recognized for incorrectness in ways that diminished their contributions' value in the group.

Group 3's directive other-regulation was similarly linked to putdowns and restricted access to contribution opportunities, but less frequently than observed for Group 2. Further, Carol used implicit rather than explicit messages to convey competence. Carol both ignored and dismissed group members' ideas. She tended to ignore group members with whom she was not directly working; both her partner Bridget and Adam were ignored when Carol was busy trying to delegitimize Dylan's argument. Relative to the group context, Carol also criticized group members' views that did not agree with or validate her own (i.e., other-negative). Several moves were made in reaction to Carol's efforts at thwarting their competence. Often, Dylan responded to Carol with other-negative comments that saliently focused on her ability. Dylan stated "You're stupid!

“You’re so stupid!” It is critical to also highlight that the group responded counterproductively to Adam’s attempts at recognizing each group member’s valuable contributions through other-positive competence positioning. In particular, when Adam recognized both Dylan and Carol’s views as having merit (“I know I know, you’re both right”), Carol simply resorted to self-presenting her own ideas (i.e. repeating “I think I’m right, guys”, while ignoring Adam’s positive feedback). In summary, competence positioning was indicative of both frustration with Carol’s directive, gatekeeping style of other-regulation and devaluing of others’ contributions. Group 3 had an overarching negative tone with attempts to broaden access to the competent contributor role; unlike Group 2, the other-regulator’s competence messages were subtle via ignoring and exclusion. However, her less direct criticism still provoked significant resistance and repositioning efforts that included explicit references to ability, which evidenced the group’s frustration with limited access to contribute.

Overall, competence messages were present in all groups, but messages varied in salience as well as focus. Competence positioning efforts dominated the interactions in Group 2. Importantly, the competence messages served more exclusionary functions in Groups 2 and 3, whereas Group 1’s messages facilitated the integration of all members’ perspectives. Group 2’s competence messages made individual competence evaluation salient because of the constant invocation of a competence hierarchy in service of David and Carla’s own attempts to use each other’s presumed incompetence to justify exclusion from contributing to the group. Finally, negotiation of competence in Group 3 was driven by Dylan and Adam’s attempts to access the competent contributor role in the face of Carol’s condescending approach to their contributions. Groups 2 and 3 show that directive other-regulation seems to yield competence threats that promote harmful group interactions.

Individual Level: Positioning Oneself as Competent

We identified self-advocating as the most frequent form of competence positioning for all groups, with students repeating their own positions seemingly to ensure they were heard and their points might be taken up by the group (Barron, 2000). Barron (2000) found that repetitions in a less academically successful group were efforts to be heard and were exclusively self-referential, with much time spent pushing for one’s own view. A more successful group had more varied uses of repetitions, one of which was repeating a view while discussing whether it was right. Consistent with Barron’s findings, the function of self-advocating varied across groups. Members of Group 1 self-advocated clarifying claims misunderstood by groupmates. For example, Bob said “I want to argue that it could be a two and not a three [quality rating of the evidence]” to reintroduce a comment which had been interrupted. Later, he re-advocated his contribution: “I never said it was a 2” to highlight the group’s misunderstanding of his claim. Group 1 can also be differentiated by their responsiveness to self-advocating. For instance, Group 1 actively listened to Bob’s repeated claims, and made efforts to both interpret and provide feedback to his ideas. Finally, it is notable that no self-positive or self-negative statements were observed, suggesting a de-emphasized focus on proving one’s individual competence to groupmates.

In Group 2, self-advocating involved repeating claims without additional elaboration or justification. For example, Carla and David continued repeating their claims across 10 turns, with Carla advocating that Explanation F was highest in quality (e.g., “I don’t like it. F. F!”), followed by David repeating that he preferred Evidence C (“It’s C!”). Carla gave some elaboration for why she preferred Evidence F, but even when some rationale was provided, repeated claims received minimal group responsiveness (e.g. David: “No it does NOT!”). Group 2 also had 8 instances of self-positive. Self-presentation comments involved demonstrating superiority of one’s competence (e.g. “Ha, I told you”; “I bet you it’s better”). Also, Group 2’s individual positioning was used to thwart opposing and competing views that may have threatened one’s own competent contribution. Moreover, self-presentation set a competitive tone to the group’s interactions that made relative ability salient.

In Group 3, Dylan self-advocated similarly to Bob from Group 1 by clarifying his claims in attempts to have his ideas included in the group task. Similarly, Adam self-advocated 6 times, with attempts to clarify his positions (e.g. adding emphasis in saying, “It says he *received* a 10 dollar bill before closing” and “I’m talking about Sam!”) and once self-advocated for task regulation (e.g. “We have to discuss the problems”). However, Carol did not work to understand or integrate their points, as we had observed for Group 1. Instead, Carol ignored each of Adam’s attempts and questioned the legitimacy of Dylan’s claims. For instance, she diminished Dylan’s contribution saying, “I don’t think that matters, but...” Additionally, in her own attempts to self-advocate, Carol directly attacked Dylan’s claims, while boosting her own by saying, “Okay, so...if he paid 15 dollars...Hello?! Are you listening? If he paid him 15 dollars, there’s no such thing as a 15 dollar bill.” There were also 5 instances of self-positive, with four instances of Carol saying she was right and one instance of Dylan saying his own reasoning was correct in the middle of an explanation. Taken together, Dylan and Adam’s individual positioning seemed to function as a means to contest a directive other-regulator’s monopolization of task access. Carol coupled self-advocating with self-presentation in ways that seemed to aim at maintaining her position as most competent contributor. While the self-presentation was not as harsh or as direct as demonstrated in Group 2, Carol’s advocacy for being right and for her own ideas communicated more subtly who was competent in the group.

Discussion

Overall, our results indicate that other-regulation can take varying forms ranging from facilitative to more directive forms. In our results, we further described degrees of directive other-regulation, with some demonstrating explicitly controlling qualities (Group 2) and others employing more subtle methods of control (Group 3) (also see Rogat & Linnenbrink-Garcia, 2011). This finding extends previous research which has primarily characterized other-regulation as supportive of understanding (Hadwin & Oshige, 2011; Li, et al., 2007; Vauras, et al., 2003). This distinction in other-regulation highlights qualitative differences in employed social regulation. First, in terms of the nature of the regulation, facilitative regulators focused on guiding group process, content understanding, and task contributions, while directive other-regulators focused on controlling and managing the task product as well as who ultimately made contributions to the work. Second, forms of other-regulation varied in quality. Facilitators more often engaged in high quality regulation, given their emphasis on ensuring equitable participation and encouraging shared understanding, while directive other-regulation was lower quality, given a focus on excluding participation and controlling the ideas integrated in the final product. Further, the harsh criticism in one group's directive other-regulation led to even lower quality regulation compared to a group with less volatile interactions.

This observed variation in other-regulation has implications for group process. In an extension of work that considered directive regulation's detrimental influence on group's socioemotional interactions (Rogat & Linnenbrink-Garcia, 2011), we observed negative impact on another group process, competence negotiation. Other regulators played a central role in defining competence for the group and thereby shaped negotiation, as group members engaged in competence negotiation in response to the tone set by other-regulation. Facilitators' solicitation of others' perspectives, advocacy for others, and treatment of everyone's views as valuable for the overall solution helped focus the group on achieving shared understanding and democratized access to the role of competent contributor. This encouraged all group members to take up opportunities to contribute by making claims as well as raising questions. In this context, group members self-advocated in a productive way resembling academically successful groups in past research (Barron, 2000). In contrast, directive other-regulators monopolized opportunities for making competent contributions through implicit and explicit competence messages that produced a hierarchy of competence (Altermatt, et al., 2002): they ignored and dismissed views, treating conflicting points as in competition with their own. These competence moves led to discussions of competence by groupmates, as well as to hostility and putdowns in reaction to directive other-regulation. Here, it is critical to highlight that while the two directive other-regulators varied in their emphasis on explicit (Group 2) versus implicit competence messages (i.e., Group 3's dismissive talk, ignoring), both seemed to have detrimental effects. Competence norms negotiated in the group had implications for how individuals positioned themselves to contribute on the group task.

What explains the emergence of directive other-regulation and the accompanying low quality regulation? It is important to highlight that these groups were observed during initial weeks of an intervention focused on collaborative reasoning. During early weeks, individual students may have still operated under a conceptualization of academic tasks as individual work and student's motivational orientations may have been focused on competition and demonstrating ability, marked by a performance goal orientation. This has several implications including that (1) groups may have still been in the process of resolving the many motivational and emotional regulation challenges required when coordinating joint work (Järvelä & Järvenoja, 2011; Rogat, Linnenbrink-Garcia & DiDonato, 2013); and (2) groups may have represented a distinct sub-context within the classroom (Pintrich, Conley & Kempler, 2003), reflecting a second system of competence for which students need to negotiate their position (Gresalfi, et al., 2009). In particular, even with competence systems at the whole class level involving disciplinary norms focused on equitable access to competence via criteria, students may attempt to assert dominance within the small group. As our findings demonstrated, a goal of besting others within the group and maintaining one's position of dominance can be antithetical to the goals of collaboration (Levy, Kaplan, & Patrick, 2004; Rogat, et al., 2013).

In terms of practical implications, our results indicate that social regulatory processes and group processes are mutually sustaining in that it is the interplay among high quality facilitative other-regulation and support for everyone making competent task contributions that promotes student learning during small group activities. This suggests that it is critical to address high quality group interactions as well as regulatory processes comprehensively to support collaboration. Taken together, these findings indicate that individual competence can look different in varying settings, and we need to conceptualize collaborative groups as activity systems nested within the classroom (Gresalfi, et al., 2009). Future research should investigate how other-regulation is initially negotiated within the group. In addition, we need to consider the role of individual differences, such as motivational orientation and student's perceptions of group work in explaining the emergence of other-regulation. Moreover, more attention should be given to the development and change in other-regulation over time, and what contextual factors and individual differences explain group members who continue to persist in providing high quality monitoring and in negotiating competence in the face of harsh feedback and sustained efforts to exclude contributions.

References

- Altermatt, E. R., Pomerantz, E. M., Ruble, D. N., Frey, K. S., & Greulich, F. K. (2002). Predicting changes in children's self-perceptions of academic competence: A naturalistic examination of evaluative discourse among classmates. *Developmental Psychology, 38*, 903-917.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261-271.
- Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of the Learning Sciences, 9*, 403-436.
- Darnon, C., Muller, D., Schrager, S. M., Pannuzzo, N., & Butera, F. (2006). Mastery and performance goals predict epistemic and relational conflict regulation. *Journal of Educational Psychology, 98*, 766-776.
- Gresalfi, M.S. (2009). Taking up opportunities to learn: Constructing dispositions in mathematics classrooms. *Journal of the Learning Sciences, 18*, 327-369.
- Hadwin, A., & Oshige, M. (2011). Self-regulation, coregulation, and socially shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record, 113*, 240-264.
- Hogan, K., Nastasi, B., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17*, 379-432.
- Järvelä, S. & Järvenoja, H. (2011). Socially constructed self-regulated learning and motivation regulation in collaborative learning groups. *Teachers College Record, 113*, 350-374.
- Kempler, T.M. & Linnenbrink, E.A. (2004, April). *Re-examining the influence of competition structures in group contexts: Implications for social and cognitive interactions in small groups*. Paper presented at the Annual Convention of the American Educational Research Association, San Diego.
- Levy, I., Kaplan, A., & Patrick, H. (2004). Early adolescents' achievement goals, social status, and attitudes towards cooperation with peers. *Social Psychology of Education, 7*, 127-159.
- Li, Y., Anderson, R. C., Nguyen-Jahiel, K., Dong, T., Archodidou, A., Kim, I.H., Kuo, L.J, Clark, A.M., Wu, X., Jadallah, M., & Miller, B. (2007). Emergent leadership in children's discussion groups. *Cognition and Instruction, 25*, 75-111.
- Nolen, S.B. (2007). Young children's motivation to read and write: Development in social contexts. *Cognition and Instruction, 25*, 219-270.
- Pintrich, P.R., Conley, A.M. & Kempler, T.M. (2003). Current issues in achievement goal theory and research. *International Journal of Educational Research, 39*, 319-337.
- 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*, 375-415.
- Rogat, T.K., Linnenbrink-Garcia, L. & DiDonato, N. (2013). Motivation in collaborative groups. In Hmelo-Silver, C., Chinn, C., Chan, C. & O'Donnell, A. (Eds.), *International Handbook of Collaborative Learning*. (pp.250-267). New York: Taylor & Francis.
- Salonen, P., Vauras, M., & Efklides, A. (2005). Social interaction—What can it tell us about metacognition and coregulation in learning? *European Psychologist, 10*, 199-208.
- Strijbos, J. W., Kirschner, P. A., & Martens, R. L. (Eds.). (2004). *What we know about CSCL: And implementing it in higher education*. Boston, MA: Kluwer.
- Vauras, M., Iiskala, T., Kajamies, A., Kinnunen, R., & Lehtinen, E. (2003). Shared-regulation and motivation of collaborating peers: A case analysis. *Psychologia, 46*, 19-37.
- Volet, S., Vauras, M., & Salonen, P. (2009). Self- and social regulation in learning contexts: An integrative perspective. *Educational Psychologist, 44*, 215-226.

The Blogosphere as Representational Space

Richard Alterman, Bjorn Levi Gunnarsson, Computer Science, Brandeis University, USA
Email: alterman@cs.brandeis.edu, bjornlevi@cs.brandeis.edu

Abstract: During the semester, the work of the instructor and students can produce a large number of representations that support student work. A significant element of the orchestration for a course is to organize the *representational space* so that it enables efficient and effective learning within and across different activities. In the case study that is presented, the students do their homework in a blogging environment. The data will show that student work in the blogosphere created alternate and progressive representations that persisted throughout the semester and were re-usable for other learning activities. As a representational space, the blogosphere enabled students to share at all phases of learning, while maintaining autonomy and ownership of their own work.

Introduction

Over the course of the semester, the work of the instructor and students can produce a large representational space in which learning activities can develop. Within any single learning activity, providing both good and alternate representations for the students to work with has significant positive effect on both reasoning and learning. Across learning activities, each new activity depends on the representations created earlier in the semester. The gradual accumulation of representations is a significant element of the effectiveness and efficiency of student learning. Thus an important part of the orchestration for a course is the design of the *representational space* used during the semester.

The representations that students share and create can potentially mediate their collaborations. A significant design goal, thus, is to create a uniform space of representations that enables students to maintain autonomy and ownership of their own work, while supporting collaboration at all stages of student work and reducing the overhead of sharing. This paper explores the utility of student work in a blogging environment as a basis for achieving this aim within learning activities and across stages of learning.

In the case study that is presented, students do their homework throughout the semester in a blogging environment. The first stage of the course was lecture and homework. During the second stage, student work was more oriented towards a team term project, but the students continued to do homework (and also reflection) in the blogosphere. Student work in the blogosphere created representations of the central content of the course that persisted throughout the semester and were re-usable for other learning activities.

As the students did their homework, they could read each other's drafts and were free to re-edit their own posts up until the deadline. For each assignment, the collective work of the class produced a case-base, with each post, and its accrued comments and meta-content, functioning as a case. Later in the semester, because the posts of each student persisted, the students were free to mine the blogosphere content for other learning activities and different stages of learning. The focus of the study is on how well the blogosphere content functioned as a representational space throughout the span of the semester.

There are three parts to the evaluation of the functioning of the blogosphere as representational space. First we look at the students doing homework before the term project was introduced; there were five assignments during that period. The second part of the evaluation examines the reading behavior of students in the blogosphere as they transitioned from attending lecture and doing homework to the term project stage. The third part of the evaluation looks at blogosphere activity after the term project proposal was finished until the end of the semester. Both survey data and a quantitative analysis of the students' reading behavior is presented.

About the Class

The class is a course on Human Computer Interaction (HCI); there were 48 students in the class, a mix of undergraduate, graduate, and post baccalaureate students. The main goal of the course was for students to learn methods for designing a human computer interaction. Most of the lectures were on methods and techniques. Some lectures included in-class design sessions or design briefs from design projects that were independent of the class. Throughout the semester the students had weekly homework assignments that were done in the blogosphere: the homework assignments counted for 30% of the grade.

The first part of the semester was lecture and homework. The lectures were focused on design methods and techniques for HCI. Towards the middle of the semester the term project was introduced. The focus then shifted from learning and practicing methods to applying them to an ongoing project; during this period the students continued to do weekly assignments in the blogosphere. The term project was to develop a design for a human interaction with technology using the methods and techniques the class had begun learning during the first part of the course. Students worked in teams of 2-4 students. There was a minimum set of methods and

techniques that each term project was required to use. The term project counted for 40% of each student's grade.

The team term project supported student learning in the ways that are associated with project based learning in general (Krajcik, et al, 2008; Blumenfeld et al 1991). It connected course material to the everyday experiences of the students, providing motivation and context for learning, while making the HCI design methods more meaningful. The project enabled the students to develop skill and knowledge in more depth than what they attained from the less contextualized homework assignments. During the project part of the course, much of the in-class time was run like a workshop with the instructor, teaching assistants, and students interacting with one another to work on the team term projects. The design artifacts that the students were producing made the progress and understanding of each team visible to other students, and thus it was a medium of sharing work, ideas, and giving each other feedback.

The project report, due at the end of the semester, asked each team to provide a narrative of how the students developed their design, emphasizing the selection and organization of methods that were used. Another requirement was to produce numerous design rationales for their final product; these design rationales were the kinds that professionals in the field would use to defend their design decisions. The design methods, techniques, and rationales were foreshadowed by student work in the blogosphere.

Student Blogging

In a student blogging community, each student has full control over the content of her blog. The blog is composed of multiple posts written by the blog owner. Blog posts can be lengthy, and they are self-contained. The format of a post or comment is flexible and adaptable to different kinds of contributions (Du & Wagner 2005). Students can browse in the blogosphere at any time, reading and commenting upon the contributions of other students. The overhead of learning to use the technology is low (Glogoff 2005).

At a very basic level, blogging is an activity composed of writing, reading, and commenting. From a more social perspective, the students' activity can be viewed as sharing (e.g., Deng and Yuen, 2011). From a third vantage point, over the course of the semester, the contributions of the students form a 'warehouse' of content that can be "mined" throughout the semester (Williams and Jacobs 2004).

Blogging has a social orientation in that each post initiates communication with other students; it fosters a sense of community and provides a channel for interaction amongst the students (Deng and Yuen 2011). Contributions to the blogosphere simultaneously maintain relevance to the course material while "retaining the self-directed, internal focus of the owner" (Cameron and Anderson 2006; Ellison and Wu 2008; Lara and Lomicka 2008). Prior studies have shown that students perceive reading in the blogosphere as improving their understanding of the course concepts (Ellison and Wu 2008), helping them to better organize ideas and consolidate knowledge (Zeng & Harris 2005), and exposing the students to alternate viewpoints (Oravec 2002; Ferdig & Trammel 2004).

The blogging environment used in the class was a complete rebuild and revision of an earlier version of the system that had been used in several classes. Early in the semester, we engaged the students in design sessions for revising the new blogging environment. Periodically updates to the blogging environment were released. After the first few weeks, almost all of the revisions to the design of the technology were completed.

On the front page of the blogosphere, the posts were shown in reverse chronological order, with the most recent posts at the top of the page. Students could also access posts by the tag label associated with each assignment. Each post included the name of the author and the date of the post, the title of the post, the assignment tag, a count of the number of comments, and a count of the number of "thumbs-up" given to a post by other students. Clicking on the author's name changed the display to show all the posts from that author. Hovering over the title of a post showed a preview of the post; clicking on the post itself showed the post in its entirety and any comments the post has accrued. A search function allowed the students to find all the posts that contained a search term.

During the semester, the students were responsible for writing 10 posts and 20 comments. Most of the posts were skill building; a few, later in the semester, were student reflections on term project work. Some assignments were preceded by in-class exercises that gave students face-to-face group time before beginning an assignment. While doing the skill building assignments, the students applied the same methods or techniques to examples of their own choice. Students were encouraged to post drafts of their work before the deadline, thus the students could collaborate on the homework by reading and commenting on each other's drafts.

After the submission deadline for each assignment, the TA assigned to each student two posts to "officially comment on"; the official comments were due a few days after the post was due. After both the posts and official comments had been graded, the TA gave "gold stars" to the best posts for a given assignment. Thus for each blogging assignment, there were several iterations on the content of the blogging assignment: some before (reading the assigned material and listening to a lecture), some during (doing the homework and browsing while doing the homework), and some after (commenting and interpreting feedback).

The first five assignments were concerned with techniques and methods for designing a human computer interaction. The sixth post was draft of the project proposal, written by each student individually. A week before the team proposal was due, each student wrote a draft version of the proposal as a blog post; so a team of three students would have three different draft versions of their project proposal written a week before the proposal was due. The draft proposal substantially overlapped with post 5 (scenarios, conceptual models, and prototypes) and to a lesser extent post 4 (data gathering plan). From the in-class discussion it was clear the students struggled with the scenario creation part.

After the draft proposal posts, there were four other posts. Two of these posts were additional skill building activities that were directly related to the requirements of the term project. The other two posts were “progress reports” written by each student on his/her team term project.

Students were encouraged to read freely throughout the semester in the blogosphere. While working on an assignment, it was perfectly fine for a student to review the posted work of other students. It was also ok for a student to revise her/his post up until the deadline.

Features of Blogging as a Knowledge Community

For the class, the blogosphere was an open space (Duval 2011) that made it easier for students to collaborate: the students broadcast and shared their work. They had access to the draft versions of each other’s post to support their own learning. The students could “work together” even though they worked from different places and/or at different times.

Doing homework in the blogosphere was a loosely coordinated activity: the students connected and shared with one another, producing sharable objects, and common (background) knowledge in a distributed fashion while collaboratively acquiring knowledge and building skills (Alterman and Larusson, 2013).

Each contribution was self-contained. As each post was authored, there was an expectation that the reader should be able to read the post by itself. Where, for example, an utterance in a chat is designed for the *recipient* (Sacks et al, 1974) with the expectation that there will be further interaction, or contributions to a discussion forum depend on other contributions, a blog post is constructed as a self-contained communication that is broadcast to the rest of the class. At the time each post is constructed, factors that would make the contribution understandable at other times, in other places, are a significant element of how the post was developed.

To a certain extent, the reasoning and learning the students did during the semester vis-à-vis posts in the blogosphere can be thought of as a form of case-based reasoning (Kolodner et al 1996; Kolodner et al, 2003). The entire collection of posts for a single assignment forms a case base. Each post “practices” the same method or technique with a different example, thus an individual post for a homework assignment on, for example, writing a questionnaire, with the commentary it accrues, can be viewed as a case. Because the students can, and do, collaborate while the posts are being written, the cases are collaboratively produced. The meta-content for, and commentary on, each post further enriches each case. Indices for retrieving content were created both as the post is developed (“encoded and inserted”) and each time it is retrieved (“retrieval time”) (Kolodner et al, *ibid*). Indices that were added as each case/post was “encoded and inserted” into the case base included the assignment tag and the name of the author of the post. Any time a student read a post, she could create an index to it by bookmarking it or just remember it or the search term she used to find it. The iterative refinement of each post/case as it was developed and the reuse of these cases for other endeavors are characteristic of the case-based approach to reasoning and learning.

Representational Space

From a larger perspective, the instructor organizes a course into a hierarchy of topics and subtopics. A sequence of instruction can be defined by a traversal of the topic tree; earlier topics of the instructional sequence prepare the students for later learning activities (Gagné, 1973). Any traversal will have both breadth and depth components (Collins et al 1991). The breadth components give the student a map of the terrain that they will be covering over a sequence of learning activities. The depth components enable the students to acquire details of the targeted skills and knowledge and a more nuanced view of their application. A course that has a sequence that is organized as breadth-first can develop in stages. Students make multiple passes through the same material, gradually deepening their understanding and improving their skill. Each stage of learning produces representations for the next stage of learning.

In the HCI course, at the level of curriculum, the term project is an iteration of working with the methods and skills taught in the first part of the course. During the first stage of instruction, the students collaboratively worked at learning techniques, methods, and argumentation of HCI. During the project stage, the students developed skill and knowledge in more depth than what they attained from the less contextualized homework assignments.

Over the course of the semester, the work of the instructor and students can produce a large representation space in which learning activities can develop. Within any single learning activity, providing

good representations for the students to work with has significant positive effect on both reasoning and learning (e.g., Larkin & Simon, 1987; Zhang & Norman, 1994; Scaife & Rogers, 1996; Suthers, 2005). Learning from multiple external representations has great additional value (Ainsworth, 2006): alternate representations can be complementary, students differ in their preferences, different tasks work better with different representations, and multiple representations encourage strategic thinking.

Across learning activities, each new activity depends on the representations created earlier in the semester. Representations produced earlier in the semester can coordinate and mediate subsequent learning activities, which produce additional alternate representations. As new representations are added to the shared space, it gradually becomes enriched. The gradual accumulation of these sorts of representation, the enrichment of the space in which the class operates, is a significant element of the effectiveness and efficiency of student learning.

The orchestration of the course has direct bearing on the media and form of the representations. For the HCI class, the students had access to many different kinds of representation: blogosphere content, reading material, electronic and hard copies of the lecture slides, in-class handouts, each student's private notes and the "remembering" of lectures and in-class exercises.

In the study presented in this paper, we will especially be interested in how well the blogosphere content functioned as a representational space both within learning activities and across stages of learning. Using the blogosphere as a forum for collaboration has direct bearing on the potential for sharing and re-using content as the students acquire skill and knowledge. Did the production of content enable collaboration and the formation of common knowledge amongst the students? Did the students leverage the content produced in one activity for another? Did the content of the blogosphere support the transition of learning to the project based stage?

Evaluation

There are three parts to the evaluation of the functioning of the blogosphere as representational space. First we look at the student doing homework before the term project was introduced. There were five assignments during that period. Each of the assignments emphasized skills and methods for designing interfaces. A survey of student attitudes and preferences is presented as well as quantitative data of the reading behavior of the students. Each time a student clicked on a post to access the full text of the post and any comment it had received, as opposed to previewing the post by hovering, we counted it as a read. Unfortunately, reading behavior data was lost for the last 12 days of the semester. During this period the students were finishing up their projects and final reports. Despite the loss of data, the evidence shows that many students continued to be leverage the accumulated content of the blogosphere during the latter part of the semester.

The second part of the evaluation examines the reading behavior of students in the blogosphere as they transitioned from the first stage of the course to the term project stage: the students first individually wrote drafts of the term project proposal and then they wrote, as a team, a more formal and complete proposal outside the blogosphere. Again, both survey and data on reading behavior are presented. The third part of the evaluation briefly looks at blogosphere activity after the term project proposal was finished.

Doing Homework

Midway through the semester, we gave a survey questionnaire to the class on their work in the blogosphere. The questions were on a 5-point Likert scale. For each question we calculated the average response and also the percentage of students who agreed/strongly agreed (see Table 1). Most of the students (69%) believed that they were learning more by doing the homework in the blogosphere (question 1). As they did their homework, a majority of the students browsed in the blogosphere to interpret the assignment, viewing how other students approached it, and looking at how it was presented (questions 2-4).

Table 1: Blogging in general.

<i>QUESTION</i>	<i>(AVERAGE; % AGREED/ STRONGLY AGREED)</i>
I believe I am learning more by doing homework in a blogosphere than I would be by doing it alone.	(3.88; 69%)
I browse in the blogosphere to help me interpret the assignment.	(3.96; 76%)
I browse in the blogosphere to see how other students are doing the assignment or part of it.	(4.04; 80%)
I browse in the blogosphere to see how other students are presenting/formatting their posts.	(3.76; 59%)

The numbers in Table 1 are for the class as whole. We compared those numbers to those of the non-native speakers. The non-native speakers in the class were even more positive about the value of doing homework in the blogosphere than doing it alone (4.29, 86%). They also agreed/strongly agreed more often that they were using the blogosphere to help interpret assignments (4.14, 76%), see how other students did the work (4.05, 85%) and presented their answers (4.10; 76%). One non-native speaker, a Masters student, said that the blogging work was better for developing her English than the ESL course that she had been required to take.

We also did a quantitative analysis of the reading behavior of the students, dividing it into three phases: the week long period the students had for completing their post for a given homework assignment, the 3-day official period for commenting (students are assigned two homework posts to comment upon), and the period afterwards that ended on the last day of the semester.

Table 2 shows how much reading, on average, each student did of the posts of other students during three different phases (columns) in the lifetime of a post – these data is for the first five posts, i.e., the homework assignments the students did before they started their term project. For each phase we calculated the average number of posts read per student per post (reads). Because a single post might have been read more than once, we also calculated the average number of different posts read (unique) by each student for each assignment. The “reads” number tells us how many times, for each assignment, on average, each student accessed cases in the case base, and the “unique” number, how many different cases/posts were accessed.

The data is divided into quadrilles (rows), from the most active readers (1st quadrille) to the least active (4th quadrille); there were 12 students in each quadrille. The quadrilles were recomputed for each phase; thus accounting for different styles of use. *All data reported in the table excludes the numbers for students reading their own posts.* As an example of how to read the table, in the first row, each student in the first quadrille read on average 24.92 posts during the time they did their assignment, and they read 15.40 different posts.

Table 2: Reading behavior during the three phases of a post/case.

	Phase 1: While doing each HW		Phase 2: While commenting		Phase 3: Thereafter	
	Reads	Unique	Reads	Unique	Reads	Unique
1st quadrille	24.92	15.40	22.19	6.90	10.26	5.81
2nd quadrille	11.60	8.42	9.07	3.51	4.00	2.40
3rd quadrille	6.42	5.22	7.04	2.79	2.92	1.72
4th quadrille	3.08	2.61	4.19	1.60	1.18	0.71

During each phase, many of the students actively used the blogosphere content. During the time the students were composing their post, the 12 most active users of blogosphere content read almost one third of the posts by other students, and the second quadrille roughly 16% of the posts. During the commenting phase, the most active readers read 6.90 different posts, two of which were formally assigned to them to comment upon.

Kaptelinin & Cole (2002) describe a “life cycle” of intersubjectivity within the classroom. In the first phase (pre-intersubjectivity), students engage in individual activities that are externally coordinated. In the second phase (intersubjectivity), the students begin to work and play together. In the third, and last, phase (post-intersubjectivity), the group work is over but the “residue” of their shared activities impacts their individual work or other collective activities. Table 2 shows that the blogosphere enables the class to share a computer-mediated collaborative intersubjective experience (phase 2) that the class can use as basis for their term project work (phase 3). Because the students had collaborated as they worked on the first five assignments, during the workshop period of the class, they were better prepared to continue to collaborate even across teams: the students had done a lot of sharing during the first stage of the course and had created a large representation space that could be leveraged for other learning activities.

Transitioning to the Project: Writing the Draft and the Formal Proposal

Students were given two weeks to write a term project proposal. At the end of the first of the two weeks, each student posted an abbreviated version of the project proposal. Table 3 summarizes the data during the week the students were writing the draft proposal. The term project description was handed out after the comments were due for post 5, so all the data is for the additional reading students did after the due date for comments. The most active readers, on average, read the draft proposals of other students 21.08 times (11.75 unique) while writing their own drafts. There were also a large number of students that read post 5 on scenarios. The most active readers of post 4 on data gathering, sampled that case-base a little bit, on average 3.42 total reads and 2.75 different posts.

Table 3: Reading behavior while composing the draft proposal.

	Data Gathering		Scenarios		Draft	
	Reads	Unique	Reads	Unique	Reads	Unique
1st quadrille	3.42	2.75	15.33	6.75	21.08	11.75
2nd quadrille	2.17	1.33	6.58	2.67	8.92	6.75
3rd quadrille	0.17	0.50	4.08	1.25	5.08	4.00
4th quadrille	0.08	0.00	1.25	1.33	2.17	0.92

After the students turned in their draft proposals, they had one week to complete the team version of the project proposal. At the time the students wrote their formal term project proposal, there were many representations available, both inside and outside the blogosphere. We surveyed the students to see what resources (representations) they felt were useful for developing scenarios and writing the data gathering plan for the proposal. See Table 4 for a summary of results. Not surprisingly a high percentage of students agreed/strongly agreed with the utility of the draft versions of their initial scenarios (84%), with the average of 4.40 on a 5-point Likert scale; ditto for writing the data gathering plan (4.23; 81%). But other representations were also clearly valuable for students in the class.

Table 4: Value of different representational resources.

<i>RESOURCE/REPRESENTATION</i>	<i>SCENARIOS</i>	<i>DATA GATHERING PLANS</i>
Draft proposal written by student	(4.40; 84%)	(4.23; 81%)
Relevant homework post	(4.14; 84%)	(4.09; 71%)
Lecture notes	(3.65; 67%)	(3.58; 56%)
Another student's homework post	(3.55; 57%)	(3.40; 49%)
Another student's draft proposal	(3.29; 43%)	(3.44; 52%)
Required reading	(2.63; 31%)	(2.77; 31%)

Figure 1 shows the distribution of students who agree/strongly agree that one or another resource was valuable in preparing either the scenario or data gathering portions of the term project proposal. The resources are listed from top to bottom, in reverse chronological order. All of the resources were valuable to somebody. For example, even though readings and lecture notes happened before the post on scenarios (or data gathering), they continued to have value for many of the students when they were writing their the final version of the project proposal; this was despite the fact they had already compiled some of the content into their draft proposals.

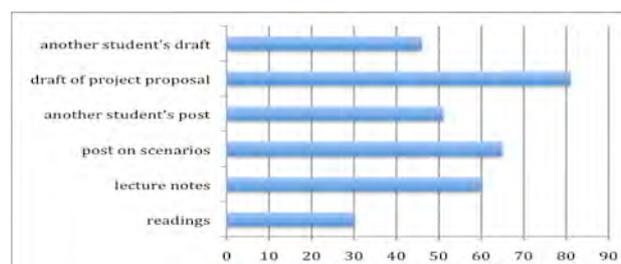


Figure 1. Distribution of students that agree/strongly agree with value of a given resource.

Table 5 shows the reading behavior data during the week the students wrote the formal proposal. There were 16 team proposals; the numbers reported are normalized for the size of each team. The first row shows the reading of the most active readers during this period. The second row shows the average amount of reading per team. The third row shows how much on average each student in the class read in the blogosphere. The draft proposal numbers include the commenting period for the draft proposal assignment (commenting on two posts was required) and teammates reading each other's posts (as opposed to, for example, emailing copies to one another). The average team size was 2.88. Subtracting these factors from the analysis of student reading of the draft proposals, the students in the 1st quadrille read 10.37 unique posts, other than the posts they were required to comment upon or the posts of their teammates. In a similar vein, the average for all teams was an additional 8.51 posts, and the average for individual students was 2.87 additional posts. It was not possible to do this analysis for the total reads. These data shows that many students continued to sample the blogosphere beyond what was required.

Table 5: Reading behavior while composing the team version of the project proposal.

	Data Gathering		Scenarios		Draft	
	Reads	Unique	Reads	Unique	Reads	Unique
1st quadrille of individual students	1.08	0.75	1.58	1.58	40.25	15.25
Average for all teams	1.44	1.06	3.25	2.13	53.06	22.56
Average individual student	0.48	0.35	1.13	0.75	18.56	7.75

The Blogosphere after the Team Proposal

There were four posts written after the due date for the formal proposal. Two of the posts were skill building, and the other two posts were progress/reflection reports; we did not do an analysis of the reflection posts.

The first skill building post was completed before the time it was relevant for the project, so for this post, the students did an analysis of an application or website other than the one they were working on for their term project. Students continued to “lean” on the blogosphere content (numbers not shown): the numbers for the 1st quadrille are commensurate with the numbers for the 1st quadrille of students for the posts that were written before the term project started; the numbers for the other quadrilles were slightly lower. The second skill building post was at the end of the semester, so the students could do the assignment by applying the techniques to data from their term project, and this is what they did. Unfortunately the data for how much reading the students did as they wrote this post was lost. However, given the difficulty of the assignment there is no reason to believe that the students did less collaborating as they wrote, and they might have done more.

Table 6 shows how much reading the students did of the first five posts after the team proposal was finished. The numbers show that a quarter of the students continued to find the representational space of the blogosphere a useful resource while the students worked on their project and then wrote their report. As to which assignment (case base) was most relevant, it depended on the student. The students undoubtedly were doing more reading than is reflected in this table for two reasons: the last 12 days of reading data were lost, and during the workshop period, teams of students, collocated in the classroom, may have been reading together.

Table 6: Reading posts 1-5 after the formal proposal was due.

	Posts 1-5	
	Reads	Unique
1st quadrille	14.33	12.33
2nd quadrille	4.25	3.25
3rd quadrille	0.42	0.33
4th quadrille	0.00	0.00

Concluding Remarks

In the study presented in this paper, the blogosphere was a representational space that supported sharing and collaboration throughout the semester, while enabling students to maintain autonomy and ownership of their own work. The data shows that there was significant use of the blogosphere as a representational space to support learning both within and across learning activities.

Many students collaborated as they did their homework. The collective work of the class collaboratively produced a case-base for each homework assignment, with each post, and its accrued comments and meta-content, functioning as a case. After the due date for a homework, the students continued to read the collected posts for that assignment, both during the official commenting phase and thereafter.

As the students transitioned to the project stage of the course, the survey data shows that many students had a preference for using the blogosphere content created earlier in the semester. The students individually wrote drafts of the project proposal as a blogging assignment, and they wrote a more formal team project proposal outside the blogosphere. In both parts of the transition, large numbers of students continued to use the blogosphere content to support their work. During the project stage, despite the loss of the last 12 days of usage data, the data shows that at least 12 of the students were very active readers of blogosphere content.

References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction* 16, 183-198.
- Alterman, R. & Larusson, J. Participation and common knowledge in a case study of student blogging. *International Journal of Computer-Supported Collaborative Learning*, 2013. (DOI) 10.1007/s11412-013-9167-2.

- Blumenfeld, P., Soloway, E., Marx, R., Krajcik, J., Guzdial, M., & Palinscar, A. (1991). Motivating Project-Based Learning: Sustaining the Doing, Supporting the Learning. *Educ. Psychologist*, 26, 369-398.
- Cameron, T. and Anderson, T. (2006). Comparing weblogs to threaded discussion tools in online educational contexts. *Int'l of instructional technology and distance learning*, 2(11):3-15.
- Collins, A., Brown, J.S. & Holum, A. (1991). Cognitive Apprenticeship: Making Thinking Visible. *American Educator*, 6-11, 38-46.
- Duvall, E. (2011). On the importance of being open. Keynote address CSCL 2011. <http://www.isls.org/cscl2011/keynote-duval.htm>.
- de Vries, E., Lund, K., & Baker, M. (2002). Computer-mediated epistemic dialogue: Explanation and argumentation as vehicles for understanding scientific notions. *Learning Sciences*, 11(1), 63-103.
- Deng, L. and Yuen, A. (2011). Towards a framework for educational affordances of blogs. *Computers & Education*, 56(2):441-451.
- Du, H. S., & Wagner, C. (2005). Learning with Weblogs: An empirical investigation. In *Proceedings of the 38th Annual Hawaii Int'l Conference on System Sciences (HICSS'05)* (pp.7b).
- Ellison, N.B. and Wu, Y. (2008). Blogging in the classroom: A preliminary exploration of student attitudes and impact on comprehension. *Journal of Educational Multimedia and Hypermedia*, 17(1), 99-122.
- Ferdig, R. E., & Trammel, K. D. (2004). Content delivery in the "Blogsphere". *T.H.E. Journal Online: Technological Horizons in Education*, 31(7), 12-16.
- Gagné, R. (1973). Learning and Instructional Sequence. *Review of Research in Education*, Vol. 1, 3-33.
- Glogoff, S. (2005). Instructional blogging: Promoting interactivity, student-centered learning, and peer input. *Innovate: Journal of Online Education*, 1(5).
- Guzdial, M. and Turns, J. (2000). Effective discussion through a computer-mediated anchored forum. *Learning Sciences*, 9(4), 437-469.
- Kaptelinin, V. and Cole, M. (2002). Individual and collective activities in educational computer game playing. In Kosmann, Hall, & Miyake, (Eds.) *CSCL 2: Carrying forward the conversation*, 303-316. LEA.
- Kolodner, J. L., Hmelo, C. E., & Narayanan, N. H. (1996). Problem-based learning meets case-based reasoning. In D. C. Edelson & E. A. Domeshek (Eds.), *Proceedings of ICLS '96*, 188-195.
- Kolodner, J., Camp, P., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S. Ryan, M. (2003). *Learning Sciences*, 12(4), 495-547.
- Krajcik, J., McNeill, K., & Reiser, B. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, 82(1), 32.
- Lara. C.D. and Lomicka, L.L. (2008). Adventures in the blogosphere: From blog readers to blog writers. *Computer Assisted Language Learning*, 21(1), 9-28.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11(1), 65-99.
- McNeill, K.L., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Research in Science Teaching*, 45(1), 53 - 78.
- Oravec, J. A. (2002). Bookmarking the world: Weblog applications in education. *Journal of Adolescent & Adult Literacy*, 45(7), 616-612.
- Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, 45(2), 185-213.
- Suthers, D. (2005). Technology affordances for intersubjective learning, and how they may be exploited. In Bromme, R., Hesse, F., & Spada, H. (editors), *Barriers and Biases in Computer-Mediated Knowledge Communication*, 295-315.
- Williams, J.B. and Jacobs, J. (2004). Exploring the use of blogs as learning spaces in the higher education sector. *Australasian Journal of Educational Technology*, 20(2), 232-247.
- Zeng, X. and Harris, S.T. Blogging in an online health information technology class. *Perspectives in Health Information Management/AHIMA*. 2, 2005.
- Zhang, J., & Norman, D. (1994). Representations in Distributed Cognitive Tasks. *Cognitive Science*, 18(1), 87-122.

Measuring 'Framing' Differences of Single-Mouse and Tangible Inputs on Patterns of Collaborative Learning

Luis A. Andrade, Joshua Danish, Indiana University,
Email: laandrad@indiana.edu, jdanish@indiana.edu

Yanín Moreno, Lenin Pérez, Universidad de San Buenaventura,
Email: ymoreno@academia.usbbog.edu.co, lhpez@academia.usbbog.edu.co

Abstract: Research on usability of tangible interfaces has predominantly focused on the *quantity* of interactions. In contrast, we argue that research on tangible and touchable interfaces must focus on the *quality* of interactions. We introduce the concept of *framing and resource activation* in studying the nature of the collaborative activity within tangible interfaces. We observed 40 five year-old children engaging in a math-series activity with either a tangible or single-mouse input, and uncovered four categories of behavioral clusters which accounted for 90% of student interactions. We found a positive correlation between exploratory talk and synchrony cluster (i.e., shared responsibility over the action) particularly on the single-mouse condition, and a negative correlation with passiveness/individual behavior (no negotiation of the actions), predominantly on the tangible condition. This suggests a tension for designers as they aim to balance the learning benefits of focused individual engagement with those that stem from collaboration.

Objectives

Traditional computer interfaces were designed with a single user in mind at a time (Grudin, 1990). However, most of us will agree that today collaboration is an important part for both working and learning. A great deal of research has shown that collaborative learning outperforms individual learning (c.f., Dillenbourg, 1999; O'Malley, 1995). Since the turn of the century, software and interface designers are moving toward facilitation of collaborative interaction in co-located groups (Stahl, Koschmann, & Suthers, 2006). For instance, studies of how people collaborate with multiple-mice inputs, touch surfaces, or tangible interfaces, have received a great deal of attention (e.g., Antle, Bevans, Tanenbaum, Seaborn, & Wang, 2011; Falcão & Price, 2009; Ha, Inkpen, Mandryk, & Whalen, 2006; Harris et al., 2009; Stanton, Neale, & Bayon, 2002). A recent idea of many software and interface designers is to embed features that increase the awareness of participants over each other activities (Dourish & Bellotti, 1992). For instance, Hornecker, Marshall, Dalton, and Rogers (2008) argue that tangible materials increase awareness by facilitating the visualization of activities of other participants.

Research on Human Computer Interaction (HCI) has shown that touchable (or tangible) interfaces can promote equal access (Hornecker et al., 2008), foment on-task talk (Harris et al., 2009), and facilitate simultaneous interaction with the digital information (Harris et al., 2009). In contrast, comparative studies of single-mouse versus multiple-mice inputs have shown that six-year-old children tend to work more collaboratively by having to share a single mouse; conversely, they tend to work in an individual, parallel way with multiple-mice inputs (Stanton et al., 2002; Stanton & Neale, 2003). Thus, there seems to be two competing hypotheses about how single- or multiple-input interfaces support collaborative learning (c.f. Harris et al., 2009). On the one hand, single-mouse allows more discussion at the interior of the group, and allows more awareness of each other's actions by forcing allocation of responsibility (Stanton & Neale, 2003); but it is prone to dominance of one individual over the others, and it has been shown that the talk content is mostly about turn-taking (Harris et al., 2009). On the other hand, touchable or tangible interfaces allow more equity to take place in the physical activity (Rick et al., 2009; Stanton et al., 2002), allow more awareness by visualizing each other's actions (Antle et al., 2011; Ha et al., 2006), and also the talk content is more task-focused (Harris et al., 2009); though, more interferences might occur (Hornecker et al., 2008).

However, the epistemological assumptions of how tangible designs might enhance collaboration must be explicitly addressed (Dillenbourg, 1999). This mixed evidence from previous research creates a paradox: Are people inherently individualistic so that we ought to design interfaces in such a way that users are forced to collaborate, or do people have a natural inclination to collaborate but traditional interfaces prevent them from doing so? Additionally, does traditional technology, designed with only a single user in mind facilitate more collaboration?

The apparent paradox is created by mistakenly identifying where the *locus* of the collaborative activity is situated. We believe that although previous literature on HCI design shows the trend has moved to support group behavior, there is no consensus on the level of the unit of analysis (Dillenbourg, 1999). This issue causes mixed and contradicted results in the literature on tangibles, tabletops, hybrid surfaces, and other shared-ware technology. In this paper we would like to address this problem by elucidating a possible answer to the question of how to study collaboration on this type of interfaces, i.e., by taking into account a dynamic unit of analysis.

Theoretical and Design Framework

We argue that theory and research on tangible and touchable interfaces has to pay attention to the level of unit of analysis. We are introducing the concept of *framing and resource activation* (Conlin, Gupta, & Hammer, 2010; Scherr & Hammer, 2009) to study the nature and quality of the collaborative activity within tangible and touchable interfaces. Framing has been defined in the anthropological literature as the structure of expectations of how people understand, moment by moment, the character of an activity; stated more simply, a sense of ‘what is going on’ (Tannen, 1993). Also, Conlin et al. (2010) introduce the idea of ontological frames as a coherent activation pattern of resources. These authors state that framing has a dynamic nature, in which the pattern of activation involves manifold resources, whether in the individual mind, across minds, or across minds and materials. Some studies of these ontological framings have pointed out that various behaviors tend to cluster together both within and across participants of an activity. Also, there seems to be only so many behavioral clusters that can account for most of the time participants interact with each other. Moreover, Conlin et al. (2010) found that one of these clusters tends to be specially correlated to a particular type of reasoning, e.g., scientific reasoning.

We intend to use this framework of framing and resource activation to first categorize the clusters of behavior in a particular activity with either tangibles or single-mouse input, and second, to correlate one of these clusters with a type of productive talk, called ‘exploratory talk’ (Mercer, 2008). Exploratory talk has been defined as “the active joint engagement of the children with one another’s ideas.” (Littleton et al., 2005, p. 5). That is, we wanted to study whether children would frame the activity to collaborate and negotiate meaning, or as an independent, parallel work.

In order to understand how this dynamic unit of analysis informs collaborative learning in tangible environments, we examined how children collaboratively worked to solve a problem of mathematical series by using either a single-mouse or tangible objects. However, we were more interested in the process of how children learned within this collaborative activity than in the product of such an activity, a method that other authors have proposed (c.f., Dillenbourg, Baker, Blaye, & O’Malley, 1996). This is one way we might better inform the design of such tangible interfaces.

Methods

Participants were 40 five-year-old kindergarten students from a public school in a capital city in South America. A between-subjects condition was employed in which we randomly assigned half of the students to a single-mouse condition and the other half to a tangible condition. We considered the mouse interface condition as a type of single-input activity, whereas the tangible condition as a type of multiple-input. For the tangible condition, children were given plastic objects (e.g. geometric figures, animal figurines, colored beads). For the mouse condition, children used virtual manipulatives available on-line at the National Library of Manipulatives (Dorward & Heal, 1999). We let students use the touchpad at the center of the laptop instead of the physical device, for it was equidistant to each participant. Many teachers use this kind of settings in which they have students work in pairs with tangible material or virtual manipulatives on a laptop that is shared by a couple of students (Rosen & Hoffman, 2009). In future studies newer technology, such as iPad’s, can be used rather than laptops. In the present study, iPad’s were not used due to limited funds.

Four tasks on mathematical series were presented to the dyads to solve together. A series task is a common mathematical activity that children solve with manipulative objects, such as geometric figures or colored objects of different sizes. In a word, a series task is an activity of finding a pattern (by shape, color, size or more than one feature) that repeats throughout a long sequence of objects. This task has also been popular with the use of virtual manipulatives, which makes it a suitable task for a comparison between a tangible-environment against a single-mouse interface. Also, such an activity is both engaging and demanding for children of this age. Further, we expected this activity would help children engage in a discussion that challenges their knowledge about mathematical relationships. The first task was to create a necklace with color beads; the second task was to create a sequence of animals; the third task was to create a sequence of colored beads with two levels ABAB or ABCABC (see figure 1); the fourth task was to complete a sequence of geometric blocks.

Students were paired up by the teacher in a way that she believed created dyads that would interact well. The pairs were balanced to include a similar number of same and mixed gender compositions. The task was conducted in a separate classroom with one dyad at a time in order to have a better setting to capture good quality video-recordings. No explicit roles were assigned or explicit instructions given on how children were to deal with turn-taking. Instructions were simple and given by the researcher (e.g., “You are to discover a pattern shown by the first elements in the sequence and then follow it”). Before the four tasks started, children were given some time to familiarize with the materials. We used this initial game as an activity to informally test children’s ability to manipulate the concrete materials or the touchpad. Although all the activities were considered for the initial steps at the qualitative analysis and interpretation, for the purpose of this paper we analyzed the third task in a fine-grain detail, in which the type of material was the most similar in both

conditions. This activity lasted about 2-3 minutes for each pair, and the four activities went on for about 15 minutes in total.



Figure 1. Two conditions of a math series task; a) tangible objects, and b) virtual manipulatives

Measures of Behavior

We developed a coding scheme to capture clusters of behavior. This coding scheme was partially based on Conlin et al. (2010) and was aligned with Collazos et al. (2007) idea of *synchrony of collaboration*. A synchronous activity can be understood as a ‘meta-cognitive’ contract in which both participants expect their messages be processed fluidly and misunderstandings resolved (Collazos & Mendoza, 2006). We delimited a span of 10 seconds as a time long enough as to include sufficient information to determine how children’s behavior converge on a certain cluster; and also short enough as to have enough number of total clusters throughout the activity. Six major categories were conceived: synchrony, interference, passiveness, individuality, distraction, and indeterminate (see Table 1). However, because of the particularities in each condition (there is no individual-parallel work on single-mouse condition) the categories of passiveness and individuality were collapsed into one single code (passive is “like” doing things on your own, without paying much attention to the group-mate). Scores on these categories represent proportion of behavior occurrence (each one 10 seconds long). Two researchers independently coded all the videos for activity three and then met to resolve differences until a 100% agreement was reached.

Measures of Talk

We also transcribed children’s talk and non-verbal language. We developed a coding scheme to categorize their talk to sort out level of engagement when negotiating a shared understanding throughout the problem-solving activity. This coding scheme was partially based on the concept of exploratory talk developed by Littleton et al. (2005). The rest of the scheme was based on a converging process of refining interpretations from several iterations of watching the videos and discussing them within the research group. Four major categories were conceived: exploratory, narrative, disputative, and other (off-task). Scores on these categories represent proportion of utterances. Each category was subdivided in various sub-codes (see Table 2). Two researchers coded half of the transcripts and a third researcher coded an overlapping 50%. Overall inter-rater agreement was 87%.

Table 1: Categories of Behavioral Clusters

Category	Description	Example
Synchronous	<ul style="list-style-type: none"> • Participants share the focus of attention and their gaze is directed towards the same point. • Their actions contribute to the result of the activity. • They implicitly share responsibility for the action. 	

<p>Passive</p>	<ul style="list-style-type: none"> • Participants may or may not share the focus of attention. • However, the authorship of the action is not shared, because the activity is not the result of the joint action of the participants. 	
<p>Individuality</p>	<ul style="list-style-type: none"> • Participants, despite having active roles, act independently, without actually articulating or negotiating their actions. 	
<p>Interference</p>	<ul style="list-style-type: none"> • Participants may or may not share the focus of attention. • However, the action of a member hinders the action of the other. 	
<p>Distraction</p>	<ul style="list-style-type: none"> • The focus of the participants does not coincide with that of the proposed activity. • Participants are distracted and look to a different place. 	
<p>Indeterminate</p>	<ul style="list-style-type: none"> • There is ambiguity about the actions of the participants. • Unable to determine with certainty the intent or purpose of the action. • Includes technical, incidental, and accidental circumstances that prevent determining with certainty the purpose of the action. • The category also serves to code actions that are not relevant to this investigation (jumping for joy, singing, dancing, etc.). 	

Table 2: Categories for utterance codes

Category	Sub-code	Example
<p>Exploratory In these types of statements the intention is to construct meaning. In a way, participants intend to make a 'shared reasoning' about the problem and its solution.</p>	Propose or suggestion	"All we need is a pink." "We have to remove this." "Put it below the other." "You are to do this." "Put it here."
	Inquiry	"What?" "How?" "Where to?" "What is this?" "Blue?" "Here?"
	Respond	Boy: "Do I do it?" Girl: "Yes, do it."
	Explain	"If there are two reds here, if you put two red ... here are two yellow, two yellow, two red, two red."
<p>Descriptive Such utterances are primarily intended to narrate the facts,</p>	Egocentric narrative	"This goes here," "here we go again," "we're almost there."
	Passive narrative	"Hey, you need to put three here!"

events or intentions of participating subjects. It is not the intention to anticipate what needs to be done.	Description	"Here's one, two, three." "All the colors are gone." "White is there already."
	Assertion	"That's it," "It is this one," "I do mine," "Here!"
Disputative These types of statements show a lack of agreement among the participants, either because they do not agree, because they do not know how to fix the activity, or because they do not like how the other is acting.	Negative	"No, not here!"
	Push back	"Let me do it!" "I do it 'cause I know how to!" "Help me!" "You do it!"
	Demand	"Why did you do that?" "Now it broke!" "You don't want to play with me." "Don't do it so hard." "Don't do that!"
Others: This category includes utterances that are not addressed explicitly or implicitly to the solution of the problem or support the interaction of participants to solve the problem.	Others	"I'm thinking I'm good at this," "I scraped myself".

Results

Finding Patterns of Behavior and Talk

All dyads completed the task correctly, and took on average about two minutes to solve it (M=2:14, SD=1:27). Participants in the virtual condition (M=2:42, SD=1:53) took longer than the tangible condition (M=1:47, SD=0:27). In general, in the tangible condition children did not talk as much (151 utterances total), and talked in an egocentric way (narrative talk). It seems they were narrating and describing what they were doing, instead of collaboratively planning future moves (see Table 3). In fact, they only included a low percentage of exploratory utterances. On a positive note, though, they were not involved in as many moments of conflict (disputative talk) as on the single-mouse condition. Nonetheless, the tangible condition seems to have elicited more independent work in the form of two parallel, individual activities (see Table 4). For instance, it is apparent from the following excerpt of the tangible condition that both children were motivated on solving the problem, but this behavior seemed more like a parallel attempt without many purposeful exchanges of negotiation.

Group1: tangibles (dyad 3)



(Child A begins to place chips on the table while B notices it)
 Child B: *I will put this one.* (Child B puts one chip while Child A keeps placing other chips) *Oh! I got this one.*
 Child B: (Child B places the remaining chips and Child A looks elsewhere). *Here, look.*
 Child A: *This one goes here.*
 Child B: *This one. Done!* (End of activity).

Table 3: Percentage of utterances by condition.

	Exploratory	Narrative	Disputative	Others	Total
<i>Tangible</i>	23 (15.23%)	84 (55.63%)	14 (9.27%)	30 (19.87%)	151 (100%)
<i>Single-mouse</i>	130 (44.52%)	69 (23.63%)	64 (21.92%)	29 (9.93%)	292 (100%)

Table 4: Percentage of behavioral clusters by condition.

	Synchronous	Passive/ Individual	Interference	Distraction	Indeterminate	Total
<i>Tangible</i>	12 (11.54%)	70 (67.31%)	8 (7.69%)	5 (4.81%)	9 (8.65%)	104 (100%)
<i>Single-mouse</i>	103 (68.21%)	23 (15.23%)	5 (3.31%)	5 (3.31%)	15 (9.93%)	151 (100%)

On the other hand, single-mouse interaction led to a greater number of utterances (292 total). Children talked about future moves, planning, proposing, asking and explaining, which we labeled as ‘exploratory talk’. Children also talked about turn-taking, but this talk was not as frequent as previous research suggests (Harris et al., 2009); in fact, only 18 over a total of 130 exploratory utterances were coded as turn-taking (7.22%). Children in this condition also talked egocentrically, which was observed about half the time compared to the tangible condition. However, children did tend to argue more and conflicts arose from time to time. Indeed, children on the single-mouse condition tended to exchange a sustained, vivid dialog to keep the activity going on. For instance, in the following excerpt, not only did these kids maintain a dialog prominent with inquiry, explanations and suggestions, but also with a small level of strain:

Group2: single-mouse (dyad 1)



Boy: *Your turn.*

(Girl manipulates the mouse)

Boy: *And there's a red... Two red ... you're doing it wrong*

(Girl continues to manipulate the mouse, while he seems focused on what she does)

Boy: *Girl, here I think two reds should follow* (points on the screen with two fingers)

(Girl silently manipulate the mouse and completes the two colors)

Boy: *No, not there ... here.*

Girl: *I know!*

Boy: *Then delete them ... those two* (scratches his head)

Girl: *But they have to be pink and have to be here.*

(Boy cautiously takes off his hand from the mouse and then points at the screen)

Boy: *Here you are missing a red, two red, look here... I'll take this, this, this, this and this* (deletes them, trying to correct some colors)

Girl: *your turn* (ceding the control of the mouse)

Correlations between Behavior and Talk

In order to understand if there were systematic differences on behavioral clusters and talk between single-mouse and tangible conditions, a one-way MANOVA with three categories of talk and four categories of behavior was conducted. With the use of Wilk's lambda criterion, the combination of these observed measures was significantly related to the condition ($F(7,12)=4.91, p=.008$), with a large magnitude of association ($\eta^2 = .741$). A series of univariate ANOVA tests with alpha = (.05/7) was performed to find significant differences on the dependent measures. We found that *exploratory talk* was significantly higher for single-mouse condition ($M=12.9, SD=8.0$) than for tangible condition ($M=2.30, SD=1.77$), $F(1,18)=16.71, p=.001$, with a large effect of association, $\eta^2 = .481$. *Passive/individual behavior* was significantly higher in the tangible condition ($M=7.0, SD=3.29$) than in the single-mouse condition, ($M=2.0, SD=2.01$), $F(1,18)=16.30, p=.001$, with a large effect of association, $\eta^2 = .475$. *Disputative talk* was marginally higher in the single-mouse condition ($M=6.20, SD=6.75$) over tangible condition ($M=1.40, SD=2.01$), $F(1,18)=4.65, p=.045$, with a medium effect of association, $\eta^2 = .205$. Also, *synchronous behavior* was marginally higher on single-mouse condition ($M=10.20, SD=10.79$) over tangible condition ($M=1.20, SD=1.23$), $F(1,18)=6.86, p=.017$, with a medium effect of association, $\eta^2 = .276$. We also found that *exploratory talk* was highly correlated with *synchronous behavior* ($r=.852, p<.001$), and with *disputative talk* ($r=.651, p=.001, R^2=.423$), and also negatively correlated with *passive/individual behavior* ($r= -.583, p=.007, R^2=.339$).

Our results provide evidence to support Stanton et al.'s (2002, 2003) hypothesis. Single-mouse interfaces for these particular tasks seem to promote more dialog, discussions and more arguments among participants. Furthermore, this kind of talk was highly correlated to a set of behaviors that imply synchrony in which both participants processed each other's messages and resolved misunderstandings fluidly. We also

noticed, as Harris et al. (2009) did, that in the tangible condition children tend to talk more about the task, but it took the form of a parallel activity instead of a collaborative effort. Finally, on the single-mouse condition children did talk more about turn taking, again noticed by Harris et al. (2009), but these conversations only account for less than 10% of the collaborative talk.

Discussion and Conclusion

This paper intends to make a contribution to the literature on evaluation of the usability of tangible interfaces and collaborative learning. Previous research on the usability of tangible interfaces has predominantly focused on the *quantity* of direct interaction (e.g., Fjeld, Schar, Signorello, & Krueger, 2002; Ha et al., 2006; Harris et al., 2009). From our point of view, by introducing the concept of exploratory talk to the study of collaboration within an object-manipulation task, we have been able to sharpen the study of the *content* of the talk this type of interaction elicits. Also to acknowledge that collaboration can occur in performance as well as talking, we measured behavioral clusters of physical actions from a framing perspective. Through this process, we have paid much closer attention to the *quality* of the interactions than previous studies have.

By using this framework of framing and resource activation, we found four categories of behavioral clusters in the math-series activity, either with tangibles or single-mouse input. These four clusters (synchrony, interference, passiveness/individuality, and distraction) accounted for more than 90% of students' interactions during the problem-solving activity. We also found a large positive correlation of exploratory talk with the synchrony cluster, particularly for the single-mouse condition. At the same time, we found that exploratory talk was negatively correlated with passiveness/individual behavior, which was predominant in the tangible condition.

What does it all mean for collaboration and learning? For collaboration this suggests that if the unit of analysis is at the level of the dyad, participants are more engaged in talk and negotiation by having to share the object of interaction. From a social-mind-unit-of-analysis perspective, it seems plausible to understand that the nature of a collective activity is directed by a shared object (literally, in this case). This shared object then shapes individual actions and other mediators within the shared activity (Engeström, 1987). On the other hand, conclusions for learning might not be as straightforward, because the narrative talk that we observed with tangible materials was a particular type of egocentric talk which is characteristic of this range of age. However, Vygotsky (1986) was one of the first psychologists to notice that egocentric talk might in fact be beneficial to the child's cognitive development. In this way, at the individual level of analysis it seems that tangible objects engage students and challenge their mathematical knowledge. Although this kind of egocentric talk is important for the child's reasoning, as Vygotsky noticed, it is not shared or negotiated synchronically with others at this developmental point yet. This suggests a tension for designers of tangible interfaces as they aim to balance the learning benefits of focused individual engagement with those that stem from collaboration. We believe that recognizing how participants tend toward a set of behavioral clusters during their interactions, and then understanding how this set of clusters correlate with a particular, desired reasoning process is an important step for evaluating the shareability of every interface.

References

- Antle, A.N., Bevans, A., Tanenbaum, J., Seaborn, K., & Wang, S. (2011). *Futura: design for collaborative learning and game play on a multi-touch digital tabletop*. Paper presented at the TEI '11 Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction, New York, NY, USA.
- Collazos, C.A., Guerrero, L.A., Pino, J.A., Renzi, S., Klobas, J., Ortega, M., . . . Bravo, C. (2007). Evaluating collaborative learning processes using system-based measurement. *Journal of Educational Technology and Society*, 10(3), 257.
- Collazos, C.A., & Mendoza, J. (2006). Cómo aprovechar el "aprendizaje colaborativo" en el aula (*How to take advantage of "cooperative learning" in the classroom*). *Educación y Educadores*, 9(2), 61-76.
- Conlin, L.D., Gupta, A., & Hammer, D. (2010). *Framing and Resource Activation: Bridging the Cognitive-Situative Divide Using a Dynamic Unit of Cognitive Analysis*. Paper presented at the CogSci, Portland, USA.
- Dillenbourg, P. (1999). *Collaborative Learning: Cognitive and Computational Approaches. Advances in Learning and Instruction Series*. New York: Elsevier Science.
- Dillenbourg, P., Baker, M.J., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In E. R. Spada, P. (Eds) (Ed.), *Learning in Humans and Machine: Towards an interdisciplinary learning science*. (pp. 189-211). Oxford: Elsevier.
- Dorward, J., & Heal, R. (1999). National library of virtual manipulatives for elementary and middle level mathematics. *Proceedings of Webnet99*, 1510-1512.

- Dourish, P., & Bellotti, V. (1992). *Awareness and coordination in shared workspaces*. Paper presented at the CSCW '92 Proceedings of the 1992 ACM conference on Computer-supported cooperative work, Toronto, ON, Canada.
- Engeström, Yrjo. (1987). *Learning by expanding. An activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsulti Oy.
- Falcão, T.P., & Price, S. (2009). *What have you done! the role of 'interference' in tangible environments for supporting collaborative learning*. Paper presented at the CSCL'09 Proceedings of the 9th international conference on Computer supported collaborative learning, London.
- Fjeld, M., Schar, S.G., Signorello, D., & Krueger, H. (2002). *Alternative tools for tangible interaction: A usability evaluation*. Paper presented at the Proceedings of the International Symposium on Mixed and Augmented Reality, ISMAR 2002, Darmstadt, Germany.
- Grudin, J. (1990). *The computer reaches out: the historical continuity of interface design*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people.
- Ha, V., Inkpen, K.M., Mandryk, R.L., & Whalen, T. (2006). *Direct intentions: the effects of input devices on collaboration around a tabletop display*. Paper presented at the TABLETOP '06 Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems, Washington, DC, USA.
- Harris, A., Rick, J., Bonnett, V., Yuill, N., Fleck, R., Marshall, P., & Rogers, Y. (2009). *Around the table: Are multiple-touch surfaces better than single-touch for children's collaborative interactions?* Paper presented at the CSCL'09 Proceedings of the 9th international conference on Computer supported collaborative learning, Rhodes, Greece.
- Hornecker, E., Marshall, P., Dalton, N.S., & Rogers, Y. (2008). *Collaboration and interference: awareness with mice or touch input*. Paper presented at the CSCW '08 Proceedings of the 2008 ACM conference on Computer supported cooperative work, New York, NY, USA.
- Littleton, K., Mercer, N., Dawes, L., Wegerif, R., Rowe, D., & Sams, C. (2005). Talking and thinking together at Key Stage 1. *Early Years: Journal of International Research and Development*, 25(2), 167-182.
- Mercer, Neil. (2008). Talk and the Development of Reasoning and Understanding. *Human Development*, 51, 10. doi: 10.1159/000113158
- O'Malley, C. (Ed.). (1995). *Computer supported collaborative learning* (Vol. 128). Berlin: Springer.
- Rick, J., Harris, A., Marshall, P., Fleck, R., Yuill, N., & Rogers, Y. (2009). *Children designing together on a multi-touch tabletop: An analysis of spatial orientation and user interactions*. Paper presented at the IDC '09 Proceedings of the 8th International Conference on Interaction Design and Children, Como, Italy.
- Rosen, Dina, & Hoffman, Jo. (2009). Integrating Concrete and Virtual Manipulatives In Early Childhood Mathematics. *YC Young Children*, 64(3), 26-33.
- Scherr, R.E., & Hammer, D. (2009). Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics. *Cognition and Instruction*, 27(2), 147-174.
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. *Cambridge handbook of the learning sciences*, 2006.
- Stanton, D., Neale, H., & Bayon, V. (2002). *Interfaces to support children's co-present collaboration: multiple mice and tangible technologies*. Paper presented at the Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community.
- Stanton, D., & Neale, H.R. (2003). The effects of multiple mice on children's talk and interaction. *Journal of Computer Assisted Learning*, 19(2), 229-238.
- Tannen, D. (1993). What's in a frame? Surface evidence for underlying expectations. *Framing in discourse*, 14, 56.
- Vygotsky, L.S. (1986). *Thought and language* (A. Kozulin, trans.). Cambridge, Massachusetts: MIT Press.

Acknowledgments

We are grateful to Diana Claros for her generous help.

Measuring Social Identity Development in Epistemic Games

Golnaz Arastoopour, David Williamson Shaffer, University of Wisconsin-Madison, 1025 West Johnson St.
Madison, WI 53706

Email: arastoopour@wisc.edu, dws@education.wisc.edu

Abstract: Learning is an inherently social process and is most effective when it is situated. Situated learning is modeled in communities of practice when new members begin their initiation into the practice and begin to develop a social identity in the context of the group. Games and simulations are one way to initiate newcomers into communities of practice. This study examines pre and post survey data and chat discourse from an engineering epistemic game to determine if students develop an engineering social identity by exhibiting forms of interdependency and depersonalization. The study concludes with how examining players' epistemic frames in CSCL environments that model real-world communities of practice, like epistemic games, can aid in the development of one of the key aspects of social identity—the process of depersonalization.

Introduction

Learning, thinking, and knowing are fundamentally social. When people make meaning of the world around them, it is a socially negotiated process because the world around is socially constructed. Thus, effective learning is situated within real-world situations and practices (Lave & Wenger, 1991) or communities of practice (Wenger, 1999). Educational researchers and practitioners are attempting to translate this concept into concrete learning environments and tools such as computer support collaborative learning (CSCL) environments. This paper examines how a CSCL learning environment can facilitate membership into a community of practice.

Theory

Decades of research in the learning sciences has shown that effective learning is situated and social (Anderson, Reder, & Simon, 1996; Bandura, 1986; Brown, Collins, & Duguid, 1989; Lave & Wenger, 1991; Vygotsky, 1978), meaning that learning takes place in the same context as which it is applied. In other words, learning is a social process where knowledge is co-constructed in some specific social environment—that is, within some community.

Analytically, such learning communities have been described as *communities of practice*—groups of people that share ways of working, thinking, and acting in the world. Such communities change over time as existing members (oldtimers) leave and new members (newcomers) begin their initiation into the practice.

Before newcomers become full members of the community, however, they need to go through a process of initiation, which Lave and Wenger (1991) refer to as *legitimate peripheral participation*. Legitimate peripheral participation is a sequence of activities through which a newcomer experiences a reduced role in the full practice—a form of simulation of the full role of a community member, which, with guidance and supervision, includes interactions that are similar in form to those of the real practice but with less intensity, pressure, and risk than the practice itself. Through this process a newcomer progressively internalizes both the practices and the identity of the community. Put another way, in legitimate peripheral participation, newcomers do not join the community simply by observing a community of practice. Instead, the three key pieces of this process are: (1) developing relationships with members, (2) understanding how members make decisions, and (3) learning the tools, language, artifacts, and routines of the community (Wenger, 1999).

Thus, a newcomer develops an affiliation with a community of practice through the development of a *social identity* within the context of the community. Tajfel (1981) and others (Abrams & Hogg, 1990; Stets & Burke, 2012; H. Tajfel & Turner, 1979) have described social identity as an individual's self-perception acquired from group membership. Jackson and Smith (1999) build on this work and posit four dimensions of social identity:

1. *Intergroup context*: the extent to which the group is different from other groups
2. *Attraction*: the positive affect one feels toward the group
3. *Interdependency*: the shared ideas toward future well-being, and
4. *Depersonalization*: thinking in terms of a group member and less as an individual.

In these terms, a key goal of legitimate peripheral participation is to facilitate these processes and thus develop a newcomer's social identity within the community of practice.

One way to help young people develop such situated social identities is through games and simulations, which can provide young people with a low risk learning environment (Gee, 2003; Shaffer, 2007) in order to increase engagement and excitement while learning domain-specific content (S. Barab & Dede, 2012; Klopfer, Education, & Squire, 2008). In the base case, educational games can help build an identity within a community of practice (S. A. Barab, Barnett, & Squire, 2012; Gee, 2003; D. Hatfield & Shaffer, 2008; Steinkuehler, 2006).

To accomplish this, some game designers have developed *alternative reality games* (Mcgonigal, 2011) or *augmented reality games* (Squire & Jan, 2012), which mix real-world scenarios and game elements. Some results suggest that such games engage and motivate young people to solve real world problems (Squire, Devane, & Durga, 2008). However, although these games are engaging and attractive to players, they typically offer only one piece of developing a situated social identity: attraction to the group. Most games are not actually situated in a realistic simulation of a real-world community of practice. In other words, players are learning domain concepts out of their real-world context. As a result, young people in these learning programs may not have an opportunity to fully develop social identities of real-world communities of practice.

Epistemic games are also CSCL environments that help young people develop positive associations with communities of practice (Bagley & Shaffer, 2011; D. L. Hatfield, 2011) But epistemic games additionally provide a space for students to experience a simulation of a professional practice. For example, in the epistemic game *Nephrotex*, students play the role of engineering interns at a medical device company. They work together to design a filtration membrane for a hemodialysis machine, in a way that is similar to how professional engineers would design, build, and test a product (Chesler, Arastoopour, D'Angelo, Bagley, & Shaffer, 2012). Previous studies have shown that after participating in *Nephrotex*, women have more positive associations with a career in engineering (Arastoopour et al., 2012). In this sense, epistemic games may provide more complete opportunities to develop social identities in that they:

1. Are simulations specific to a particular professional practice (intergroup context),
2. Generate positive affect (attraction to the group),
3. May motivate young people to imagine a future within the profession (interdependency), and
4. Are situated within professional communities of practice which may allow young people to learn the epistemology of a practice (depersonalization).

Shaffer (2004, 2006, 2007) has operationalized the learning that occurs in epistemic games (and in communities of practice more generally) in terms of an *epistemic frame*. Epistemic frame theory suggests that every practice has unique collections of skills, knowledge, identities, values, and epistemologies that construct an epistemic frame. Members in a practice rely on domain-specific skills and knowledge to make and justify decisions. They have characteristics that define their identities as members of the group, as well as a set of values they use to identify important issues and problems in the field. Developing an epistemic frame means making a network of connections between these skills, knowledge, identities, values, and epistemological elements that are characteristic of the community. Using epistemic frame theory, we can thus examine newcomers' epistemic frames to determine whether they are learning to link knowledge, skills, and values to make and justify decisions in ways that model those of oldtimers of a community.

Previous studies of epistemic games have shown that they are able to accomplish the first two aspects of social identity development: intergroup context (D. L. Hatfield, 2011) and positive affect (Arastoopour et al., 2012). In this study we consider whether such games can create the interdependency and depersonalization that facilitate initiation into a community of practice.

To accomplish this, we look at data from the epistemic game *Nephrotex* and ask:

- 1) Does *Nephrotex* facilitate the development of interdependency: that is, after playing *Nephrotex*, are students more confident in their ability to succeed, and committed to pursuing the field of engineering?
- 2) Is this process of interdependency connected to the process of depersonalization: that is, do students who make more connections to epistemological elements in their epistemic frames wind up more confident and committed to engineering than those who do not?

Methods

Participants

We collected survey data from 268 students from three different studies. Students were either in the experimental *Nephrotex* group or a control group. Students in the control group were enrolled in an introductory engineering course. Figure 1 summarizes and compares activities between *Nephrotex* and the introductory course.

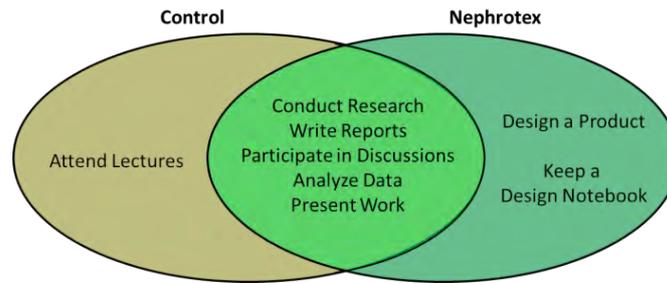


Figure 1. Venn diagram of control group and *Nephrotex* activities.

We collected pre and post survey data from all students in the experimental and control group and discourse data from students in the experimental group. Participant information is summarized in table 1.

Table 1. Participant information from three studies including experimental group and number of participants

Study	Total
University of Wisconsin-Madison, Control (Traditional course)	130
University of Pennsylvania, Experimental (<i>Nephrotex</i>)	102
University of Wisconsin-Madison, Experimental (<i>Nephrotex</i>)	36
Total	268

Pre and Post Survey

Students in the *Nephrotex* condition and in the control group answered 20 Likert-scale questions on their perceptions of engineering careers and their commitment to the field in a pre- and post-survey. We identified questions that were correlated with the survey question *How committed are you to a career in engineering?* or with *I feel confident in my ability to succeed in engineering.* at greater than .25. This resulted in a subset of eight highly correlated items: *The future benefits of studying engineering are worth the effort*, *Someone like me can succeed in an engineering career*, *I like the professionalism that goes with being an engineer*, *Engineers are innovative*, *From what I know, engineering is boring* (negatively correlated), and *I enjoy taking liberal arts courses more than math and science course* (negatively correlated).

We conducted a principal component analysis on this subset of eight items. This resulted in one significant principal component (accounted for 48% of the variance)—Commitment and Confidence.

Discourse Analysis

All chat discourse from the virtual internship was segmented by utterance. An utterance, in this case, was when a student sent a single instant message in the chat program. We coded the discourse using a set of 20 codes. The codes were developed from ABET criteria for undergraduate engineering program outcomes (ABET, 2011) and using epistemic frame theory as a guide for professional practices. Each utterance segment was coded separately (1 = present, 0 = absent) for evidence of the codes. The excerpt in figure 2 is an example of a coded utterance.

<i>Putting patient's care as a first priority, our team agrees that reliability and flux are the most important characteristics of a device. We put them as one because with higher flux, reliability increases as well. Together, higher flux and reliability decreases the amount of time a patient has to spend in dialysis care improving the patient's quality of life and treatment.</i>	Value of Client	Valuing the client/patient or stating that their needs are important
	Skill of Collaboration	The action of collaborating or participating in a team action.
	Knowledge of Attributes	Referring to attributes: reliability, flux, biocompatibility, marketability, and cost.
	Epistemology of Design	Justifying decisions using design references such as device specifications ranking/priority of attributes, or tradeoffs in design.
	Epistemology of Client	Justifying decisions by referring to the client's or patient's safety or health.

Figure 2. Example of an utterance from *Nephrotex* chat data coded for five engineering epistemic codes.

We used the method of epistemic discourse coding to code for the presence of all 20 codes. This automated coding process has been validated by comparing hand-coded utterances by multiple, independent

human coders, and by comparing hand-coded utterances to the automated coding system. Cohen's kappa scores were between .80 and .98 between the automated system and human coders. These results compare favorably to previous human-to-human coder outcomes, and, in some cases, outperform them (D'Angelo et al., 2011).

ENA measures relationships between epistemic frame elements by quantifying the co-occurrences of those elements in discourse (Orrill & Shaffer, 2012; Rupp et al., 2009; Rupp, Gustha, Mislavy, & Shaffer, 2010; Shaffer et al., 2009). We used ENA in the epistemic game, *Nephrotex*, to measure students' development of connections made between skills, knowledge, identity, values, and epistemology, and not simply quantify the isolated occurrences of these elements. For this analysis, the data was segmented into stanzas defined by all utterances from each student in each class session. We created an adjacency matrix for each stanza that identified if codes co-occurred with each other. Each of the adjacency matrices were unwrapped into adjacency vectors, normalized to a unit sphere, and a principal components analysis was conducted on the normalized data.

Results

At the end of the course, *Nephrotex* students were significantly more confident and committed to engineering than students in the control group. There was a significant difference between *Nephrotex* student ($M=.067$, $SD=1.54$) and control student ($M=-.52$, $SD=1.91$) delta scores; $t(247.9)=2.79$, $p=.005$.

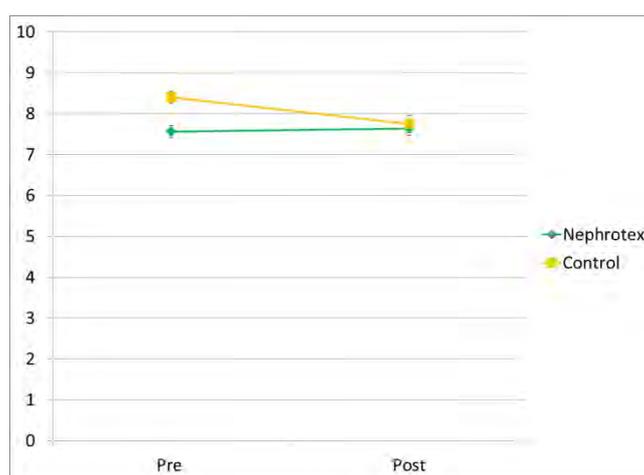


Figure 3. *Nephrotex* and control pre and post survey scores on Commitment and Confidence (Mean \pm Standard Error)

For example, one student responded to a short answer post survey question:

Before Nephrotex, I was unsure of the field of bioengineering. But I think that this internship has allowed me to understand what a bioengineer does in their day to day lives and I have realized that this is something that I like and want to do in the future.

Nephrotex students who had higher delta scores on confidence and commitment were focused mostly on epistemology of engineering during the virtual internship. There was a positive correlation between the first component of ENA (high score = focus on epistemology of data analysis and engineering design, low score = focus on other elements) and the first component of the survey (high score = high confidence and commitment, low score = low confidence and commitment), $r = .185$, $n = 135$, $p = .030$.

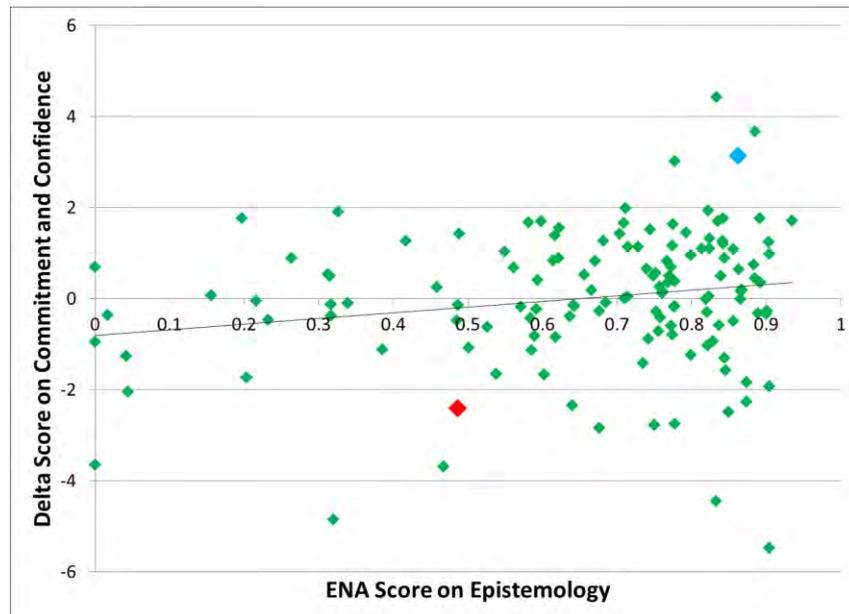


Figure 4. Scatterplot of *Nephrotex* student delta scores on *Confidence and Commitment* vs. ENA scores on *Epistemology*. Red point is a student with a low delta score and low epistemology score. Blue point is a student with a high delta score and high epistemology score.

During the second design iteration activity, interns discussed and designed final devices with their teams to send to the lab for testing. We selected two students as representative samples of students that had low commitment and confidence and low epistemology scores (Janice is represented by the red point) and of students that had high commitment and confidence and high epistemology scores (Allison is represented by the blue point). Janice (blue point) justifies her decisions with design parameters and data. She also draws on her knowledge of carbon nanotubes and attributes. On the other hand, Allison (red point) makes descriptive statements about materials, data, and manufacturing processes without justifying her statements or making epistemological statements.

Janice	<i>Yes, there are always going to be trade-offs, which is why we had to rank the attributes.</i>
High commitment and confidence High epistemology ◆	<i>but based on the conclusions from the other devices, I would recommend using 2% CNT because it lowers the cost and has the same results in terms of quality</i>
Allison	<i>The most reliable device had mediocre marketability, flux, and biocompatible values, and a high cost of \$100/unit</i>
Low commitment and confidence Low epistemology ◆	<i>The most marketable PSf product had a value of 500,000. It is also the most biocompatible of the models, with a value of 65.56. It was treated with a steric hindering surfactant and manufactured through phase inversion.</i>

Figure 5. Chat discourse samples from a design activity in *Nephrotex* from two students.

Discussion

Our results thus show that students in *Nephrotex* were more confident and committed to engineering than those in a control group. Further, students who became more confident and committed to engineering were the ones who had made more connections between epistemology and other frame elements in their epistemic frames. In the context of Jackson and Smith’s framework for social identity development, *Nephrotex* provides an opportunity to develop the final two aspects of a social identity: interdependency (in terms of commitment to the group) and depersonalization (in terms of making connections to epistemology in the epistemic frame). In particular, examining connections to epistemology in epistemic frames was an effective way to determine an individual’s depersonalization within a group.

Depersonalization, in fact, has been widely assumed to be the central cognitive process when developing a social identity (Hogg & Terry, 2012; Stets & Burke, 2012; Turner & Oakes, 2011). Change in the

perception of the self as an embodiment of the group is a critical component of entering into a new community. In turn, determining a person's level of depersonalization provides a useful measure of their social identity development relative to a particular group.

Thus, our study has two significant findings:

- (1) A key element of using a CSCL game such as *Nephrotex* to initiate affiliation with a community of practice is to offer a *realistic simulation of the epistemology of the community*; and
- (2) Epistemic frame theory can be used to *assess depersonalization*, the central piece of developing a social identity with a group.

Therefore, epistemic frames can play a central role in assessing the extent to which a CSCL environment is providing opportunities for legitimate peripheral participation and a context for situated learning.

Further studies will be needed to examine the links between students' development of attraction, interdependency, and depersonalization in a single study. This will provide further insight into social identity develop in real-world situated computer supported collaborative learning environments.

References

- ABET. (2011). *Criteria for Accrediting Engineering Programs*.
- Abrams, D. E., & Hogg, M. A. (1990). *Social identity theory: Constructive and critical advances*. Springer-Verlag Publishing.
- Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated Learning and Education. *Educational Researcher*, 25(4), 5–11.
- Arastoopour, G., Chesler, N. C., D'Angelo, C. M., Opgenorth, J. W., Reardan, C. B., Haggerty, N. P., Lepak, C. G., et al. (2012). Nephrotex: Measuring first year students' ways of professional thinking in a virtual internship. *American Society for Engineering Education (ASEE)*. San Antonio, TX.
- Bagley, E. A., & Shaffer, D. W. (2011). *Stop talking and type: Mentoring in a virtual and face-to-face environmental education environment*. Department of Learning Sciences. University of Wisconsin-Madison, Madison.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall, Inc.
- Barab, S. A., Barnett, M., & Squire, K. (2012). Developing an Empirical Account of a Community of Practice: the Essential Tensions Characterizing. *Journal of The Learning Sciences*, 11(4), 489–542.
- Barab, S., & Dede, C. (2012). Games and Immersive Participatory Simulations for Science Education: An Emerging Type of Curricula. *Journal of Science Education and Technology*, 16(1), 1–3.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–42.
- Chesler, N. C., Arastoopour, G., D'Angelo, C. M., Bagley, E. A., & Shaffer, D. W. (2012). Design of professional practice simulator for educating and motivating first-year engineering students. *Advances in Engineering Education*.
- D'Angelo, C. M., Arastoopour, G., Lepak, C., Flores-Lorca, I., Breck, E., Witherspoon-Johnson, A., Burkett, C., et al. (2011). Epistemic discourse coding. Epistemic Games Group, University of Wisconsin-Madison.
- Gee, J. P. (2003). *What Video Games Have to Teach Us About Learning and Literacy*. New York: Palgrave/Macmillan.
- Hatfield, D. L. (2011). *The right kind of telling: an analysis of feedback and learning in a journalism epistemic game*. Department of Educational Psychology. University of Wisconsin-Madison, Madison, WI.
- Hatfield, D., & Shaffer, D. W. (2008). Reflection in professional play. *ICLS*. Utrecht, Netherlands.
- Hogg, M. A., & Terry, D. J. (2012). Social Identity and Self-Categorization Processes in Organizational Contexts, 25(1), 121–140.
- Jackson, J. W., & Smith, E. R. (1999). Conceptualizing Social Identity: A New Framework and Evidence for the Impact of Different Dimensions. *Personality and Social Psychology Bulletin*, 25(1), 120–135.
- Klopfer, E., Education, T., & Squire, K. (2008). Environmental Detectives – The Development of an Augmented Reality Platform for Environmental Simulations. *Educational Technology Research and Development*, 56(2), 203–228.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, MA: Cambridge University Press.
- Mcgonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. Penguin Press HC.
- Orrill, C. H., & Shaffer, D. W. (2012). Exploring connectedness: Applying ENA to teacher knowledge. *International Conference of the Learning Sciences (ICLS)*.

- Rupp, A., Choi, Y., Gushta, M., Mislevy, R. J., Theis, M. C., Nash, P., Hatfield, D., et al. (2009). Modeling learning progressions in epistemic games with epistemic network analysis: Principles for data analysis and generation. *LeaPS* (pp. 24–26). Iowa City, IA.
- Rupp, A., Gustha, M., Mislevy, R., & Shaffer, D. W. (2010). Evidence-centered design of epistemic games: Measurement principles for complex learning environments. *Journal of Technology, Learning and Assessment*, 8(4).
- Shaffer, D. W. (2004). Epistemic frames and islands of expertise: Learning from infusion experiences. In Y. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.), *Proceedings of the Sixth International Conference of the Learning Sciences* (pp. 473–480). Santa Monica, CA: Erlbaum.
- Shaffer, D. W. (2006). Epistemic frames for epistemic games. *Computers and Education*, 46(3), 223–234.
- Shaffer, D. W. (2007). *How Computer Games Help Children Learn*. New York: Palgrave.
- Shaffer, D. W., Hatfield, D., Svarovsky, G., Nash, P., Nulty, A., Bagley, E. A., Franke, K., et al. (2009). Epistemic Network Analysis: A prototype for 21st century assessment of learning. *The International Journal of Learning and Media*, 1(1), 1–21.
- Squire, K. D., Devane, B., & Durga, S. (2008). Designing centers of expertise for academic learning through video games. *Theory into practice*, 47(3), 240–251.
- Squire, K. D., & Jan, M. (2012). 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.
- Steinkuehler, C. (2006). Massively Multiplayer Online Video Gaming as Participation in a Discourse. *Mind, Culture, and Activity*, 13(1), 38–52.
- Stets, J. E., & Burke, P. J. (2012). Identity Theory and Social Identity Theory, 63(3), 224–237.
- Tajfel, H., & Turner, J. (1979). An Integrative Theory of Intergroup Conflict. *The Social Psychology of Intergroup Relations, Monterey*, 33–47.
- Tajfel, Henri. (1981). *Human groups and social categories: Studies in social psychology*. Cambridge University Press.
- Turner, J. C., & Oakes, P. (2011). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*, 25(3), 237–252.
- Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes*. *Mind* (14th ed.). Harvard University Press.
- Wenger, E. (1999). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press.

Experiences as Resources for Sense Making: Health Education Students' I-positioning in an Online Science Philosophy Course

Maarit Arvaja, University of Jyväskylä, P. O. Box 35, 40014 University of Jyväskylä, Finland, maarit.arvaja@jyu.fi

Abstract: This study on dialogical learning explored how different philosophical approaches and reflections on these issues became a resource for health education students for understanding their work and discipline-related experiences. Especially, the focus was on exploring what "I-positions" the students constructed, and how different "voices" encountered in the students' discourse. The subjects, 11 health education students, attended an online course on the philosophy of science in the context of higher education. In order to study students' internal and external dialogue in terms of analyzing multivoicedness in their sense making process, a discourse analysis was used. Results showed that by reflecting their experiences in the light of different scientific approaches the students became aware of the different voices that underlie their thinking and activities and seemed to cause tensions in their professional positioning. This resulted in refining, strengthening and re-constructing their professional and scientific I-positions, and constructing a We-position.

Theoretical background

The dialogical perspective on learning and instruction has stressed the importance of acknowledging students' personal lives as resources for meaning-making. Building on Dialogical Self theory (Hermans, 2001), Akkerman and van Eijck (2011) talk about I-positions, which people use to express their different voiced perspectives. These I-positions vary within single individuals and are connected to the person's cultural and historical experiences and social relationships. In educational contexts, the fact that students come from various groups, communities and social networks, which are essentially reflected in their different voices and I-positions, should be therefore valued and recognized. One way to nourish students' different I-positions in educational contexts is by encouraging them to use their personal life experiences as resources and subjects for learning (Arvaja, 2012). This can be seen as promoting student identity work through giving personal meaning to learning activities. Therefore, as Ligorio (2010) argues, an educational process should help the students make sense of culture and communities, as well as themselves as active members of these communities.

According to Ligorio (2010), in education the dialogical nature of learning is often related to external dialogue between individuals, whereas internal dialogue within individuals, namely between I-positions of the self, are excluded. As a consequence most analyses and also efforts to promote learning in interaction have typically focused on the talk-in-interaction. Stressing the value of external dialogue only may lead people to assume and insist that learning situations must always be organized around some sort of conversation between individuals. This, in turn, might lead into neglecting the role of individual reasoning and internal dialogue in students' learning activities (Ligorio, 2010; Linell, 2009). However, if we consider learning from a dialogical perspective, it broadens our focus from the situated talk-in-interaction to the wider sociocultural and material aspects, and to the dialogism of the self (Akkerman & van Eijck, 2011).

From a dialogical point of view human sense making processes are profoundly interactional and contextual in their nature (Grossen & Salazar Orvig, 2011; Linell, 2009). Interaction comprises not only talk-in-interaction, but also interaction 'with the world'. Therefore, even a solitary activity, such as thinking, writing, reading, and sense making in general, is also interactional in its nature. More particularly, dialogism stresses the occurrence of several voices within individual persons as they are thinking on their own or being engaged in an external dialogue. This means that the self is based on multivoicedness manifested in expression of various and heterogeneous voices that have different social and cultural origins (Akkerman & van Eijck, 2011; Grossen & Salazar Orvig, 2011; Koschmann, 1999). The contextual aspect of human sense making highlights the fact that meaning making is anchored in both a physical and sociocultural world. Therefore, human interactions always take place in concrete situations that are mediated by artifacts and institutional rules and norms, for example (Grossen & Salazar Orvig, 2011). This stresses the importance of understanding the characteristic and "framing" of the learning situation in understanding the subject's behavior and discursive activity.

Talking about the dialogical approach to learning Akkerman and van Eijck (2011) speak for horizontal learning seeing "a learner as continuously shifting between different worlds, hybridizing and negotiating insights from different sites". They continue that learning should be seen as a dialogical practice through which students' different I-positions are stimulated. This presupposes recognizing the simultaneity of different positions that are part of students' identity and thereby part of their meaning making process. Ligorio (2010), argues that acknowledging students' personal perspectives by giving them a space to voice themselves and their perspectives in the school context can enhance students identity (de/re)construction.

Keeping the above notions in mind, in this study the health education students were encouraged to reflect on their discipline and work experiences when studying the reading material. The texts dealing with the philosophy of science provide different perspectives, “ideological voices” to reflect on one’s own multivoiced experiences. Through the reading material and discussions with one another the students come into contact with several voices (multivoicedness) that provide a context for negotiation and meaning making. Learning occurs when different voices (from different origins) are negotiated in the learning situation. Making learners contrast their own voices or perspectives with several other voices promotes reflective processes that are important for learning (Tsang, 2007). Reflecting individually and together on philosophical texts from the perspective of one’s work related experiences may also provide resources for students’ identity re-negotiations.

This study on dialogical learning explores how the health education students make sense of the philosophical knowledge through their own experiences and positions, and how they see themselves and their activities in the light of this knowledge. Furthermore, the focus is on analyzing what kinds of negotiations occur between different voices encountered in the students’ discourse. This, in turn, targets our attention to learning as identity building. It allows us to explore how the different philosophical approaches and reflections on these issues become a resource for understanding oneself and one’s own experiences, and thereby for identity construction. In sum, the aim of this study was on studying what I-positions the students took and constructed, and how different voices were encountered in the students’ internal and external dialogue during the course.

Methods

Participants and context of the study

The subjects of this study comprised 11 health education students (all female) attending an online course on the philosophy of science in the context of higher education. The students had a job, working as physiotherapists and action therapists, for example, and pursuing additional studies in health sciences, e.g. aiming at teacher qualification in health-education. They were thus studying part-time by distance learning while also working full-time. The students participated in the course from all over the country, and they met only virtually. The course utilized a web-based learning environment consisting of an asynchronous discussion tool, a tool for making text documents, and folders containing course material.

The course consisted of six learning tasks, all of which dealt with historical approaches in the philosophy of science. Each task was a reasoning task where the students were first supposed to read a given text (or texts) dealing with a particular approach within the philosophy of science. In reasoning about the task, the students were asked to use their prior experiences or conceptions about their own field of science or work as resources in interpreting the texts. Based on these tasks, each of the students was first supposed to write an *individual reasoning text*. In the next phase, the students posted their individual writings onto a shared web-based (asynchronous) discussion forum, and their task was first to read each other’s writings and finally to have a *shared discussion* based on these.

Data sources and analytical approach

One type of data consisted of students’ individual writings based on the reasoning tasks. There were altogether 66 writings. The other data consisted of shared asynchronous discussion postings. Across the six tasks, there were altogether 52 such discussion postings that commented on other’s writings or responses.

In order to study students’ internal and external dialogue in terms of analyzing multivoicedness in their sense making process, a discursive approach is necessary (Gee, 1999; Gee & Green, 1998). Discourse analysis is an analysis of social language, not an analysis of language per se (Gee & Green, 1998). According to this approach, language is seen as a socio-cultural practice and social resource of groups, and the focus of analysis is basically on what the participants accomplish through their discourse rather than on the linguistic forms or functions used as such. This allows for studying social meanings embedded in discourse as well as examining the cultural models, social and discourse practices that the participants draw on in learning and making sense of new situations. In this study, particularly the feature that the students are encouraged to discuss and reflect on their work and discipline related experiences is supposed to elicit discussion with the transpersonal dimension of the self (Grossen & Salazar Orvig, 2011). Therefore, zooming into the students’ discourse reveals how they position and see themselves through and in relation to their work, study and scientific communities, and how they display the norms, values, and beliefs attributable to these communities and/or to their Self.

In analyzing the individual writings and discussion postings around and about the learning tasks, the first thing to do was to extract from the data those work-, study-, and discipline-related experiences that were made discursively relevant (Linell, 2009) in interpreting and making sense of the phenomena under reasoning. These “tellings of personal experience” (Ochs & Capps, 1996, p. 21) were manifested in explicit discursive references to *texts, people, discourses, contexts, practices, activities, values, norms, and conceptions* related to the students’ work, studies, or discipline. These data segments, consisting of thematic meaning units or episodes

were separated upon further analysis. Then, a more intense focus was set on those data segments where students were expressing their I-positioning that depicted a personal tendency of thinking or doing (Akkerman et al., 2012). The analysis focused on what kinds of I-positions the students constructed and reflected and how these were related to other positions and voices that emerged in the course through different resources, such as the philosophical reading material, other students' contributions, and other "third parties".

Results

This section sums up the different I-positions found in the analysis of the students' discourse including both the individual writings and discussion around these writings. When it comes to interpreting the findings, attempts are made to illustrate the dialogicality of the I-positions, and the encounters of different voices (multivoicedness) in the student discourse. The section is divided into four themes that form a narrative of the course, serving also as a timeline that enables tracking possible continuities, changes and developments in the students' I-positions. For the purposes of this paper the presentation of the findings is particularly focused on exemplifying one of the students', Aino's, internal and external dialogue.

The role of science in defining one self

In the beginning of the course the most of the students define themselves as professionals whose work is strongly based on research, theories and scientific knowledge. The following example represents this position:

Aino: "Correctly and well conducted studies are of primary importance in my own work, for instance, and I always try to find an explanation for what I am doing, so that I can plead to "research-based knowledge". So there's great confidence in research."

In the example above Aino describes how research based knowledge is manifested in her everyday practices and guiding her actions and thinking in everyday work. Therefore, through her discourse it seems evident that scientific I-positioning is a strong part of her professional self. In the study, theories, frameworks and research-based knowledge can be seen as third parties that play a significant role in defining the way students think and act in their professions. Therefore, the way of seeing their activity through "science" can be seen as a relatively stabilized perspective in the professional self. This position highlights the transpersonal dimension of the self. In other words, positions within the self are not solely personal but connected to the social groups or communities (generalized other) that set rules, norms and beliefs in Me.

While the student discourse portrays a professional whose work is highly defined by scientific knowledge and theories, there are also traces of internal dialogue that questions the dominance of science in positioning oneself. The next dialogue between Aino and Niina is an example of a situation where the students are shifting between positions in their internal dialogue:

Aino: "We can think that in physiotherapy, for instance, the reliability of the therapies applied need to be proved scientifically before they can be used. This, I think, is just for the reasons of effectiveness and safety. And also for the reason that the particular field of science becomes better known and gains credibility. I myself always try to find research-based knowledge before therapy. But isn't the success of a process after all dependent on the client's/patient's subjective feelings rather than on a bunch of studies? If the patient is dissatisfied and finds that the therapy was of no help, even if it were scientifically proven, does it make any difference, then? Then again, if we reach the desired goal, does it matter what means were used?"

In her discourse Aino first defines herself as a person whose work as a physiotherapist is strongly based on scientific evidence. However, in her internal dialogue she raises another perspective for succeeding in a therapy process. She contrasts the patient's perspective ("subjective feelings") with "a bunch of studies". This shows how different stances struggle in Aino's overt talk and her inner dialogue seems to contain several 'voices' (Linell, 2009). Therefore, despite a strong research-orientation in her work, Aino questions the dominant role of science in defining the therapy process by taking up a patient's perspective, thereby accounting for the patient's personal voice. This situation illustrates how there are two points of views struggling within the same person. Later in the same writing Aino categorizes herself as a mathematical and numerical type:

Aino: "I am such a mathematical and numerical type myself that sometimes at work I have had to consciously broaden my views from staring at gauges. Do angle degrees make any difference for the end result in terms of the patient's functioning? Sure they do, but they need to be used together with broader thinking."

This self-categorization is a mark of a transpersonal self (Grossen & Salazar Orvig, 2011) through which Aino relates her activity to certain kind of science (but not the other). However, she questions her own activity “staring at gauges” in her internal dialogue. Her internal dialogue reveals that she finds this positioning as too narrow; “I have had to consciously broaden my views”. The inner confusion manifested in Aino’s internal dialogue is supported by Niina’s telling of personal experience from her own work:

Niina: “Aino, you took up an important issue about the effectiveness of therapy. Just how many intervening factors there are, considering human-related research! This week at my work again I got to find several times the fact that no matter how strictly according to the rules of science/knowledge you worked (i.e. as a therapist), human mind is always taking you by surprise. Indeed, what an impact does the psyche have on recovery! Once the “core” of a person is all right, many other things will settle and improve. Often my clients ask about so-called alternative treatments of different kinds (massage of neural pathways, zonal therapy, etc.). As you know, conventional medicine does not acknowledge these alternative treatments, so there is no so-called research evidence for these. Yet, some clients have got help from these, would this be worth noting after all?”

Niina grasps Aino’s internal dialogue and continues her discourse in the same frame. Thus, Niina identifies an internal dialogue in Aino’s utterances and verbalizes it further. It seems that the external dialogue provides a space for supporting the contradiction in an inner dialogue. Also in Niina’s internal dialogue there are traces of conflicting I-positions. She sees a contradiction between “the rules of science” and the unpredictability of the “human mind”. Furthermore, as in Aino’s case, also in Niina’s internal dialogue there is a struggle between her own faith in science and the client’s different perspective. Therefore, Aino and Niina are both voicing a patient’s/client’s perspective. Even though they position science as a strong part of Me, they also define themselves as practitioners whose work with patients is defined by their interpersonal encounters with them, and cannot be solely defined by the rules of science. In their personal tellings they define these encounters as unpredictable, subjective and situated and hence opposite to what they see science to be. In this way the students question the authoritative voice (the rules of science) that dominates and defines their activity as physiotherapists. Furthermore, we can see that their I-position shifts from strong faith in science to questioning the meaning of science. This demonstrates well their inner struggle when facing fundamental questions relative to their professional practice. According to Markova (2006), internal dialogue usually involves personal issues that require reflection and evaluation of one’s own and other people’s behavior, both past and present. It can be seen as an attempt to solve a person’s inner conflict between different ‘voices’. This can also be seen as a struggle between the collective and personal positions (Hermans, 2001), where the collective position represents the norms adopted by the generalized other, while the personal position represents true subjective feelings that may contradict with the collective position. Therefore, there is a tension between the students’ social position, outlined by societal definitions and expectations, and the students’ personal position.

A frequently used device in the students’ discourse was categorization. In the next example quite opposite to the “reliance on science” position that most of the students took in their discourse is a position where the emphasis of science is questioned. In her comment Nea indirectly criticizes others’ (physiotherapists) strong need to define their own field as science-based. In doing this she uses we-categorization including herself to the same collective (collective self, cf. Hermans, 2001), as if to soften the criticism toward the others:

Nea: “I wonder if we physiotherapists have a particularly strong need to prove our field as more scientific, so that people just wouldn’t include it to the so-called scam treatments? I haven’t heard very many physicians, nurses or, say, social workers to ponder whether this and this treatment was scientifically investigated and effective. As far as I understand, not many of the present generally used treatments/therapies/interventions have been studied exhaustively, not more than any other – if I recall it right, of medicines only penicillin has a ‘clean record’.”

In her criticism Nea emphasizes the categorization by distinguishing physiotherapists from physicians, nurses and social workers, who according to her observation do not speculate whether the treatment is based on scientific research. She also brings forward the idea that many of the things bearing influence in their work have not been (thoroughly) investigated. Aino answers Nea’s criticism by acknowledging the role of education in defining their (physiotherapists) thinking and acting as “evidence-based”:

Aino: “In our training, at least, evidence-based was emphasized so heavily that it may have become an obsession to physiotherapists. I guess we are such type of people that one must always be a little better, more efficient and more careful.”

Like Nea, Aino uses we-categorization in her positioning. In this statement Aino explicitly highlights the role of institutional education in the construction of the self. Through this transpersonal dimension Aino associates ‘physiotherapists’ with their educational community and its values and norms in defining their professional self. Therefore, leaning on evidence-based research stressed in education is regarded as a culturally correct activity for a physiotherapist and as such as a norm of generalized other. Evidence-based research can be seen as a part of a professional genre, which consists of activities with a history in their educational and work community. The term evidence-based, which is constantly used in many students’ discourse, is therefore part of a social language or professional jargon of the students’ communities. Aino also categorizes physiotherapists as a certain kind of “human type” with certain qualities. In this way she integrates the professional self with the personal self. In a way Aino defends her ideas presented before by relating them to the generalized other, which has become part of Me (a physiotherapist) or We (physiotherapists). Through these categorizations the students build their professional identity; on the one hand the categorizations define what they are, but also what they are not. Therefore, these categories have a social origin and are based on activities and practices in the students’ educational and work communities. According to Grossen and Salazar Orvig (2011), (institutions and institutional) categories are kept alive through discourse and interaction and contribute to the construct of a person’s identity and as a result lead the person to define him/herself through these categories.

The examples show how students make sense of the meaning of science in their work by referring to their experiences and past activities and practices. In their discourse the students construct different positions to science. Furthermore, there are struggles, integration and shifts between different voices and positions in the student discourse. The examples illustrate how the dialogical self is polyvocal, taking up different positions in internal and external dialogue, and may therefore appear quite fragmented at times. In particular, the examples highlight the role of institutions in the construction of the self. The examples demonstrate how science, education and work practices are all third parties or generalized others that define the students’ beings and doings as professionals. On the other hand, they are also challenged, which is manifested in internal dialogue between different voices and perspectives. The discursive space between the students seem to arise mostly from their shared or similar experiences and perceptions and acknowledging the same position.

The boundaries of (natural) sciences too narrow

Aino: ”In this week I’ve been reading texts from web and on paper, and frankly speaking I feel that my head is somewhat overloaded. One doesn’t really know anymore what to think of what issue, and now one is questioning one’s own work and science and research and whatever it was and I cannot make any sense of this, there are simply too many ideas. So I decided to look once more at this ‘what is science’ issue, on the basis of Niiniluoto’s article, because it bears most relevance to me personally. I have always considered myself a type very much oriented to science and especially to natural science, and being somehow schematic and mathematical. For this reason it feels somehow overwhelming to question everything now. Admittedly at the same time really interesting, too. What’s hard for me is that one can keep elaborating the idea endlessly and never reach a solution.”

In the beginning of her text Aino explicitly describes an inner tension and confusion she is facing when reading the philosophical course material. This results in “questioning one’s own work, science and research”. Again she categorizes herself as “a type oriented to natural science”, and it seems that from that position she is facing a challenge when being introduced with fundamental questions of the essence of science. It seems that her orientation to natural science (according to her constant self-categorization) represents a dominant voice adopted from authorities in her working and study environment. According to Linell (2009), an authoritarian voice is often like cultural assumptions that the individual does not question and once the ideas of this voice are internalized it often becomes a kind of self-discipline. However, it seems that when Aino is introduced with texts that offer different perspectives or alternative conceptions her beliefs pertaining to her “natural science oriented” position becomes questioned. Later in her internal dialogue this tension is externalized and she juxtaposes oriental and western medicine and questions the dominant voice of western medicine:

Aino: ” Oriental medicine treats a person as a whole and takes the psyche more intensely into account than western medicine. The effect of the psyche cannot be measured or questioned, still it is a fact. How much is getting well based on what our society considers appropriate treatment? In western countries we are used to think in a different way compared to the Orient. Along with globalization, sciences and pseudosciences are mixing and changing, however. In my opinion we should remain open to take into consideration also matters that we regard as pseudoscience, and look at them with equal interest to what we are able to show toward “our own science”.”

It seems that when Aino is introduced with ideas (voices) that do not support her current beliefs (an authoritative voice adopted) these beliefs become questioned and negotiated in her internal dialogue. It seems that in her particular work, science and study communities Aino have been involved mostly with research pertinent to natural science, and mostly from the perspective of physical activity. The research dealing with the psyche, therefore, seems to be in a minor position or in a position of the pseudoscience in her scientific positioning. The inner confusion that Aino faces is, therefore, a result of questioning the voice of a generalized other or authoritative voice she has adopted. She sees this dominant perspective as too narrow from her personal position or from the position of another perspective (e.g. oriental science). The examples show that when confronted with different perspectives and ways of thinking, the students also reflect and become aware of the underlying voices in their (collective) thinking and practices.

Constructing holistic We-position through heterodialogue

At the end of the course the students' are introduced with two philosophical texts representing two different points of view as regards human and natural sciences. In his text Varto emphasizes the radical difference in the basis and methodology between natural and human sciences, whereas Raatikainen finds a lot of similarities in their scientific approaches and research practices. On the one hand, the texts seem to help the students to conceptualize and construct their positions through providing definitions and qualities of different scientific approaches. On the other hand the students are engaged in *heterodialogue* (Linell, 2009) with the ideas stated in the texts, which serve as resources for strengthening the students' position either by supporting (Raatikainen) their ideas or by contradicting (Varto) some notions presented in the reading material. Through Varto's and Raatikainen's texts, the students construct a 'holistic' We-position, where they acknowledge the equal significance of human and natural sciences, qualitative and quantitative research as well as strict and exact sciences in their professions. Next example represents this position:

Anita: "You're discussing [Satu and Tia] the utilization of the perspectives of human and natural sciences at your own work on the basis of Raatikainen's article. Good reflections from both of you. In my own field as an action therapist I see this issue very largely in the same way. At work one just often comes across with what for example the referring body (doctors) value and appreciate (research consistent with natural science; unfortunately)."

Aino: "Yes, indeed, this is precisely the way I see it in physiotherapy and for my own work. The problem just lies specifically in that, for example, at work people have too high regard for the views of natural science. One has to measure mobility and muscular strength etc. and compare the results and assess effectiveness in that way. [...] Another issue I face at work is compilation of statistics. If I spend time at the ward discussing with a patient, talking about goals and motivation, listening to the person and evaluating her emotional state, without performing actual physiotherapy i.e. muscular or mobility or functional exercises, so can I then record the visit as a physiotherapeutic visit? As I didn't actually perform any therapy, but as much time was spent and after the discussion the patient is likely to be more motivated to engage in rehabilitation and more cooperative when we start actual training. [...] There's only the problem that I can't really mark on my daily nursing record sheet just that 'discussed about therapy'. Then one will cheat and take say a stretching or a quick inspection in the end. That's how it is; the emphasis is too much on natural science. :)"

The above discussion thread can be seen as an example of internal sharing of an I-position (Hermans, 2001). The students recognize in themselves and in others the same position as regards different scientific approaches and their role in their professions. Through a dialogue they construct then a We-position. Even though Anita and Aino take a similar 'holistic' position as Satu and Tia, they also acknowledge that this position is not supported in their work communities, and in the dominant practices, norms and values of the communities. In Anita's and Aino's discourse it is evident that they experience a contradiction between their own values and the values ("what the doctors value") and practices ("compilation of statistics") of the work community. Thus, Anita and Aino share the view about the dominance of the natural sciences and describe work practices that reflect and support this dominance. Therefore, their positions as an action therapist and physiotherapist are in conflict with the prevailing practices in their workplace. Thus, while their discourse is based on agreement and shared viewpoints in general, they also share the same critical position toward their work practices. The students' external dialogue reveals a struggle between their professional/personal voice and the authoritative voice in their work community. In other words, the students' are engaged in 'collective criticism' against the authoritative voice they recognize in their work and related science practices. Here, others (doctors) come to function as defining positions in what is not Me and practices (compilation of statistic) as defining practices with which I do not agree (Akkerman & Meijer, 2011). Hence, through these conceptions that the students voice one can interpret a strong support for the view that the technical-rational model (Tsang, 2007)

is inappropriate when it comes to their professional practice where they need to take into account not only the body but also the mind; feelings, motives, and values in inter-personal interactions.

The (dis)continuity of the I-positions

An interesting finding concerning Aino's discourse is that while throughout her discourse she categorizes herself as a natural science type - as a person who has adopted the practices, norms and beliefs of natural sciences dominant in her work and study communities, yet, at the end of the course she ends up criticizing those practices. However, as we can see Aino is engaged in an internal dialogue at various points of the course, where different voices from different perspectives meet. Therefore, by engaging in reflective talk in which one can reflect on one's own beliefs, values and practices, these become explicit and re-negotiable. A change in Aino's I-position is also explicitly stated in her last writing:

Aino: "At least for me this course has taught a quite different way of thinking for doing research and broadened my approach to science in general. It seems that I started from Sharply positivistic notions and ended up in a fairly broad and open view on the importance of qualitative research and human sciences, for example. It's good to stop and reflect on things and their meanings every now and then. At work one is often measuring just for the fun of it and it bears no significance, after all, to the patient let alone for science. Actually it may have been the most important lesson for me in this course; to consider what really significant science is. It is by no means about angle degrees and gauges but consideration of causal relationships more broadly and consideration of humans and interaction. Although it sometimes feels that thinking was really tangled, in the end one must say that this has been a good process."

Aino's statement clearly demonstrates how the things discussed, read and written in the course had an impact on "what I am". Aino implicitly states that what I think now is different of what I thought before, therefore indicating re-negotiation of an I-position. As previous examples demonstrate, Aino constantly categorized herself as being "a natural science type". However, exposure to the diversity through the different approaches and perspectives challenges and changes her scientific I-position. It seems that in this process she loses "the sense of dogma" that characterizes her earlier conception of science, and she now sees science as negotiated and changeable (cf. Ligorio, 2010).

In the course all the students refined their scientific I-position by conceptualizing their experiences in the light of the philosophical knowledge, and through considering and becoming aware of the various voices underlying their thinking. Therefore, providing a forum for 'identity talk' (Cohen, 2010), the writings and discussions and reflections around the learning materials gave an opportunity to *strengthen, refine and reconstruct* one's own I-positions. In other words, they provided the students with a forum for the recognition and construction of their disciplinary and professional identity.

At the end the criticism toward the dominance of science in general in the first task is refined into criticism toward the dominant position of quantitative research and natural sciences as well as toward the emphasis on physical aspects in their work and discipline. Hence, now the internal conflicts that the students highlighted in their earlier discourse are explicitly stated and conceptualized, and their origin is understood. The students recognize the tension between different value systems that is reflected in their professional practice.

Discussion and conclusions

Supporting the students' agency by acknowledging their own lives and experiences as resources for learning can lead to a learning situation where the (institutional) knowledge provided by the teacher becomes a mediating tool for understanding those lives rather than being an end in itself. The analysis revealed that the philosophical knowledge seemed to serve as a resource *through* which the students learned also about their own 'lives' and not only *about* the philosophical knowledge content. Prompting the students to use their professional experiences as a resource in interpreting the different approaches described in the philosophical texts led them into a discourse where they were able to reflect on and analyze their work and discipline related perceptions. By engaging themselves in reflective talk (Cohen, 2010), through which they could reason and share beliefs, values, and practices associated with their professional identity, the students made visible to one another who they are and what they are doing (Gee, 1999). Making one's understanding explicit and reviewing the current practices in the light of new knowledge led to new awareness and, for some students further to refined and wider perspectives and new identity constructions.

Throughout the course it seemed that the contrastive pattern the students used in discussing human and natural sciences as well as qualitative and quantitative research helped them understand the philosophy of science and its different approaches. It also served as a resource for making explicit and critically evaluate the contradictions they perceived with regard to work and science, and also facilitated conceptualization of their

professional and study-related experiences. By analyzing their experiences in the light of different scientific approaches the students became aware of the different voices and value systems that underlie their thinking and activities and seemed to cause tensions in their professional positioning. This provided possibilities to strengthen, refine and re-construct their own professional I-position. Therefore, the study showed how material resources such as texts can be meaningful in introducing the students with different voices, and how these can be a powerful resource for their meaning making and identity negotiation processes.

Altogether, it seemed that the reasoning tasks were able to raise internal dialogue, that is, negotiation between different I-positions of the self or heterodialogue with the texts (Linell, 2009). This was manifested in tensions and contradictions in the students' overt discourse. When the students were discussing their experiences, the nature of their external dialogue was mostly harmonious and characterized by sharing, acknowledging and taking the same position. However, even though the students were not engaged in argumentative talk, which is often regarded as necessary for collaborative learning, it can be argued that their external dialogue had an important role in their discussion and learning. According to Tsang (2007), while internal dialogue has a "self-generative capacity" leading to new possibilities for thought and action, external dialogue can be seen as essential for the validation of one's ideas. For example, in this study the views of others seemed to shed light on some inner tensions and add weight to a particular option in settling the inner struggles. Through sharing and constructing a We-position the students strengthened their professional identity. The way of defining one's own identity in the discussions was based not only on identifying with similar positions or shared views, but also on differentiating oneself from those who "are part of what I am not" (Akkerman & Meijer, 2011, p. 315). This kind of contrastive discourse was targeted against different "others". In their shared discourse the students' were questioning, criticizing and challenging different third parties, voices in the texts, and generalized others, even though their external dialogue remained harmonious. It can be said that the students shared a similar professional identity through recognizing the similar values, beliefs and practices present in each other's discourse.

It can be argued that the online environment provided an ideal context for the students' reflective discourse that combined institutional and personal knowledge. In the online environment the students' writings and related discussions, were available for elaboration and reflection at any time. Linell (2009) has argued that some communication types favor reflective processes more than others. Reflection presupposes that one can take an observer's role in the flow of discussion. One advantage of an asynchronous discussion forum was that it supported the occurrence of such reflective processes. It gave time to observe and reflect on one's own and others ideas, as these were readily present and available for the whole process of learning. Therefore, the online environment used in the course provided a supportive context for reflection and identity discourse.

References

- Akkerman, S., Admiraal, W. & Simons, R.J. (2012). Unity and diversity in a collaborative research project. *Culture & Psychology, 18*(2), 227-252.
- Akkerman, S. & Meijer, P. (2011). A dialogical approach to conceptualizing teacher identity. *Teaching and Teacher Education, 27*, 308-319.
- Akkerman, S. & Van Eijck, M. (in press). Re-theorising the student dialogically across and between boundaries of multiple communities. *British Educational Research Journal*.
- Arvaja, M. (2012). Personal and shared experiences as resources for meaning making in a philosophy of science course. *International Journal of Computer-Supported Collaborative Learning, 7*(1), 85-108.
- Cohen, J. (2010). Getting recognized: teachers negotiating professional identities as learners through talk. *Teaching and Teacher Education, 26*, 473-481.
- Gee, J. (1999). *An introduction to discourse analysis: Theory and method*. London: Routledge.
- Gee, J., & Green, J. (1998). Discourse analysis, learning and social practice: A methodological study. *Review of Research in Education, 23*, 119-169.
- Grossen, M. & Salazar Orvig, A. (2011). Dialogism and dialogicality in the study of self. *Culture & Psychology, 17*(4), 491-509.
- Hermans, H. (2001). The Dialogical Self: Toward a theory of personal and cultural positioning. *Culture & Psychology, 7*(3), 243-281.
- Koschmann, T. (1999). Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration. In C. Hoadley & J. Roschelle (Eds.) *Proceedings of the CSCL 1999 Conference* (pp. 308-313). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ligorio, B. (2010). Dialogical relationship between identity and learning. *Culture & Psychology, 16*(1), 93-107.
- Linell, P. (2009). *Rethinking language, mind and world dialogically: Interactional and contextual theories of human sense making*. Charlotte, NC: Information age publishing.
- Marková, I. (2006). On the "inner alter" in dialogue. *International Journal of Dialogical Science, 1*(1), 125-147.
- Ochs, E. & Capps, L. (1996). Narrating the self. *Annual Review of Anthropology, 25*, 19-43.
- Tsang, N. M. (2007). Reflection as dialogue. *British Journal of Social Work, 37*, 681-694.

Supporting school group visits to fine arts museums in the 21st century: A CSCL concept for a multi-touch table based video tool

Moritz Borchers, Knowledge Media Research Center, Schleichstrasse 6, 72076 Tuebingen, Germany,
m.borchers@iwm-kmrc.de

Philipp Mock, Department of Computer Science, University of Tuebingen, Sand 14,
72076 Tuebingen, Germany, philipp.mock@uni-tuebingen.de

Carmen Zahn, University of Applied Sciences and Arts – Northwestern Switzerland, Riggenbachstrasse 16,
4600 Olten, Switzerland, carmen.zahn@fhnw.ch

Jörg Edelmann, Knowledge Media Research Center, Tuebingen, j.edelmann@iwm-kmrc.de
Friedrich W. Hesse, Knowledge Media Research Center, Tuebingen, f.hesse@iwm-kmrc.de

Abstract: School visits to art museums are a vital aspect of art education and cultural participation. In this paper we present a CSCL concept designed to encourage students to observe closely and reflect on art works in art museums. The concept is based on a multilevel perspective on learning in art museums. It is implemented by means of a multi-touch tabletop and a video tool with functions enabling students to collaborate on processing digital reproductions of art works in small groups during a collaborative visual design task.

Introduction

Visits to arts museums and exhibitions are a vital aspect of arts education and cultural participation (for indicators of cultural participation, see Morrone, 2006). School visits to arts museums provide access to arts and high culture for a broad range of children and youth – including those coming from family backgrounds where parents are not able to or not interested in visiting fine arts museums. Hence, school museum visits are considered particularly suitable occasions for framing participation in public spaces (Nespor, 2000).

School visits are supported by museums in various ways – be it free entrance for school groups, special guided tours, or high quality educational services for K–12 educators on museum webpages. For example, The Metropolitan Museum of Art (NY) offers among other services detailed topic-specific lesson plans that relate explicitly to the National Visual Arts Standards and comprise suggestions for advanced learning activities to explore and reflect on art objects in the museum's collections (1).

Targeting the intersection of museum and school education, we propose in this paper a CSCL concept to support analysis and interpretation of art works during school visits to fine arts museums. Specifically, we investigate a concept that fosters students' close observation and reflection of art works, while sustaining the experiential flavor of the museum visit (cf. Duke, 2010). It is implemented by means of a multi-touch tabletop (MTT) and a video tool for collaborative visual design tasks in small groups. The concept has been developed through close interdisciplinary cooperation of psychologists, computer scientists and museum educators. Our work is based on our previous CSCL research which shows that advanced digital technologies offer specific opportunities for fostering knowledge construction and learning during *collaborative design activities* (e.g., Zahn, Pea, Hesse, & Rosen, 2010) and for *knowledge communication in museums* (Knipfer, Mayr, Zahn, Schwan, & Hesse, 2009). It is also firmly rooted in related CSCL perspectives, assuming multiple levels of learning ranging from individuals to communities, and assuming small group interaction to be the primary unit that mediates between individual learning and community learning (Engeström, 1999; G. Stahl, Koschmann, & Suthers, 2006; G. Stahl, 2006). In the following sections we will describe the theoretical and technical underpinnings of our concept. We will first review curricular and museum educational requirements. Subsequently, we will elaborate on how the collaborative visual design approach backed by MTT hardware can meet these requirements. The tool's implementation and functionality will be described thereafter.

Theoretical Framework

K–12 visual art education and museum educational goals

Among the most important skills which should and can be learned in visual art education, are the analysis and interpretation of visual material (National Art Education Association, 1994; Winner, 2007). In this respect, Winner (2007) differentiates between *observation* and *reflection*. Observation encompasses attending to aspects and details of works of art, especially when they are not obvious. Reflection refers to activities such as questioning, explaining and evaluating one's own works and those of others (Winner, 2007). For example, students may "explain *what* some part of their drawing depicted, *how* they had achieved a certain effect, *why* they had made something the way they did, and *what* changes they were planning in their work" (Winner, 2007, p. 28; italics in original). In addition to these academic learning objectives, museum educators stress the point that the museum visit provides an opportunity to have unique (aesthetic) experiences (cf. Pekarik, Doering, &

Karns, 1999) and to learn from those experiences beyond the “right answer paradigm” (Duke, 2010, p. 271). In this way, Barrett (2008) emphasizes the importance of multi-perspectivity in the interpretation of works of art. Such a constructivist approach to museum education is in line with the current trend toward shifting interpretative power from authorities like curators and art historians to the common visitors (cf. Barrett, 2008; Hooper-Greenhill, 2000; Trant, 2006). In the specific case of school group visits, the goals of academic visual arts education and museum education coincide in a collaborative learning situation. The specific challenge is to support such collaborative learning in fine arts exhibitions showing inestimable and untouchable (!) original art works.

We argue here from a CSCL perspective that advanced media applications in museums are suitable for supporting close observation and reflection, as well as active and collaborative learning (cf. Knipfer et al., 2009). More specifically, we argue that learning through *collaborative visual design* with digital video tools is a promising approach for meeting those requirements in a fine arts museum. In the following sections we will explain why.

Supporting observation and reflection in the art museum – a CSCL concept

Building on CSCL theory (e.g., Engeström, 1999; G. Stahl et al., 2006; G. Stahl, 2006), we develop our concept for school class visits in a fine arts museum according to the following framework:

We assume a multilevel structure of observation and reflection in a museum environment distinguishing five levels: (i) a cognitive level, where observation and reflection of art works is performed by the individual student in front of an original art work in “silent dialogue”, guided by individual knowledge and information processing abilities (cf. Leder, Belke, Oeberst, & Augustin, 2004). (ii) a socio-cognitive level, where observation and reflection is a collaborative activity of two or more students engaging in what Pea (2006, p. 1332) termed the “look-notice-comment cycle (LNC)” – the iterative sequence of observing, directing another person’s attention and commenting (Pea, 2006). In these cycles, one person’s observation becomes the starting point for a discussion, which in turn leads to further discoveries and interpretations by other group members and so forth. (iii) a socio-constructivist or: small group level, where the whole process of the students’ joint observation and reflection is held and supported by a task structure, which transforms those processes into a lasting artifact, mediated by technological affordances (cf. Suthers & Hundhausen, 2003). (iv) a class level, where observation and reflection and their manifestation in students’ group products (artifacts) form a collective museum experience that is guided by curricular goals and standards (cf. National Art Education Association, 1994) and moderated by the teacher. (v) a socio-cultural or: community level, where observation and reflection consist of a museum’s contextualized activities for knowledge building (Scardamalia, 2002), e.g., collecting, storing and presenting art works as cultural heritage, and providing for constant dialogues between artists and viewers/visitors mediated by expert curators.

We assume the five levels in our art educational concept for school class visits in a fine arts museum as being connected by *small group student interaction*. In accordance with related CSCL research, we thus put small groups at the core of the concept: Small group interaction is the primary unit that mediates between the multiple levels ranging from individual learning to community learning (G. Stahl et al., 2006; G. Stahl, 2006). Specifically, we suggest that learning through *collaborative visual design* (Zahn, Krauskopf, Hesse, & Pea, 2010) is a task that connects individual cognition to participation in knowledge building (Zahn, Krauskopf, Hesse, & Pea, 2009). By accomplishing design tasks together, students are cognitively active, they collaborate and they produce new knowledge products for the community – thereby deepening their own knowledge and adding to the community knowledge. Moreover, collaborative visual design is in line with curricular and museum educational needs: It is suitable for fostering observation and reflection skills and – as a constructivist approach – it is capable of sustaining the experiential flavor of the museum visit.

In order to allow students to learn through collaborative visual design in the art museum of the 21st century, it is crucial to provide for hardware and software solutions, which are not only appropriate for the task (cf. Pea, 2006; Zahn, Pea, et al., 2010) but also meet the special requirements of a museum environment (cf. Hinrichs, Schmidt, & Carpendale, 2008; Jacucci et al., 2010). In search of appropriate hardware that can support a collaborative design task in the museum, we explore the potential of MTTs, which have already made their way into many museums. In terms of software, we rely on digital video tools that have a tradition of being used for observation and analysis skills development, including those from our own research (e.g., Goldman, 2004, 2007; Pea, 2006; Salomon, 1974; Spiro, Coulson, Feltovich, & Anderson, 1994; Zahn, Pea, et al., 2010).

Providing task structures for school group visits: Learning through collaborative visual design

Our concept builds on the learning through collaborative design approach for small groups of students (similar to *learning through design*, Kafai & Resnick, 1996). For decades, computer supported design tasks have been successfully implemented in schools (Harel, 1990; Kafai & Ching, 2001; Kafai & Resnick, 1996; Kolodner et al., 2003; Lehrer, Erickson, & Connell, 1994; Papert & Harel, 1991). Studies investigating collaborative visual design tasks in the domain of history have repeatedly shown how students acquire substantial historical

knowledge and visual skills when they design websites for a virtual history museum, using a historical newsreel and advanced video editing tools in their history lesson (Zahn, Krauskopf, et al., 2010; Zahn, Pea, et al., 2010). The conceptual underpinnings of the collaborative visual design approach have been described in earlier CSCL, JLS and iJCSC papers (Zahn, Pea, et al., 2010; Zahn et al., 2005; Zahn, Krauskopf, Hesse, & Pea, 2011, 2012). The basic assumption is that visual design leads to a deep and meaningful engagement with content, since it links people, form and content in a combined design space. People engaging in a collaborative design task have to negotiate not only the content (what should be designed) but also the form (how it should be designed for a specific audience; cf. Harel, 1990; Kafai & Ching, 2001), while coordinating their collaborative process at the same time (Zahn et al., 2012). In this sense, collaborative design constitutes a form of complex problem solving that is distributed over the cognitive systems of different people in a joint problem space (Zahn, Krauskopf, et al., 2010; cf. Roschelle & Teasley, 1995).

Providing digital tools with socio-constructivist potential: Multi-touch tabletop technology

Our concept includes the use of a multi-touch tabletop system (see Figure 1). Multi-touch tabletops are horizontal displays that allow simultaneous interaction of several people by touch input (Harris et al., 2009). They have received a lot of attention recently – not only by museums around the world, which incorporate MTTs in their visitor information systems and exhibits (e.g., Correia, Mota, Nóbrega, Silva, & Almeida, 2010; Geller, 2006; Hornecker, 2008), but also from the CSCL community (cf. Dillenbourg & Evans, 2011; Higgins, Mercier, Burd, & Hatch, 2011): MTTs are assumed to afford a more collaborative and constructivist working mode, thereby favorably suited for learning activities (cf. Kaplan et al., 2009). Dillenbourg & Evans (2011, p. 491) have characterized this as the “socio-constructivist flavor” of tabletops. The authors identify four key features, which may set MTTs apart from other computing devices in terms of their potential as learning tools: MTTs are designed for (i) co-located (ii) multiple users, who interact with each other via (iii) multiple forms of communication (i.e. gestures, talk, and actions), while primarily engaging in (iv) hands-on problem solving activities (i.e. manipulation of virtual objects). Building on this socio-constructivist potential and their increasing availability in museums, we consider MTTs to be promising candidates as tools for learning through visual design in museum art education. As a matter of fact, MTTs have already been used as a basis for design and video editing tasks (e.g., de Sa, Shamma, & Churchill, 2012; Rick, Rogers, Haig, & Yuill, 2009; Warnecke, Dohrmann, Jürgens, Rausch, & Pinkwart, 2011). However, as Dillenbourg & Evans (2011, p. 500) point out, although the medium may lend itself to a certain use, the technology itself does not have any “intrinsic pedagogical effects”. Whether or not a device can be utilized as an effective learning tool depends mainly on the respective task and on the conditions of its usage. As we intend to illustrate below, learning through collaborative visual design offers a framework that allows for meaningful tasks which make full use of the MTTs’ socio-constructivist potential.

Providing digital tools to support observation, analysis and reflection

Research in the learning sciences has provided ample evidence for using digital video technology to support a variety of socio-cognitive functions. Since early research on the educational value of films, where Salomon (1974) found that filmic coding elements can facilitate individual students’ mastery of mental skills necessary to attend to details of art works, video was repeatedly suggested as a tool for observation, analysis and reflection (Goldman, 2004, 2007; Pea, 2006; Pea et al., 2004; Spiro et al., 1994). It was shown in experiments how different video tools influence collaborative epistemic activities (grounding, negotiation, comparison and interpretation processes) for students using those video tools (e.g., Zahn, Pea, et al., 2010): Results from different studies show that the affordances of specific video tools (e.g., WebDiver™, Pea et al., 2004) better support learners’ interactions in making them more productive, compared to interactions performed with simple technological solutions. The results were improvements in learning outcomes and observation abilities (Zahn, Pea, et al., 2010). A field study further revealed that the differences in learners’ interactions persist in the real, “noisy” history classroom with 16-year old students (e.g., Zahn, Krauskopf, et al., 2010). In these studies it was the student’s task, to design a website for a virtual history museum, based on their analysis and decomposition of an original newsreel about the 1948 Berlin Blockade. In their products, the students reflected and commented upon the different camera and cutting techniques and the respective effects that these techniques evoked. Yet there are no studies available concerning the use of advanced video tools in art education or in art museums.

Learning through collaborative visual design in the art museum

The task concept

Derived from our multilevel CSCL framework, we have developed the following task structure: During a school visit, students individually browse the museum’s collection (cognitive level, see above), guided by an art educational group task (for possible tasks see scenario below and table 1), and collect digital reproductions of art works from the museum space (community level) using smartphones. The reproductions then can be

transferred to the MTT and processed by means of a video tool in order to produce a video clip in each group (socio-constructivist or small group level, see above). Thereby, students discuss with other students, which reproductions should be selected and how they should be further processed in the video clip production (socio-cognitive level). The clip will eventually be saved, which enables classmates, teachers or other people to share, discuss and enrich the product (class and community level).

In the next section, we will elaborate more deeply on how students will be supported during video clip production through socio-cognitive tool functions. The task workflow is depicted in Figure 1. For details regarding the general design task structure please refer to Zahn, Krauskopf, Hesse, & Pea (2010).

Implementation of socio-cognitive functions in the MTT video editing tool

We implemented specific video functions to support observation and reflection (socio-cognitive level, see above). The video functions consist of three basic socio-cognitive functions that are based on related research in cognitive psychology and the learning sciences: counterfactual image manipulations, highlighting of aspects within images, and linking images to create sequences. All functions constitute epistemic actions, as described by Kirsh and Maglio (1994). *Counterfactual image manipulations* are image modifications based on digital imaging filters, which change an image's formal appearance; for instance, a painting's strong light and dark contrasts can be changed to more subtle lighting differences and vice versa (see example scenario below). Counterfactual image manipulations were implemented in accordance with the concept of counterfactual thinking (e.g., Byrne, 2005) and the understanding of manipulations as cues to causality (cf. Woodward, 2005). We assume the following: Because generating mental alternatives to reality (counterfactual thinking) can assist people in thinking about the causes of effects and events (causal reasoning; e.g., Spellman & Mandel, 1999), creating alternatives to visual material (counterfactual image manipulations) can help to reflect upon the causes of certain visual effects. We implement a tool based strategy here, since it can be difficult to mentally manipulate a painting (i.e. to create a counterfactual painting; cf. Brandimonte, Hitch, & Bishop, 1992; Chambers & Reisberg, 1985). So far, four specific image filters have been implemented in our tool: Saturation, lightness, color temperature and vertical orientation (reversal of left/right orientation). For possible filter uses see Table 1. *Highlighting* is a well-known function that helps to bring out specific features of (visual) information. This function has already been investigated in detail with respect to arts education (cf. Salomon, 1974) and implemented in video analysis tools (Pea et al., 2004). Figure 1e shows an example of a highlighting function: An important detail of an art work is marked by a freehand drawing tool. Other planned functions contain spotlight, zoom and text annotation functionalities. *Linking images* to create new sequences is a tool function known from hypertext and hypervideo tools research in the learning sciences (Spiro & Jehng, 1990; E. Stahl & Bromme, 2004; Zahn & Finke, 2003). In our tool, reproductions of paintings can be linked to compose a slideshow, thereby facilitating, for example, comparisons regarding content, epoch, style and artist. For instance, the linking function offers a convenient way to tell the history of portrait painting or to give an overview of variations on the same motif throughout different epochs. The activity of linking images creates (new) connections among them. We assume that this leads to mental integration and, finally, to building respective knowledge structures.

According to our CSSL concept for supporting school visits to art museums, and as a basis for further research, we develop example learning scenarios in cooperation with a fine arts museum in Germany (2).

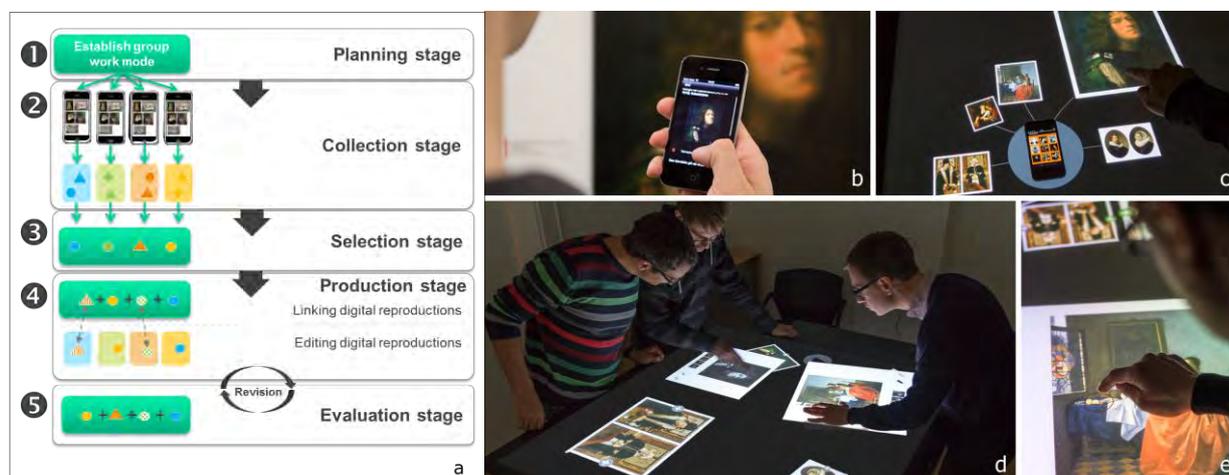


Figure 1. Task workflow (a; see also Zahn, Krauskopf, et al., 2010) and tool functions for collaborative visual design in a fine arts museum: Collecting digital reproductions in the museum via smartphones (b). Selecting collected reproductions from the smartphone at the MTT (c). Processing digital reproductions collaboratively at the MTT (d). Detail view of a highlighting function (drawing a circle around an important detail) (e).

Example Scenario: Light and Shadow

Let's say that an advanced art class visits a fine arts museum in order to study the effects of different light-shadow techniques. In order to learn about the use of light in art, the teacher asks the arts class to accomplish a collaborative visual design task to compare the effects of strong light-dark contrasts (a technique known as *chiaroscuro*) to more subtle light-dark contrasts. Students are divided into two groups. Each group will be encouraged to focus on one technique (*chiaroscuro* vs. subtle contrasts) and to "collect" digital reproductions of appropriate art works with smartphones while browsing the museum's exhibition (see Figure 1b). After returning to the MTT, the students can transfer their collected art works to the MTT and discuss in small groups which of them should be included in the video clip (see Figure 1c). Following this selection phase, the students will start the video editing (cf. Figure 1e). They use different image filters in order to find out how the appearance and the art work's effect will change due to alteration of contrast and lightness: Shifting an image towards a strong light-dark contrast evokes a *chiaroscuro* with dramatic appeal, while a reduction creates the opposite effect. If necessary, images can be annotated with short texts. Parallel to the editing, the students can decide on the sequence (cf. Figure 1d) in which the art works will be presented within the clip. A playback function enables the users to preview and revise their work. Subsequent to the editing phase, the students can save the clip. Finally, the clip will be made accessible through the museum's website and can be used for further discussion in class. The described scenario is just one example, for other possible applications please refer to Table 1. Also please note: The examples provide only a basis for supporting teachers and museum staff, who will decide on how the tool will be integrated in a particular school visit. Under ideal conditions, a specific learning scenario could be prepared in class, played out in the museum using the MTT and reinforced, again, in class. Art teachers can also invent different tool-based scenarios, which allow for a more spontaneous use of the tool on-site. Moreover – after careful reviewing by museum staff – successful user-generated scenarios will be provided at the MTT for other visitors, thereby establishing a scenario database to be used in the long run.

Table 1. Examples for topics and possible scenarios supported by MTT and video editing tool.

Topic	Possible scenario and technology support (MTT video editing)
Light and shadow	Students tell the history of shadow depiction in visual art (cf. Gombrich, 1995) in a video clip, or they illustrate the difference between cast and attached shadows (cf. Jacobson & Werner, 2004) with different reproductions.
Details/Motifs	Students look out for art works with interesting but less obvious details (e.g., symbols, motifs). They collect them and use highlighting functions to bring out the details in digital reproduction of the art works. By linking images in a slideshow, students recognize the same motif in art works across different artists, styles, epochs etc.
Composition	Students identify the compositional structure of a painting: They test hypotheses regarding the composition by sketching overlays directly on the reproduction. By flipping reproductions horizontally, students can learn about the importance of left-right orientation in paintings (cf. Bennett, Latto, Bertamini, Bianchi, & Minshull, 2010; Zaidel & Fitzgerald, 1994).
Color	Students collect reproductions of particular warm and cold colored paintings and discuss how a change in color temperature affects the impression and meaning of these paintings.

Discussion and Outlook

In this paper we have presented a multilevel CSCL concept for learning through collaborative visual design with an MTT based advanced video editing tool that can be adopted in fine arts museums to foster the observation and reflection required in academic art education. Our concept constitutes a method that can be tailored to further topics and learning objectives. Although the learning through visual design method is empirically validated, the present concept warrants further research: In a next step we intend to investigate the concept and the tool within a sample scenario. For further experimental research, multilevel approaches will be considered. Notwithstanding future research, in our view the concept proposed in this paper offers at least three advantages for museum and curricular art education: (i) Museums and schools are both social places which serve an educational purpose – our CSCL concept can account for that fact by supporting "21st century collaborative learning". (ii) The educational concept meets both curricular and museum educational requirements. It complements other art educational strategies and is highly adaptable to specific learning goals. (iii) Building on the high popularity of interactive displays among young adults (cf. nielsenwire, 2012), the present concept may help to bridge the gap between young students and old masters through new media. However, some curators are concerned that digital media could actually hinder appropriate art appreciation. Since this concern is justified in certain cases (cf. Hsi, 2003), it should be deliberated whether the use of a certain technology offers an added value (cf. Buder, 2007). As original paintings cannot be altered, annotated or even viewed from close distances,

we consider the added value of our video tool to be evident. Furthermore, our task ensures a close relationship between digital reproductions and original art works.

Endnotes

- (1) <http://www.metmuseum.org/en/learn/for-educators/lesson-plans-and-pre-visit-guides>
- (2) Herzog Anton Ulrich-Museum, Braunschweig, Germany.

References

- Barrett, T. (2008). Interactive Touring in Art Museums: Constructing Meanings and Creating Communities of Understanding. *Visual Arts Research*, 34(2(67)), 76–84.
- Bennett, K. M., Latto, R., Bertamini, M., Bianchi, I., & Minshull, S. (2010). Does left-right orientation matter in the perceived expressiveness of pictures? A study of Bewick's animals (1753-1828). *Perception*, 39(7), 970–981.
- Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. M. (1992). Manipulation of visual mental images in children and adults. *Journal of Experimental Child Psychology*, 53(3), 300–312. doi:10.1016/0022-0965(92)90041-4
- Buder, J. (2007). Net-Based Knowledge Communication in Groups. *Zeitschrift für Psychologie / Journal of Psychology*, 215(4), 209–217. doi:10.1027/0044-3409.215.4.209
- Byrne, R. M. J. (2005). *The Rational Imagination: How People Create Alternatives To Reality*. Cambridge, MA, US: MIT Press.
- Chambers, D., & Reisberg, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception and Performance*, 11(3), 317–328. doi:10.1037/0096-1523.11.3.317
- Correia, N., Mota, T., Nóbrega, R., Silva, L., & Almeida, A. (2010). A multi-touch tabletop for robust multimedia interaction in museums. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces 2010* (pp. 117–120). New York, NY, US: ACM.
- De Sa, M., Shamma, D., & Churchill, E. (2012). Live Mobile Collaboration for Video Production: Design, Guidelines, and Requirements. *Personal and Ubiquitous Computing*.
- Dillenbourg, P., & Evans, M. (2011). Interactive tabletops in education. *International Journal of Computer-Supported Collaborative Learning*, 6(4), 491–514. doi:10.1007/s11412-011-9127-7
- Duke, L. (2010). The Museum Visit: It's an Experience, Not a Lesson. *Curator: The Museum Journal*, 53(3), 271–279. doi:10.1111/j.2151-6952.2010.00028.x
- 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). Cambridge, UK: Cambridge University Press.
- Geller, T. (2006). Interactive Tabletop Exhibits in Museums and Galleries. *IEEE Computer Graphics and Applications*, 26(5), 6–11.
- Goldman, R. (2004). Video perspectivity meets wild and crazy teens: a design ethnography. *Cambridge Journal of Education*, 34(2), 157–178. doi:10.1080/03057640410001700543
- Goldman, R. (2007). Video Representations and the Perspectivity Framework: Epistemology, Ethnography, Evaluation, and Ethics. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 3–36). Mahwah, NJ, US: Lawrence Erlbaum Associates.
- Gombrich, E. H. (1995). *Shadows: The Depiction of Cast Shadows in Western Art*. London: National Gallery Company Ltd.
- Harel, I. (1990). Children as Software Designers: A Constructionist Approach for Learning Mathematics. *Journal of Mathematical Behavior*, 9(1), 3–93.
- Harris, A., Rick, J., Bonnett, V., Yuill, N., Fleck, R., Marshall, P., & Rogers, Y. (2009). Around the table: Are multiple-touch surfaces better than single-touch for children's collaborative interactions? In *Proceedings of the 9th international conference on Computer supported collaborative learning* (pp. 335–344). ISLS.
- Higgins, S. E., Mercier, E., Burd, E., & Hatch, A. (2011). Multi-touch tables and the relationship with collaborative classroom pedagogies: A synthetic review. *International Journal of Computer-Supported Collaborative Learning*, 6(4), 515–538. doi:10.1007/s11412-011-9131-y
- Hinrichs, U., Schmidt, H., & Carpendale, S. (2008). EMDialog: Bringing Information Visualization into the Museum. *IEEE Transactions on Visualization and Computer Graphics*, 14(6), 1181–1188. doi:10.1109/TVCG.2008.127
- Hooper-Greenhill, E. (2000). Changing Values in the Art Museum: rethinking communication and learning. *International Journal of Heritage Studies*, 6(1), 9–31. doi:10.1080/135272500363715
- Hornecker, E. (2008). "I don't understand it either, but it is cool" - visitor interactions with a multi-touch table in a museum. In *TABLETOP 2008: Proceedings of IEEE International Workshop on Horizontal Interactive Human Computer Systems* (pp. 113–120). Amsterdam, the Netherlands: IEEE.

- 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. doi:10.1046/j.0266-4909.2003.jca_023.x
- Jacobson, J., & Werner, S. (2004). Why cast shadows are expendable: Insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art. *Perception*, 33(11), 1369–1383. doi:10.1068/p5320
- Jacucci, G., Morrison, A., Richard, G. T., Kleimola, J., Peltonen, P., Parisi, L., & Laitinen, T. (2010). Worlds of information: designing for engagement at a public multi-touch display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2267–2276). New York, NY, US: ACM. doi:10.1145/1753326.1753669
- Kafai, Y. B., & Ching, C. C. (2001). Affordances of Collaborative Software Design Planning for Elementary Students' Science Talk. *Journal of the Learning Sciences*, 10, 323–363. doi:10.1207/S15327809JLS1003_4
- Kafai, Y. B., & Resnick, M. (1996). *Constructionism in Practice: Designing, Thinking, and Learning in A Digital World*. Mahwah, NJ, US: Routledge.
- Kaplan, F., DoLenh, S., Bachour, K., Kao, G. Y., Gault, C., & Dillenbourg, P. (2009). Interpersonal Computers for Higher Education. In P. Dillenbourg, J. Huang, & M. Cherubini (Eds.), *Interactive Artifacts and Furniture Supporting Collaborative Work and Learning* (Vol. 10, pp. 1–17). New York, NY, US: Springer.
- Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4), 513–549. doi:10.1016/0364-0213(94)90007-8
- Knipfer, K., Mayr, E., Zahn, C., Schwan, S., & Hesse, F. W. (2009). Computer support for knowledge communication in science exhibitions: Novel perspectives from research on collaborative learning. *Educational Research Review*, 4, 196–209. doi:10.1016/j.edurev.2009.06.002
- Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., ... Ryan, M. (2003). Problem-Based Learning Meets Case-Based Reasoning in the Middle-School Science Classroom: Putting Learning by Design™ into Practice. *The Journal of the Learning Sciences*, 12(4), 495–547.
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95(4), 489–508. doi:10.1348/0007126042369811
- Lehrer, R., Erickson, J., & Connell, T. (1994). Learning by designing hypermedia documents. *Computers in the Schools*, 10(1-2), 227–254. doi:10.1300/J025v10n01_15
- Morrone, A. (2006). *Guidelines for measuring cultural participation*. Montreal, Quebec, Canada: UNESCO Institute for Statistics.
- National Art Education Association. (1994). *The National Visual Arts Standards*. Reston, VA, US: National Art Education Association.
- Nespor, J. (2000). School field trips and the curriculum of public spaces. *Journal of Curriculum Studies*, 32(1), 25–43. doi:10.1080/002202700182835
- nielsenwire. (2012, February 23). Introducing Generation C: Americans 18-34 Are the Most Connected. Retrieved from http://blog.nielsen.com/nielsenwire/online_mobile/introducing-generation-c/
- Papert, S., & Harel, I. (1991). Situating Constructionism. In S. Papert & I. Harel (Eds.), *Constructionism*. Norwood, NJ, US: Ablex Publishing Corporation.
- Pea, R. (2006). Video-as-Data and Digital Video Manipulation Techniques for Transforming Learning Sciences Research, Education, and Other Cultural Practices. In J. Weiss, J. Nolan, J. Hunsinger, & P. Trifonas (Eds.), *The International Handbook of Virtual Learning Environments* (pp. 1321–1393). Dordrecht, the Netherlands: Springer.
- Pea, R., Mills, M., Rosen, J., Dauber, K., Effelsberg, W., & Hoffert, E. (2004). The Diver project: interactive digital video repurposing. *IEEE MultiMedia*, 11(1), 54 – 61. doi:10.1109/MMUL.2004.1261108
- Pekarik, A. J., Doering, Z. D., & Karns, D. A. (1999). Exploring Satisfying Experiences in Museums. *Curator: The Museum Journal*, 42(2), 152–173. doi:10.1111/j.2151-6952.1999.tb01137.x
- Rick, J., Rogers, Y., Haig, C., & Yuill, N. (2009). Learning by doing with shareable interfaces. *Children, Youth and Environments*, 19(1), 321–342.
- Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer supported collaborative learning* (Vol. 128, pp. 69–97). Germany: Springer.
- Salomon, G. (1974). Internalization of filmic schematic operations in interaction with learners' aptitudes. *Journal of Educational Psychology*, 66(4), 499–511. doi:10.1037/h0036753
- Scardamalia, M. (2002). Liberal Education in a Knowledge Society. In B. Smith (Ed.), *Collective Cognitive Responsibility for the Advancement of Knowledge* (pp. 67–98). Peru, Illinois, US: Open Court Publishing.
- Spellman, B. A., & Mandel, D. R. (1999). When Possibility Informs Reality. Counterfactual Thinking as a Cue to Causality. *Current Directions in Psychological Science*, 8(4), 120–123. doi:10.1111/1467-8721.00028

- Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1994). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading (4th ed.)* (pp. 602–615). Newark, DE, US: International Reading Association.
- Spiro, R. J., & Jehng, J.-C. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. J. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (pp. 163–205). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Stahl, E., & Bromme, R. (2004). Learning by Writing Hypertext: A Research Based Design of University Courses in Writing Hypertext. In G. Rijlaarsdam, H. Bergh, & M. Couzijn (Eds.), *Effective Learning and Teaching of Writing* (Vol. 14, pp. 547–559). Dordrecht, the Netherlands: Springer.
- Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge, MA, US: MIT Press.
- Stahl, G., Koschmann, T., & Suthers, D. D. (2006). Computer-supported collaborative learning: An Historical Perspective. In R. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 409–426). Cambridge University Press.
- Suthers, D. D., & Hundhausen, C. D. (2003). An Experimental Study of the Effects of Representational Guidance on Collaborative Learning Processes. *Journal of the Learning Sciences*, 12(2), 183–218. doi:10.1207/S15327809JLS1202_2
- Trant, J. (2006). Exploring the potential for social tagging and folksonomy in art museums: Proof of concept. *New Review of Hypermedia and Multimedia*, 12(1), 83–105. doi:10.1080/13614560600802940
- Warnecke, T., Dohrmann, P., Jürgens, A., Rausch, A., & Pinkwart, N. (2011). Collaborative Learning through Cooperative Design Using a Multitouch Table. In Y. Luo (Ed.), *Cooperative Design, Visualization, and Engineering* (Vol. 6874, pp. 22–29). Germany: Springer.
- Winner, E. (2007). Visual thinking in arts education: Homage to Rudolf Arnheim. *Psychology of Aesthetics, Creativity, and the Arts*, 1(1), 25–31. doi:10.1037/1931-3896.1.1.25
- Woodward, J. (2005). *Making Things Happen: A Theory of Causal Explanation*. New York, NY, US: Oxford University Press.
- Zahn, C., & Finke, M. (2003). Collaborative knowledge building based on hyperlinked video. In B. Wason, R. Baggetun, U. Hoppe, & S. Ludvigsen (Eds.), *Proceedings of the International Conference on Computer Support for Collaborative Learning - CSCL 2003, COMMUNITY EVENTS - Communication and Interaction* (pp. 173–175). Bergen, NO: InterMedia.
- Zahn, C., Hesse, F., Finke, M., Pea, R., Mills, M., & Rosen, J. (2005). Advanced digital video technologies to support collaborative learning in school education and beyond. In *Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!* (pp. 737–742). International Society of the Learning Sciences.
- Zahn, C., Krauskopf, K., Hesse, F. W., & Pea, R. (2009). Participation in knowledge building “revisited”: reflective discussion and information design with advanced digital video technology. In *Proceedings of the 9th international conference on Computer supported collaborative learning - Volume 1* (pp. 596–600). International Society of the Learning Sciences.
- Zahn, C., Krauskopf, K., Hesse, F. W., & Pea, R. (2010). Digital video tools in the classroom: How to support meaningful collaboration and critical advanced thinking of students? In M. S. Khine & I. M. Saleh (Eds.), *New Science of Learning: Cognition, Computers and Collaboration in Education* (pp. 503–523). New York, NY, US: Springer.
- Zahn, C., Krauskopf, K., Hesse, F. W., & Pea, R. (2011). Digital Media in the Classroom: A Study on How to Improve Guidance for Successful Collaboration and Learning in Student Teams. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings*. (Vol. 1, pp. 152–159). Hong Kong: International Society of the Learning Sciences.
- Zahn, C., Krauskopf, K., Hesse, F. W., & Pea, R. (2012). How to improve collaborative learning with video tools in the classroom? Social vs. cognitive guidance for student teams. *International Journal of Computer-Supported Collaborative Learning*, 7(2), 259–284. doi:10.1007/s11412-012-9145-0
- Zahn, C., Pea, R., Hesse, F. W., & Rosen, J. (2010). Comparing Simple and Advanced Video Tools as Supports for Complex Collaborative Design Processes. *Journal of the Learning Sciences*, 19(3), 403–440. doi:10.1080/10508401003708399
- Zaidel, D. W., & Fitzgerald, P. (1994). Sex of the face in Western art: Left and right in portraits. *Empirical Studies of the Arts*, 12(1), 9–19.

Acknowledgments

The project is funded by the Leibniz Association as part of the „Wettbewerbsfonds 2011“.

Navigating Through Controversial Online Discussions: The Influence of Visualized Ratings

Jürgen Buder, Knowledge Media Research Center, Tübingen Germany, j.buder@iwm-kmrc.de
Christina Schwind, Knowledge Media Research Center, Tübingen Germany, c.schwind@iwm-kmrc.de
Anja Rudat, Knowledge Media Research Center, Tübingen Germany, a.rudat@iwm-kmrc.de
Daniel Bodemer, University of Duisburg-Essen, Media-Based Knowledge Construction, Germany, bodemer@uni-due.de

Abstract: Awareness tools using visualized ratings of other people can help recipients to find their way in crowded information spaces. This article reports on a laboratory experiment that investigates how the design of an awareness tool impacts the navigation behavior of recipients within an online discussion forum. In the experiment, 127 participants read through a forum discussion for which posts were rated by average agreement and average quality. Depending on the experimental condition, posts were visualized along continua indicating average agreement ratings (absent vs. present) and/or average quality ratings (absent vs. present). It was hypothesized that an agreement visualization will increase the salience of both high-agreement and low-agreement posts, whereas a quality visualization will only increase the salience of high-quality posts. Measuring reading times for discussion posts, these hypotheses could be confirmed. However, neither recipients' attitude strength nor recall of discussion arguments was affected.

Introduction

One of the hallmarks of computer technologies is their ability to create social situations over a distance. Email, online discussion forums, virtual environments, multiplayer games, social networks, microblogging, and wikis: All these technologies enable communication and collaboration among persons who are spatially separated. But communicating over a distance is not without its problems, as nearly 30 years of research has shown that computer-mediated communication is impoverished in comparison to the richness of face-to-face environments (Kiesler, Siegel, & McGuire, 1984). For instance, the lack of copresence implies that it is much more difficult in a computer-mediated environment to refer to spatial relations, for example, pointing at an object (Kramer, Oh, & Fussell, 2006). Similarly, other non-verbal cues like facial expressions, intonation or pitch, are not (or not sufficiently) transmitted (Walther & Tidwell, 1995). And finally, social cues about the person one is interacting with (e.g., gender, age) might be missing. In the early 1990s researchers in the field of computer-supported cooperative work began to address this problem by developing tools that try to re-create some of the richness of face-to-face environments (e.g., Heath & Luff, 1992). These so-called group awareness tools (Gutwin & Greenberg, 2002) were able to show who is present in an environment, indicate the activities that other persons are currently involved in, or provide background information about communicators through profile pages.

Starting in the mid-2000s, researchers in the field of computer-supported collaborative learning (CSCL) have started to adapt, develop, and test group awareness tools with the goal of supporting collaborative learning processes. In conjunction with this shift of application fields came a shift in the design of group awareness tools. Rather than trying to imitate the contextual richness of copresent scenarios, CSCL awareness tools provided information that would be difficult or even impossible to yield in face-to-face environments, thus lending an added value to computer-mediated communication (Buder, 2007). For instance, CSCL awareness tools provide information about the knowledge of learning partners (Engelmann, Dehler, Bodemer, & Buder, 2009), their degree of comprehension (Dehler, Bodemer, Buder, & Hesse, 2011), their assessments on tests (Sangin, Molinari, Nüssli, & Dillenbourg, 2011), perceived personality traits (Phielix, Prins, Kirschner, Erkens, & Jaspers, 2011), or their opinions (Buder & Bodemer, 2008).

In particular, Buder and Bodemer (2008, 2011) conducted two experiments that investigated the use of group awareness tools for controversial online discussions. Small groups of learners (3-4 members) were required to discuss two conflicting hypotheses from various science topics. By providing differential access to learning material prior to discussion, it was ensured in all conditions of both experiments that one learner (the informed minority) of each group favored a "correct" hypothesis. However, this learner was confronted with a dissenting majority who favored an "incorrect" hypothesis. The task of the groups was to discuss the conflicting evidence before making a decision about the better hypothesis. Among other things, both experiments compared groups using an awareness tool and groups without a tool. In the tool conditions, group members were required to rate the discussion posts of their collaborators on two dimensions, namely, agreement and novelty. The awareness tool computed the average ratings for each contribution and fed these data into a two-dimensional visualization in real time where each dot represented a discussion post. On the horizontal axis of the visualization, dots were displayed along an agreement dimension. The rationale behind this was to visually

separate majority contributions (high agreement) from minority contributions (low agreement) in order to create an awareness that different opinions were available. On the vertical axis of the visualization, dots were displayed along a novelty dimension. The idea behind this was that minority contributions garner higher novelty ratings than majority contributions, and therefore become highly salient. The rationale behind these visualizations can be related to the notion of representational guidance (Suthers & Hundhausen, 2003). According to this concept, different representational formats lead to differences in how individuals and groups process information: Representations constrain the way in which learners think about an object (e.g., in terms of agreement and novelty), and they make some parts of a representation more salient than others (e.g., novel minority posts). The results of both experiments indicate that without awareness tool, post-discussion preferences of groups and individuals were leaning toward the incorrect majority viewpoint. However, in groups that were supported by the awareness tool, the social influence of minority members could be strengthened, resulting in a higher likelihood that the correct hypothesis was chosen by individuals and groups. In sum, these experiments showed that carefully designed awareness tools can have an impact on the behavior of collaborating groups.

The current work builds on these prior findings, but takes the project in two new directions. The first new direction is associated with a change of the learning setting. While the prior studies by Buder and Bodemer (2008, 2011) involved small groups of learners in formal contexts, we are now investigating potentially large groups in informal contexts, that is, controversial discussions in online forums that are part of the Web portals of newspapers or magazines. The rationale for this shift is that group awareness tools can act as filters for information seeking (Dourish & Chalmers, 1994). In small groups, learners can actually attend to all information that is produced by their collaborators. However, in a large online forum there can be literally thousands of discussion posts on a single topic, and participants can only attend to a small portion of the available information.

The second departure from the prior experiments has to do with a deeper understanding of the mechanisms that help explaining the effectiveness of awareness tools. While the experiments from Buder and Bodemer (2008, 2011) showed that visualizations by a rating-based awareness tool can be effective in shaping the behavior of a group, the present work is focused on the way in which these visualizations are actually used. In other words, the current study investigates how visualizations of rating dimensions impact the navigation of recipients. The general idea is that different rating dimensions exert different types of representational guidance (Suthers & Hundhausen, 2003).

In particular, we distinguish between two types of behavior termed bipolar and unipolar navigation. If discussion posts are visualized alongside a continuum, bipolar navigation means that both ends of the continuum become salient. This should be the case for the visualization of agreement ratings, as such a visualization gives insights into the pros and cons of a discussion. In contrast, unipolar navigation means that only one end of a continuum becomes salient. In previous studies, this was accomplished with the novelty dimension. However, as the current study involves discussions in online forums where participants can enter at different times, novelty becomes a relative concept. Therefore, the current study tries to trigger unipolar navigation through the use of a quality visualization. Visualization of quality ratings is likely to exert representational guidance in a way that high-quality posts become more salient than low-quality posts.

Hypotheses

The present experiment investigates these issues by confronting individuals with the content of a large online discussion, with only limited time to read all discussion posts. If the posts are visualized along an agreement rating continuum, it might become likely that recipients will read posts with both high and low agreement ratings (bipolar navigation) as this provides insights into the different viewpoints of the controversial discussion (pro vs. con). A good balance between attention towards the pros and cons of a controversial discussion should not only be found with regard to the frequencies of opened discussion posts, but also with regard to subsequent reading times. We measured navigation by reading times because they provide better insights into the processing of information. As a consequence of bipolar navigation, recipients might achieve a more balanced view on the controversial issue. This, in turn should be accompanied by an attenuation of a recipient's post-discussion attitude strength compared to the pre-discussion attitude strength. This leads to the following hypotheses:

Hypothesis 1: Participants who are provided with an agreement visualization should show smaller reading time differences between pro and con discussion posts than participants who are not provided with an agreement visualization.

Hypothesis 2: Participants who are provided with an agreement visualization should show a weakening in attitude strength, whereas participants who are not provided with an agreement visualization should show an enhancement in attitude strength.

If the discussion posts are visualized along a quality rating continuum, it might become likely that recipients will read high-quality posts at the expense of low-quality posts (unipolar navigation). As high-quality

contributions are most likely to contain the main arguments of a discussion, the availability of a quality visualization should also lead to better memory for these crucial arguments. This leads to the following hypotheses:

Hypothesis 3: Participants who are provided with a quality visualization should have longer reading times for the top quality contributions than participants who are not provided with a quality visualization.

Hypothesis 4: Participants who are provided with a quality visualization should better recall the crucial discussion arguments than participants who are not provided with a quality visualization.

Method

Participants

Data were collected from 127 participant volunteers (38 male). All participants were students, recruited via a university mailing list. For their participation in the experiment, which took 60 minutes, participants were paid 8 €. Alternatively, the students could get a certificate of their participation if needed for course requirements. Participants' age ranged from 18 to 58 years ($M = 24.94$ years, $SD = 5.22$, one missing value).

Design

This lab study used a 2 x 2 design to explore the effect of the two visualizations of agreement ratings and quality ratings. Thus, we set up four experimental conditions in which the visualizations of agreement ratings and/or quality ratings were either available or not (see Table 1): a condition without any visualization (*no Visualization*), a condition with only a visualization of the agreement ratings (*Agreement only*), a condition with only a visualization of the quality ratings (*Quality only*), and a condition with a combination of both visualizations, agreement ratings and quality ratings (*Combination*).

Table 1: 2 x 2 factorial design of the experiment.

Availability of rating visualization		Quality	
		NO	YES
Agreement	NO	<i>no Visualization</i>	<i>Quality only</i>
	YES	<i>Agreement only</i>	<i>Combination</i>

Material

Discussion forum: We used a discussion from an online forum that is part of the Web portal of a German news magazine. The discussion is about pro and cons of the three-tier school system, which is a controversially discussed topic in German public discourse. We took the first 137 original posts and coded them for their position (pro or contra the discussed three-tier school system, Interrater-Reliability Cohen's $\kappa = .64$) and the arguments contained therein.

Arguments: Based on the material we identified eleven main arguments, speaking for or against the three-tier school system (e.g., "Leveling down is bad and has negative consequences. It results in overload or underload." as an argument pro three-tier school system; "It is bad to select pupils too early because their potential might be unexploited." as an argument contra three-tier school system).

Ratings: In a pre-study ($N = 67$), all discussion posts have been rated on their quality on a five point Likert-scale (1 *very low* to 5 *very high*). Out of the pro-posts and out of the contra-posts we identified the ten qualitative best posts (so called *TOP 10* posts), five of pro and contra each. All main arguments can be found in these TOP 10 posts. For reasons of experimental systematicity, the ranges of ratings on agreement and quality have been adapted in a way that, for example, all posts could be identified separately in the visualization.

List display: The discussion was presented in a simulated online forum environment and the discussion posts were presented in all conditions in a list of their headings in their original chronological order. In the no Visualization condition only the list display was presented.

Visualizations: Participants in the three other conditions (Agreement only, Quality only, and Combination) were additionally provided with embedded visualized rating dimensions. The visualization represented all discussion contributions as dots along continua, with the position of each dot indicating the average ratings that a contribution received.

Measures

As dependent variables we measured reading times, attitude strength, and recall. Reading times include three different types: the reading time of the pro-posts and the contra-posts, and the reading time of the TOP 10 posts.

Reading time PRO-CON. Reading time PRO-CON was measured for testing the impact of different types of visualizations on bipolar navigation. We measured and compared reading time PRO-CON by calculating the difference of the reading time of pro-posts and the reading time of the contra-posts.

Attitude strength. The attitude indication was represented by a word pair ranging from *weak* (= 1) to *strong* (= 5) attitudes about the subject matter. The attitude strength was calculated as the difference between the pre-discussion attitude strength and the post-discussion attitude strength. Zero indicates that there was no difference in attitude strengths, whereas negative values indicate an enhancement of attitude strengths, and positive values indicate a weakening of attitude strengths.

Reading time TOP 10. We measured reading times of the TOP 10 posts. In this way, it was possible to test the impacts of different types of visualizations on unipolar navigation.

Recall. We measured free recall of arguments. For that, participants wrote down as many arguments out of the whole discussion they could remember. For measurement, we coded the number of recalled arguments with a self-developed coding scheme (Interrater-Reliability Cohen's $\kappa = .72$).

Procedure

We recruited participants from a database of all university students from a German university via mailing list and asked them to take part in an “online discussion forum - study” where they would have to read an online discussion forum. All instructions and materials were presented on a computer screen. Each participant was randomly assigned to one of the four experimental conditions. After having read the instructions, participants indicated their pre-attitudes about the three-tier school system. Then, all participants had 20 minutes time to navigate freely in the online discussion forum. To avoid them rushing through the posts without reading, they were told it is hardly possible to read all 137 posts within 20 minutes. Participants in the no Visualization condition saw an online forum without any ratings, but only with a list display. Participants in the Agreement only condition were additionally provided with visualized average ratings of agreement of all discussion posts. Each dot in the visualization represented one corresponding post (see Figure 1). In contrast, participants in the Quality only condition saw a visualization with quality ratings of all discussion posts. Again, each dot in the visualization represented the average rating the corresponding discussion post received. Participants in the Combination condition were provided with visualized ratings of both, agreement and quality (see Figure 2) additionally to the list display. Again, in the visualization each dot marked one corresponding discussion post. The visualization in the Combination condition was two-dimensional. That means, on the horizontal axis the agreement ratings were shown, and on the vertical axis the quality ratings were shown.

Participants could access forum posts either by scrolling through the list display, or by clicking on a dot in the visualization. Only by clicking, the corresponding post opened on the bottom of the screen. In all visualization conditions, the headings of the posts could be made visible by moving the mouse over a dot of the visualization. Posts that have been read were marked in orange in both the list display and the visualization.

After 20 minutes with the online forum, post-attitudes about the three-tier school system were measured. Then, participants performed the recall task. For that, they were told to keep in mind that arguments should not be confused with posts as one single post could include none, one or more arguments. At the end, participants were thanked and debriefed.

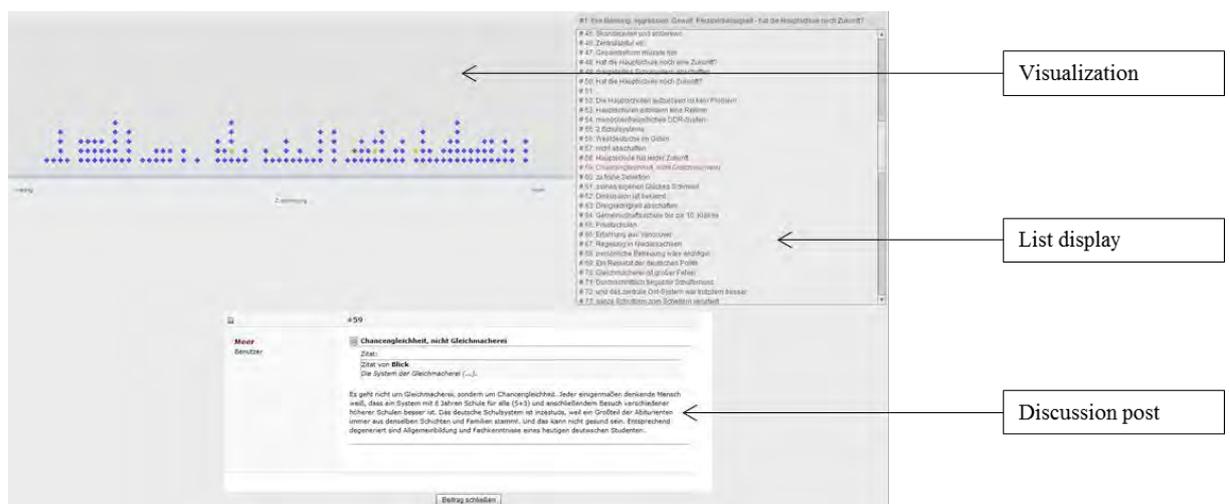


Figure 1. Screenshot of the presented online forum in the Agreement only condition.

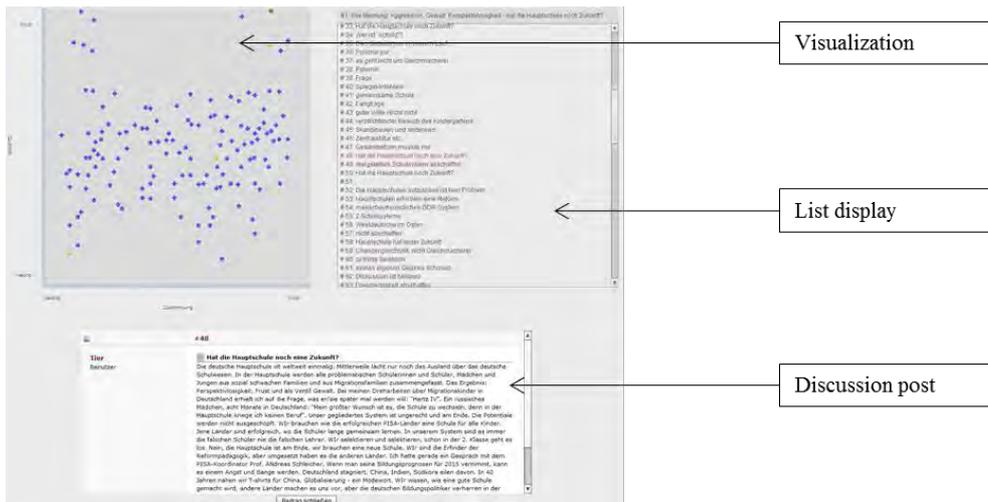


Figure 2. Screenshot of the presented online forum in the Combination condition.

Results

The preliminary analysis and the main results were based on logfile data (opening frequencies, reading time PRO-CON and reading time TOP 10). Navigation and recall measures were analyzed using analyses of variance (ANOVAs). Before addressing the main results, the preliminary analysis of the opening frequencies investigated in how far participants in the three conditions with visualizations (Agreement only vs. Quality only vs. Combination) used the visualization vs. the list display for navigation.

Preliminary analysis. We explored if participants show differences in using the visualizations or using the list display to navigate within the discussion. The measurement expressed the difference between the amount of post openings via the visualization and the amount of post openings via the list display. This means that positive values indicate that participants used mainly the visualization, whereas negative values indicate that participants used mainly the list display for navigation.

To explore the difference in opening frequencies via the visualization and via the list display, a one-factorial ANOVA with condition (Agreement only vs. Quality only vs. Combination) as independent variable and the difference of opening frequencies (opening frequencies for the visualization minus opening frequencies for the list display) as dependent variable was conducted. Results show a significant main effect for condition ($F(2, 92) = 5.75, p = .004, \text{partial } \eta^2 = .11$). Subjects in the Agreement only condition ($M = 5.81, SD = 16.94$) as well as subjects in the Quality only condition ($M = 1.87, SD = 16.02$) mainly used the visualization, whereas subjects in the Combination condition ($M = -8.19, SD = 18.02$) used the list display for navigation (see Figure 3).

Pairwise comparisons using Tukey-HSD revealed that there was no difference between the Agreement only condition and the Quality only condition ($p = .630$) regarding the opening frequencies via visualization and via list display. However, subjects in the Combination condition used the visualization less frequently compared to subjects in the Agreement only condition ($p = .004$) and marginally less frequently compared to subjects in the Quality only condition ($p = .055$).

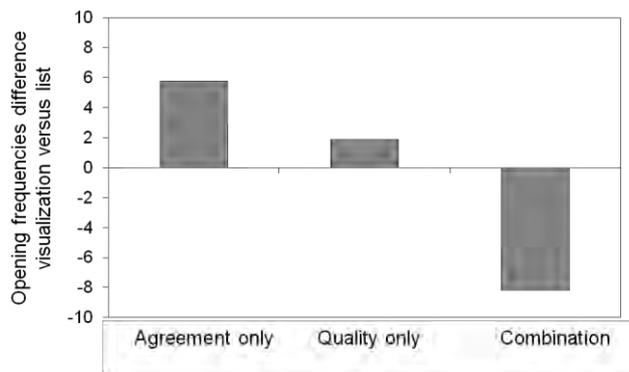


Figure 3. Opening frequencies difference (visualization versus list display) for the three conditions with visualization.

Reading time PRO-CON. In Hypothesis 1, it was predicted that the presence of an agreement visualization will support bipolar navigation: Participants in the conditions with agreement visualization will show more balanced reading times, meaning that they equally distribute their overall reading time to pro posts and to con posts. To test this prediction, we computed a two-factorial ANOVA with Agreement (agreement visualization: no vs. yes) and Quality (quality visualization: no vs. yes) as independent variables and reading time PRO-CON as dependent variable. In line with the hypothesis, the analysis for the reading time PRO-CON revealed a significant main effect, $F(1, 123) = 8.37, p = .005$, partial $\eta^2 = .06$. As shown in Figure 4, participants in the conditions with agreement visualization ($M = 27.27, SD = 198.85$) showed more balanced reading time PRO-CON than participants in the conditions without agreement visualization ($M = 130.33, SD = 199.08$).

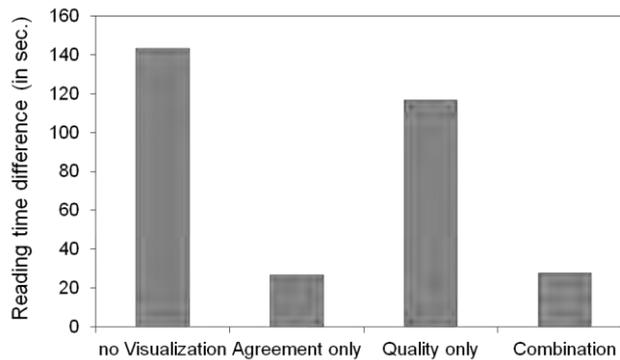


Figure 4. Reading time PRO-CON (in sec.) in the four conditions.

Attitude strength. We hypothesized that participants in the conditions with the agreement visualization will show a weakening in attitude strength, whereas participants in the conditions without the agreement visualization will show an enhancement in attitude strength (Hypothesis 2). We analyzed this hypothesis using a two-factorial ANOVA with Agreement (agreement visualization: no vs. yes) and Quality (quality visualization: no vs. yes) as independent variables and the difference of pre-discussion attitude strength and post-discussion attitude strength as dependent variable. Contrary to our expectations, no difference between the four conditions was revealed: Participants with the agreement visualization ($M = -0.03, SD = 0.87$) showed the same attitude strength difference between pre-discussion and post-discussion as did participants in the conditions where the agreement visualization was not present ($M = -0.19, SD = 0.78$). There were no differences for agreement visualization ($F(1, 123) = 1.18, ns$), for quality visualization ($F(1, 123) < 1, ns$) and no interaction effect ($F(1, 123) < 1, ns$). This means that no change in attitude strength was found in any of the four conditions.

Reading time TOP 10. It was hypothesized that the quality visualization will support unipolar navigation. This means that for conditions with quality visualization, it was predicted that participants will show longer reading times for the TOP 10 posts than participants in the conditions without quality visualization (Hypothesis 3). A two-factorial ANOVA with Agreement (agreement visualization: no vs. yes) and Quality (quality visualization: no vs. yes) as independent variables and reading time TOP 10 as dependent variable was conducted. The analysis yielded the expected highly significant main effect, $F(1, 123) = 56.91, p < .001$, partial $\eta^2 = .32$. Participants in the conditions with quality visualization ($M = 379.16, SD = 214.05$) read TOP 10 posts longer than participants in the conditions without quality visualization ($M = 157.58, SD = 103.08$); see Figure 5.

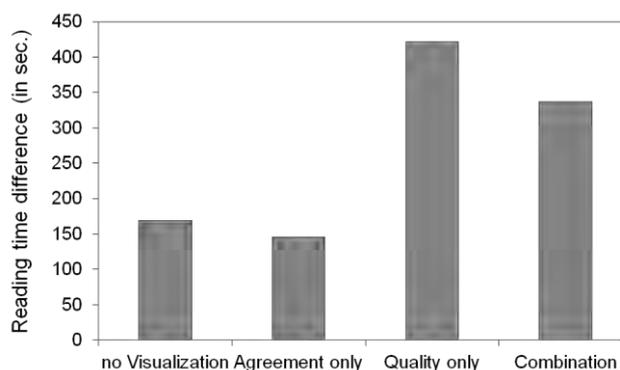


Figure 5. Reading time TOP 10 (in sec.) in the four conditions.

Recall. Since TOP 10 posts cover all crucial arguments in the discussion, we expected a similar result pattern for reading time TOP 10 and for recall. Therefore, it was hypothesized that participants in the conditions with quality visualization will show better recall for the crucial discussion arguments than participants in the conditions without quality visualization. To test Hypothesis 4, we computed a two-factorial ANOVA with Agreement (agreement visualization: no vs. yes) and Quality (quality visualization: no vs. yes) as independent variables and recall as dependent variable. Contrary to our expectations, no differences between the four conditions were found (no Visualization: $M = 3.56$, $SD = 1.32$; Agreement only: $M = 3.50$, $SD = 1.63$; Quality only: $M = 3.74$, $SD = 1.83$; Combination: $M = 3.46$, $SD = 1.67$). Participants with the quality visualization ($M = 3.46$, $SD = 1.76$) recalled the same amount of arguments as did participants in the conditions where the quality visualization was not present ($M = 3.53$, $SD = 1.47$). There were neither differences for agreement ($F(1, 123) = 1.15$, *ns*), nor for quality ($F(1, 123) < 1$, *ns*) and no interaction effect ($F(1, 123) < 1$, *ns*). This means that overall participants in all four conditions recalled the same amount of arguments.

Discussion

We have tested an awareness tool that employs visualized user ratings of discussion posts. In order to test whether different forms of visualizations give rise to different navigation styles, we conducted an experiment that varied the presence or absence of an agreement rating visualization and a quality rating visualization. A number of results provide insights into the design of awareness tools and its implications for learner navigation. First, results indicate that making agreement ratings available leads to bipolar navigation, that is, readers of the discussion forum tended to read both pro and con contributions of the controversial discussion to the same degree. Ensuring balanced attention to both viewpoints pertaining to a controversial issue is an important antecedent of unbiased opinion formation and critical thinking (Schwind, Buder, Cress, & Hesse, 2012). Hence, the current study gives some insight into how the design of awareness tools can contribute to an open-minded stance of learners towards a controversial topic. Second, the study could show that the availability of quality ratings leads to unipolar navigation, that is, readers of the discussion forum spent more time reading the TOP 10 posts once this information was made salient. Third, our results show that combining both rating visualizations in a two-dimensional display comes with a cost: This could be the main reason why learners in the Combination condition seemed to prefer a standard list display for navigation. Taken together, the navigation results show that the design of awareness tools can exert different types of representational guidance, but that the complexity of a graphical representation could be a roadblock that might prevent a group awareness tool from unfolding its full potential. For instance, in the present study the group awareness tool did *not* affect attitude strength and recall. The reason for the lack of effects on cognitive variables might be explained by the status of our participants as relatively passive recipients who could only read discussion posts. Our prior experiments involving active participants have shown that an awareness tool also impacts individual and perceived group preferences. This would be in line with the general notion that active participation is a key to effective collaboration (Cohen, 1994). However, the lack of cognitive effects can also be explained by the specific paradigm used in the study: The measurements for the cognitive variables (i.e. attitude strength and recall) were delayed measurements. Therefore, it might be interesting to investigate in follow-up studies if it is possible to affect cognitive variables when directly measured after the confrontation with the group awareness tool. Another question for further research is whether learners with higher involvement and/or interest in the topic show more effects on cognitive variables than the student population in the present study. Thus, topical involvement and interest might be moderators for the found results.

The current study employed an experimental approach that allowed us to carefully control for opening frequencies and reading times with regard to each discussion post. While the experimental control had the advantage that navigational behavior could be assessed in greater detail than before, it also gave rise to a number of limitations having to do with the selection of participants and the temporal nature of the study. First, participants were not self-selected. This is in contrast to real-world online forums in which it is likely that readers will have a personal interest in the topic under discussion. It would be interesting to see whether the results that we have found (and even more, the results that we haven't found) could be shown with a sample of participants who have a vested interest in the controversy. Second, participants were confronted with a static picture of a controversial group discussion. While it could be said that even in real-world online forum each reader only sees a static picture of the discussion when logging in, it cannot be denied that there is an ongoing dynamic in a discussion that unfolds over repeat visits in a forum. It would be fascinating to see how a rating-based group awareness tool would fare as a support mechanism of a large, controversial online forum, particularly at the early and formative stages of discussion. Will the distribution of agreement ratings lead to changes in the type of posts that are made? Will high-quality posts spur or inhibit further discussions? These questions can only be tapped into by employing a rating-based group awareness tool in an actual online forum. Consequently, transferring group awareness tools into the "real world" is an obvious step once the mechanisms of these tools are better understood.

Collecting user ratings has become a widespread phenomenon in Web 2.0 environments. However, most systems make use of relative primitive ratings (5-star ratings or “Like it” ratings). We set out to explore whether different forms of representational guidance can be achieved by using more specific rating dimensions. Our results lend some support to the idea that navigation of learners in large online environments can be shaped in educationally meaningful ways.

References

- Buder, J. (2007). Net-based knowledge-communication in groups. *Zeitschrift für Psychologie*, 215(4), 209-217.
- Buder, J., & Bodemer, D. (2008). Supporting controversial CSCL discussions with augmented group awareness tools. *International Journal of Computer-Supported Collaborative Learning*, 3, 123-139.
- Buder, J., & Bodemer, D. (2011). Group awareness tools for controversial CSCL discussions: Dissociating rating effects and visualized feedback effects. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings* (Vol. I, pp. 358-365). Hong Kong, China: International Society of the Learning Sciences.
- Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64, 1-35.
- Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior*, 27, 1068-1078.
- Dourish, P., & Chalmers, M. (1994). Running out of space: Models of information navigation. In G. Cockton, S. W. Draper, G. R. S. Weir (Eds.), *Proceedings of the HCI Conference*. Cambridge, UK: Cambridge University Press.
- Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior*, 25, 949-960.
- Gutwin, C., & Greenberg, S. (2002). A descriptive framework of workspace awareness for real-time groupware. *Computer-Supported Cooperative Work*, 11, 411-446.
- Heath, C., & Luff, P. (1992). Collaboration and control: Crisis management and multimedia technology in London Underground Line control rooms. *Journal of Computer-Supported Cooperative Work*, 1, 24-48.
- Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist*, 39, 1123-1134.
- Kramer, A. D., Oh, L. M., & Fussell, S. R. (2006). *Using linguistic features to measure presence in computer-mediated communication* (Research Paper 53) Retrieved from Human-Computer Interaction Institute website: <http://repository.cmu.edu/hcii/53>
- Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior*, 27, 1087-1102.
- Sangin, M., Molinari, G., Nüssli, M.-A., & Dillenbourg, P. (2011). Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in Human Behavior*, 27, 1059-1067.
- Schwind, C., Buder, J., Cress, U., & Hesse, F. W. (2012). Preference-inconsistent recommendations: An effective approach for reducing confirmation bias and stimulating divergent thinking? *Computers & Education*, 58, 787-796.
- Suthers, D., & Hundhausen, C. (2003). An empirical study of the effects of representational guidance on collaborative learning. *Journal of the Learning Sciences*, 12, 183-219.
- Walther, J. B., & Tidwell, L. C. (1995). Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. *Journal of Organizational Computing*, 5, 355-378.

Acknowledgments

This research was financed and supported by the ScienceCampus Tübingen “Informational Environments”. We would like to thank Susanne Stoll and Tim Höfling for their support.

Constructive use of authoritative sources among collaborative knowledge builders in a social science classroom

Fei-Ching Chen, Chih-Hsuan Chang, Cheng-Yu Yang, National Central University, No 300, Chung-Da Rd.,
Chung-Li City, Taoyuan County, Taiwan
fcc@cc.ncu.edu.tw, chsuchang@gmail.com, chengyu0911@gmail.com

Abstract: *Constructive use of authoritative resources* has been one of the important principles in knowledge building activities. However, how knowledge builders work together on their respective ideas and on external authoritative resources is understudied, especially in the social sciences when distinguishing advances made in the development of diverse ideas on a specific topic is a highly complex process for knowledge builders. Instead of using the conceptual inquiry thread as the unit of analysis, this study explores the responsive engagement of knowledge builders in each thread in order to reveal how achievement of deeper levels of knowledge advancement either were, or failed to be, achieved. In other words, the communal growth will be examined by evidence of authentic exchanges among knowledge builders. Results indicate that revised ideas were productive but knowledge building tended to remain incompatible with them. Possible explanations for the reluctance to incorporate idea improvement in social sciences are discussed.

Introduction

Knowledge building has long been regarded as a promising way to achieve quality learning. Although it has received much attention from science learning and teaching in the k-12 classroom for more than two decades, we know little about how it is carried out across curricula. Recently this issue has understandably been raised by Bereiter and Scardamalia in relation to the quality of learning. “Knowledge building”, particularly in the social sciences, is described by them as an approach to quality learning of conceptual content in which a depth of understanding is achieved through creating and improving explanatory theories (Bereiter & Scardamalia, 2012). In comparison with the definition of knowledge building in general (Scardamalia & Bereiter, 2003), it seems that this revised idea emphasizes a distinction between improving ideas relating to “general theories” in the natural sciences and to “theories of the case” in the social sciences. How do students in the university classroom contribute their notes on advancing theories that explain particular events and conditions? How do they adduce sources to support their case/event explanations? This study has two foci. First, we implement a principle-based design in a university social science classroom and try to identify the characteristics and potential challenges of knowledge building in different domains. Second, we are interested in how the principle of *constructive use of authoritative resources* is carried out by students in order to examine the relationship between their experience with idea improvement and the effect of this experience on their subsequent beliefs and their future approach to knowledge building.

Literature review

Advance of ideas in the social sciences

There is a period of reflection on the very idea of social science as a science modeled on the natural sciences (Flyvbjerg, 2001). The natural sciences excel at conducting decontextualized experiments to understand abstract and generalizable law-like relationships, while the social sciences conduct contextualized studies involving field research that produces intimate knowledge of localized understandings of subjective human relationships (Flyvbjerg, Landman, & Schram, 2012). Based on this distinction, we suggest that when people engage in revising ideas, one would expect to find many differences between the science classroom and the social science classroom when it comes to knowledge building.

At the school level, researchers speculate about factors related to the challenges of knowledge building in the social sciences (Bereiter & Scardamalia, 2012). The major one is the relative weakness of cognitive rewards for inquiry in comparison to those present in the natural sciences. They claim that raising the level of complexity with which students approach social issues is a more promising objective than striving for bisociative “big ideas”. Accordingly, pursuing explanations in a progressive way and producing new knowledge of value to their community is the principal work of students in knowledge building. Given that social sciences produce the kind of knowledge that grows out of intimate familiarity with practice in contextualized settings, knowledge builders will consequently contribute respective local knowledge emerging out of their own practice (Flyvbjerg, Landman, & Schram, 2012). Complexity is certainly expected but raising the discussion to a higher level becomes a new challenge.

Little empirical research has taken place on knowledge building in the social sciences. Some research projects have examined university courses in teacher education (Hong, Chen, Chai, & Chan, 2011); others have

explored psychology courses. The characteristics of threads in Knowledge Forums generated in the social sciences included lengthy notes and build-ons, multiple and diversified perspectives, a garbled set of directions and unidentifiable advances in ideas (Chen, 2012). On the one hand, some of these characteristics benefit knowledge building, but on the other hand, some inhibit it. These findings are, to some extent, consistent with what Scardamalia and Bereiter (1991) found – that the knowledge building platform lends itself nicely to divergent processes but lacks support for convergence. Following this line of inquiry, this study aims to go one step further and identify exactly how idea improvement is either achieved, or fails to be achieved, in this specific domain.

Research on constructive use of authoritative sources

Knowledge workers build on and advance the knowledge assets of their community by engaging in idea-centered discourse involving multiple perspectives, constructive criticism, progressive discourse and using a wide variety of resources (Sternberg, 2003; Bereiter & Scardamalia, 1993). In comparison to the natural sciences, it is more difficult for knowledge builders to identify authoritative sources in the social sciences than in the natural sciences. We therefore focus on the principle of *constructive use of authoritative sources* in this study.

The principle of *constructive use of authoritative sources* has been modified and augmented into “To know a discipline is to be in touch with the present state and growing edge of knowledge in the field. This requires respect for and understanding of authoritative sources, combined with a critical stance toward them.”(Scardamalia, 2002). Research has investigated how 12 principles supported quality learning in science (Zhang, Hong, Scardamalia, Teo, & Morley, 2011; Moss & Beatty, 2010; Lam & Chan, 2010). Research has also examined how this specific principle is used by knowledge builders in a PBL science learning activity (Yeo & Tan, 2010), in the understanding of optics (Zhang, Scardamalia, Lamon, Messina, & Reeve, 2007) and in a Peer Tutoring Project (Law & Wong, 2003). It is crucially important to scrutinize how constructively or how critically external sources are used by knowledge builders when it comes to the social sciences.

There are several issues which can be triggered by this principle, the first of which is achieving a balance between local and authoritative sources. Authoritative sources are traditionally considered to be books, experts (Zhang, et.al., 2007), Internet sources and teachers (Yeo & Tan, 2010). Yet external sources could go beyond them. Audio-visual sources such as YouTube and Facebook have much more powerful circulation than text-based Internet/hardcopy sources. In addition, distributed expertise plays an important role in the generation and improvement of local knowledge. It is debatable whether or not experts are more authoritative than young people themselves when dealing with adolescent affairs. If this kind of indigenous account were considered as local sources, it is obvious that local community sources have been given less attention than have external or authoritative sources.

The second issue involves the role of external sources in cycles of idea improvement (Chernobilsky, DaCosta, & Hmelo-Silver, 2004). Previous findings described the notes that contain authoritative sources and sorted them into two categories according to usage: introducing resources and going beyond resource material (Zhang, et. al., 2007). Others discovered and described instances in which authoritative sources were used to mediate the science meaning making process in a PBL activity (Yeo & Tan, 2010). Also some reported little evidence of further interpretation or of keeping a critical stance towards such materials (Law & Wong, 2003). Little to no research examined in detail the different usage of varied sources by knowledge builders and to what extent these sources produce further cycles of revised ideas.

Method

Participants

This study was conducted in a university course entitled “Adolescent Psychology” which was offered by the university’s Center of Teacher Education in Taiwan. The university is ranked as one of the best in the nation. Consequently, the students enrolled were all academically high achievers. Participants in the present study were 21 teacher-education students (14 females) who were pursuing majors in Mathematics (47%) , Physics (5%), English Literature (29%), and Chinese Literature (18%) other subjects. Thirty six percent of them were graduate students and 64 % were undergraduate students.

Principle-based design and implementation

This study employed a specific design to investigate the role of authoritative sources on idea improvement. By engaging students in this new form of pedagogy, three main instructional goals were: (1) to engage students in the revision of existing textbooks and in developing state-of-the-art knowledge about adolescent bullying; (2) to help students gain a more informed and practical understanding of knowledge building; and (3) to help students deepen the quality of asynchronous discourse via a Knowledge Forum (Scardamalia 2003). To these ends, a tutorial workshop about knowledge building theory, pedagogy and principles and how to use the Knowledge Forum for knowledge building was presented at the beginning of the semester. Focusing on the specific

principle *Constructive use of authoritative resources*, the major instructional activities included: (1) critical comments on bullying issues codified in any available textbook of adolescent psychology at the beginning and a revised paragraph related to bullying at end of this semester; (2) selected movies (i.e., *Odd girls out*) and clips on adolescent bullying introduced various types of bullying; (3) transcripts of each student's 1-hour interview with one local adolescent about their experiences and stories of bullying; and (4) most importantly, a 7-week long sustained online peer discussion about the issue of bullying.

Data sources

To address questions of how students use authoritative sources, what role they play in idea improvement, and how students changed (or did not change) their views about knowledge building and why, we collected the following sets of data: (1) students' online discourse which was recorded in a Knowledge Forum database, (2) a survey on acceptance and feasibility of knowledge building, and (3) interviews.

A multi-level analysis of discourse was performed on the recorded dataset in the Knowledge Forum. First, using notes as the unit of analysis, we were focusing on conversation structuring (Lonchamp, 2012). Notes were identified as belonging to certain categories based on the contributor who takes some aspect of the note or trace of activity of a prior contributor as being relevant for the present contributor's ongoing activity (Suthers & Desiato, 2012). In other words, in the present study, a single note was coded not only by the content of the note itself, but also by its relationship to adjacent notes. All the notes were sorted into two sets, one (set A) containing notes that involve authoritative sources, and the other (set B) containing notes without authoritative sources. By separating them and assigning different labels to them at the initial stage, we can later trace the interweaving relationships between A and B to discover whether authoritative sources inform and produce further cycles of idea improvement.

To explore the question of what kinds of authoritative sources were used by university students to develop new understanding, we examine the kinds of sources that they use as referenced in their notes. Authoritative sources can be divided conceptually into local sources and external sources (Zhang, et. al., 2007). Considering the nature of knowledge building in the field of adolescent bullying, we purposely identify information generated by *adults* as high level external authoritative sources (H), such as textbooks and Internet sources (i.e., movies). On the other hand, information generated by *adolescents* and late adolescents were identified as low level local community sources (L). This included interview transcripts with local adolescents and reflections from university students' personal experiences.

To explore the question of how different kinds of authoritative sources were used by university students, notes in set A were identified as belonging to one of three categories: A0 refers to notes that contain merely authoritative sources but do not express explicitly the contributor's claim or idea. A1 refers to notes reflecting the fact that the contributor is basically in agreement with the authoritative sources cited. A2 refers to notes that show that the contributor kept a critical stance or questioned the authoritative sources cited. Accordingly, notes in set B were further identified as belonging to 5 categories: B0 refers to an initial note of a thread or a note that does not relate to any previous one. B1 refers to notes in which the contributor is basically agreeing with the previous note. B2 refers to notes in which the contributor keeps a critical stance or raises questions toward the previous one. B3 refers to notes upon which the contributor elaborates or augment the previous ones. B4 refers to notes in which the contributor tries to conceptualize or theorize a concept based on the previous ones.

Table 1: Eight categories and their descriptions

Set	Category	Description
Notes with authoritative sources	A0	Use sources without revealing personal opinion upon them, either pro or con
	A1	Agree with the sources
	A2	Partly disagree with / or question the sources
Notes without authoritative sources	B0	Initiate a new claim /not related to the previous note
	B1	Agree with the previous note
	B2	Partly disagree with / question the previous note
	B3	Elaborate upon/augment the previous note
	B4	Draw conclusions or make inferences based upon previous notes

(A = authoritative, B = build-on)

Second, we use threads as the unit of analysis to trace both individual growth as well as communal growth on idea improvement. A2 and B2 were considered to be indicators of an attempt to revise ideas critically. The four categories, A0, A1, B0, and B1, are assumed to indicate playing a passive role in idea improvement. Alternatively, the last two categories, B3 and B4, are treated as playing a positive role in idea improvement. To evaluate idea-improvement collectively, we divided 7 weeks of discourse dataset evenly into two phases. A pair-

sample t-test was performed to examine whether there were any significant differences in each category in terms of idea improvement between the two phases. Ideally A/B 0 and A/B 1 would decrease significantly while A/B 2, B3 and B4 would increase significantly if the individuals were making progress in idea improvement.

Communal growth on idea improvement was traced by identifying “rounds of idea improvement (RII)”. We use the term “round” as a unit for counting the number of the emerging efforts on idea improvement in each thread as well as in the whole semester. Rounds are defined to be a series of adjacent notes in a thread starting with A2 and B2 and ending with A0, A1, B0, and B1. Then, if A2 and B2 re-emerge in a thread, this would be identified as a second round of idea improvement in that thread. In this way, A/B 0 and A/B 1 notes were treated as an interrupts of collective efforts toward idea improvement. In most of the chat or quasi-synchronous chat analysis, researchers generally cannot assume that a note is taking up the one before it (Suthers & Desiato, 2012). Nevertheless, there is an average lifespan of 32.3 days with an average of 4.3 notes per thread in the present study. Each thread has a much longer lifespan and therefore we assume that contributors have to some extent taken up adjacent contributions. A variation in RII of the two time phases is examined and the role of authoritative sources in RII is also be reported.

Exploring the process of idea improvement is merely one part of the story in knowledge building. Surveying the change in beliefs before and after the knowledge building activities is another. The lengthy description of each KB principle provided by Scardamalia (2002) has been divided into three sub-points based on its meaning. Questionnaires were developed using a five-point Likert scale (1 = strongly disagree; 5 = strongly agree) and containing 36 items to assess the students’ opinions on the acceptance and feasibility of 12 knowledge building principles (Chen, 2012). A pre-test was conducted after the tutorial workshop and the post-test was conducted at the end of the semester when students had finished their final assignments on a revised paragraph in the textbook.

The statistical t-test results of questionnaires will be incorporated together with interview data to see if there were pre-post changes in students’ views and why. Based on our observation notes, ten students representing heterogeneous attitudes towards this course were recruited to be our interviewees. The interview data were transcribed verbatim and used to help reveal student views on impediments and potential benefits to progress in the implementation of knowledge building.

Results & Discussion

The types and uses of authoritative sources

Types of authoritative sources: The distribution of authoritative sources in this course is shown in Table 2. Throughout the whole semester, students contributed a total of 113 threads containing 433 notes with a mean count of 20.6 notes generated per person. Due to the need to trace adjacent pairs in collective idea improvement, 21 threads containing only a single note and 17 rise-above threads lacking an original connection with previous notes were excluded. As a result, 75 threads containing 320 notes (Mean =4.3) were used as our sample and were sorted into categories. Two raters independently coded all the data and the inter-coder reliability was 0.82 ($p < .01$).

Table 2: Distribution of types and uses of authoritative sources

Type/Use	A0	A1	A2	Total
H1	2 (5.1%)	18 (46.1%)	19 (48.8%)	39 (29.5%)
H2	1 (4.5%)	13 (63.7%)	7 (31.8%)	21 (16.7%)
L1	2 (4.4%)	25 (60%)	17 (35.6%)	44 (34.1%)
L2	3 (11.5%)	17 (57.7%)	8 (30.8%)	28 (19.7%)
total	8 (6.1%)	73 (56.1%)	51 (37.8%)	132 (100%)

H1: movies; H2: text-based Internet sources & textbooks; L1: adolescent interview; L2: late adolescent experiences

The results show that the percentage of A-notes in whole data (132/320) is 41.2%. In terms of types, the teen-ager interview sources (L1) which were collected by each of the university students were the most cited ones. The adolescent movies (H1) were also much mentioned but H2 were the least referenced sources. The fact that L2 was used more frequently than H1 reveals that bullying is an authentic problem for these late adolescents. Two interesting points demanded attention. First, student use of authoritative sources in the digital age has shifted from text-based to video-based sources. Second, students preferred local and communal sources (L1&2) to external authoritative sources (H1&2). The findings of this study suggest that the conventional notion of authoritative sources, as perceived by most of the researchers in this community, should be reconsidered.

As regards the usage of these cited sources, the majority (56.1%, A1) of the notes reflect agreement with the sources, while there was substantive evidence (37.8%, A2) of notes disputing or challenging sources, based on information cited. In comparing with A1 (46.1%) and A2 (48.8%) within H1, we found that student agreement and disagreement were evenly divided on the detail in adolescent movies. Evidence shows that

movies as a kind of authoritative source were a very powerful means to motivate students' constructive use of sources.

Uses of authoritative sources: In general, notes were found to be distributed variously among the eight categories of notes (see Table 3). The percentage of all B-notes was 58.8%. Of all A and B notes, 23.4% were determined to be B2; 22.8% were A1; 19.4%, B3 and 15.9%, A2. It was surprising to discover that B2 notes occurred with greatest frequency. Going by the classification of the eight categories – A0, A1, B0, and B1 were treated as playing a passive role while A2, B2, B3, and B4 were defined as playing a positive role in idea improvement – 32.2% of notes fell into the former, passive categories while the latter active categories contained 67.8%. In other words, this class in general was intensively engaged in idea improvement activities. Nevertheless, while this tally of single notes provides a general picture of how students worked online in this database, it does not reveal much about how students actually built knowledge collaboratively with one another. To better understand the interweaving of A-notes with B-notes in their knowledge building, a series of thread analyses were performed.

Table 3: Distribution of notes on 8 categories.

	A0	A1	A2	B0	B1	B2	B3	B4	Total
# of notes	8	73	51	8	14	75	62	29	320
% of notes	2.5	22.8	15.9	2.5	4.4	23.4	19.4	9.1	100
	7	2	4	7	6	1	3	5	

The process of collective idea improvement

The collective idea improvement pattern. Pre-post comparisons were made between two phases from the early to the late stages of idea improvement. The results showed (Figure 1) a significant decrease in category A1 ($t=-2.73$, $p<.05$) and A2 ($t=-3.75$, $p<.01$) between the two stages; also there was a significant increase in category B2 ($t=2.19$, $p<.05$) and B3 ($t=2.50$, $p<.05$). It shows, at the early KB stage, that student responsive engagement tended to focus more on citing authoritative sources with either consenting or dissenting comments; while they shifted more to arguing with and elaborating upon previous notes at the latter KB stage. It seems that A-notes first played a scaffolding role in KB activities, and then these A-notes were able to elicit powerful B2 and B3 notes to improve ideas in a collective fashion.

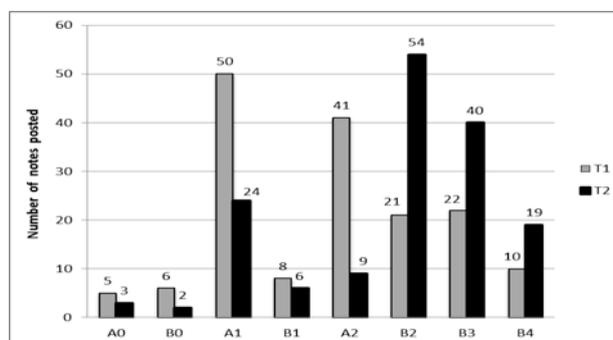


Figure 1. Changes in students' responsive engagement among 8 categories over a semester.

The rounds of idea improvement (RII): Threaded analyses were conducted to understand the interweaving relationships between A-notes and B-notes so that the effectiveness of authoritative sources on idea improvement could be examined clearly. In order to investigate the patterns of eight categories in each thread, threads with less than 5 notes were excluded from further analysis. The remaining left 28 long threads which, with each note categorized, were then made into a figure (Fig. 2) to illustrate the adjacent notes in each thread. Each line represents one thread. The order of the threads is determined by the posting time of the first note. Shading represented A-notes and Black represented B-notes. The number in each category was represented by the height of the bars with a line, single height, double height, triple height, and quadruple height representing 0, 1, 2, 3 and 4 respectively.

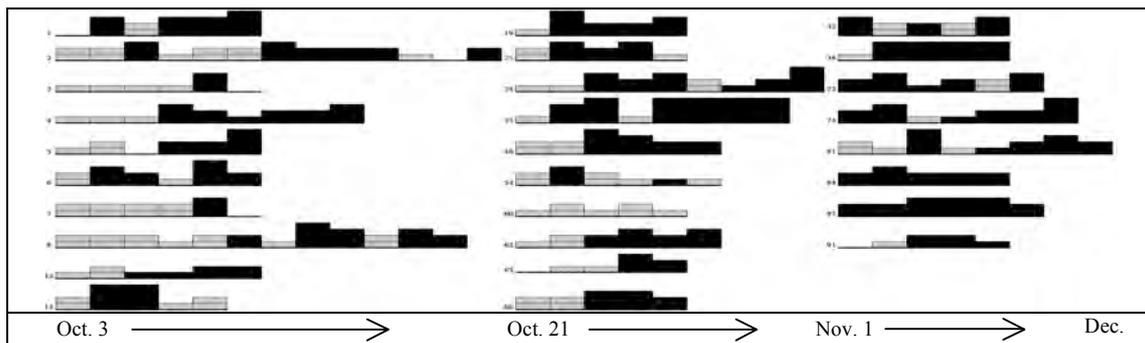


Figure 2. Rounds of idea improvement (RII) in 28 longer threads

The number of RII was further analyzed. Notes identified as either A2 or B2 were considered to be the trigger of idea improvement while A/B 0 and A/B 1 notes were treated as an interruption of idea improvement. In other words, RII ended with either an A0, B0, A1, or a B1. Then, if A2 and B2 re-emerged in a thread, it would be identified as a second RII in that thread. For example, the number of RIIs is 3 in thread #8, 2 in thread #28, etc.

Overall, there were 73 RII found in 57 threads and a total of 177 notes reflecting engagement in continuous efforts in idea improvement. On average, there were 2.4 notes within each RII. Given the average of 4.2 notes per thread in the present study, 2.4 was a fairly reasonable figure for the most important activity in knowledge building. It indicates that, with the constructive use of authoritative sources, A2 and B2 did inform adjacent notes and are shown to have had a more productive effect upon idea improvement.

On the other hand, threads without RII also deserve further attention. Eighteen threads which contained no RII had a total of 60 notes. Interestingly, half of them were B3 (23.3%) and B4 (20%). Although the threads without RII contained no indication of critical exchanges, B3 and B4 collectively achieved an even higher level of idea improvement. Moreover, examining the structure of a thread revealed more details about the students' style of engagement in knowledge building. Of the 75 threads, 45% of the initial notes were A1 and 27% of them ended with B2. This suggests that given the interweaving process of RII circles, the revised products remain divergent. It is not easy to collectively reach a convergent product.

Obviously the abovementioned multiple sources of evidence indicate that students in this class were collectively engaged very positively on idea improvement. The results so far included 1) The most frequently posted notes were B2, 2) the RII was substantive, and 3) the whole community shifted their energy from A1 and A2 to B2 and B3. However, evidence from the questionnaire suggests a different interpretation.

Student view of idea improvement in social sciences

Changes in students' perceived acceptance and feasibility of knowledge building. It was found that students tended to consider knowledge building to be both recognizable ($M = 4.4$) and feasible ($M = 3.9$) as these means were both higher than similar samples (Hong, et. al., 2011). To further understand if engaging students in knowledge building practice has an impact on their views about the acceptance and feasibility of knowledge building, t-tests were conducted. Notably, the results showed there was no significant difference between the pre- and post-tests (acceptance: $M=158.85$, $SD=12.33$ vs. $M=157.95$, $SD=12.84$, $t=0.43$, $p>.05$; feasibility: $M=142.00$, $SD=17.44$ vs. $M=141.15$, $SD=20.34$, $t=0.21$, $p>.05$). When looking into specific principles, the only significant increase in feasibility was found in principle #9, *constructive use of authoritative sources* (feasibility: $M=11.20$, $SD=2.07$ vs. $M=12.25$, $SD=1.77$, $t=2.50$, $p<.05$). Nevertheless, what is more important to know is why in general there was no significant difference after students had engaged in knowledge-building practice for a semester.

Messiness and difficulties in idea improvement in the social science. As the above findings reveal, although students had engaged in idea improvement in a fairly productive fashion for the whole semester, their perceived pre-post acceptance and feasibility remained the same. One explanation could be that their perceived acceptance and feasibility was already quite high, therefore there was little room for improvement. Another explanation could be that questionnaires are only part of the whole picture of their knowledge building experience. Our end-of-class interviews were conducted by one student in the class who later became our research assistant. It was very valuable to discover how two classmates discussed the practice of knowledge building on a collegial footing. Three concerns regarding the muddied trajectories encountered in knowledge building in social sciences arose.

Multiple foci within one lengthy note.

One of the features that differ from most Knowledge Forum databases in the natural sciences is that, in the social sciences, the length of each note tends to be much longer in comparison to those of regular discussion databases.

A: That note is too lengthy.

D: Is that right...?

A: So people couldn't find the focus. If you have three points, you need to post three notes instead of one note. I remember many notes in that thread were lengthy. Because of its length, every one read it differently and paid attention to different directions. Everyone tried to post something, the point which was sensitive to himself therefore differed.

Lengthy notes in general indicate greater effort on the part of the contributor. The knowledge builders in this class, however, felt that it was not convenient to figure out what was meant. Explanation is the main activity in the social sciences. Contributors quite often described particular bullying events and emphasized the specific contexts, conditions and details before addressing the points necessary to develop the argument. Thus, an issue emerges involving explanation-based knowledge building. When lengthy notes are replied to, a series of the build-on notes followed resulting in rather fragmented discourse. Each contributor emphasized a point relevant to the previous note, but these were often comparatively less relevant to the one that had in turn preceded it. Instead of a deeper evolution of understanding via threaded discussion, these threads of decreasing and sporadic relevance produced merely segmented and disjoint idea improvement in the long run.

No "better" theories at all.

It is difficult for social science students to consider themselves to be improving upon or surpassing the ideas of others. What are the effects of segmented criticism on knowledge builders?

C: I feel that if I reply to a note, it is just a different point of view.

D: So you think it is just a trivial opinion and makes discussion even messier. It is not necessary to post because there is no right or wrong involved.

C: Exactly. What everyone contributed is based on his/her own experience. And then these experiences after all represent just a case or two.

D: When others reply to your note, were their notes really irrelevant to yours?

C: There is something relevant. But they usually dispute about a very small part of my argument. I don't feel we are building anything, rather, I feel that by finding a loophole in my argument, it seems that he disagrees totally with this idea. If so, I don't feel we are discussing anything.

From the viewpoint of students, their real knowledge building experience in the social sciences was frustrated. They could go beyond expert knowledge and provide adolescent bullying events either from the movies or the interviews as local knowledge to develop their argument but the exchange of anecdotes often led to divergent dialogues and left many underdeveloped issues. Knowledge building simply degenerated into a game of "find the loophole".

Theories "among" the cases.

If depth of understanding is achieved through improving explanatory theories as quoted by Scardamalia (2012), how exactly can generalized theories be developed through analysis of specific cases in the social sciences?

D: It's just an exception. He is adding an exception in addition to your points. So do you think can "hit back"... a "hit back"-like response be a kind of knowledge building? People add exceptions in addition to your main argument...At this moment, what we need is not to put our knowledge together but to differentiate the difference between the general statement and a statement of exception. ...But these two are not compatible, but we have to incorporate both of them into a sort of well-developed knowledge. Do you think the co-existence of the two can be a complete knowledge building?

C: It seems that...many notes dealt with the exceptions.

D: Yes, there are many.

C: Thus viewed, it seems that these general statements were wrong. ...All are so confused.

D: The secondary supersedes the primary.

What kind of work does a social science community do in general? In progressing toward explanations of phenomena, students did endure messy situations and confusion about how to develop either general theories, theories of a specific case or theories among the cases (Flyvbjerg, Landman, & Schram, 2012). Within the scope of adolescent psychology, the students are seeking to generate a few salient claims or conclusions based on the bullying anecdotes of adolescents'. They compare cases from each other's repositories but then find no way to generate valid and valuable claims and inferences based on them. They are confused by this "messy" level of complexity in their KF discourse and consequently remain unconvinced by KB theory and disinclined to pursue them.

Conclusion

Our research acknowledges the significance of constructive use of authoritative sources but argues that traditional notions of authoritative sources should be reconsidered. New types of external sources such as movies on YouTube and community sources such as interviews with local adolescents are promising at present. We have explored the types and uses of authoritative sources by university students in idea improvement and have found that counter-arguments combined with authoritative sources scaffold well during the beginning stages; build-on notes with counter-arguments and elaborations followed in the latter stages. The unit of analysis, “Rounds of idea improvement”, was developed to measure the authentic responsiveness in a micro-analysis fashion. The RII showed productive work, while the pre-post survey revealed students’ frustration when encountering co-construction of knowledge in social sciences. There is clearly much more work that could be done to develop instructional design and to facilitate social science students in their efforts to build theories based upon cases or to produce intimate knowledge of localized and contextualized understanding of the social world.

References

- Bereiter, C., & Scardamalia, M. (1993). *Environments for collaborative knowledge building*. Toronto: Centre for Applied Cognitive Science Ontario Institute for Studies in Education.
- Bereiter, C., & Scardamalia, M. (2012). Theory building and the pursuit of understanding in history, social studies, and literature. In M. J. Lawson & J. R. Kirby (Eds.). *The quality of learning* (pp. 160-177). New York: Cambridge University Press.
- Chen, F. C. (2012, May). From interaction to co-construction: University students’ experiences in Taiwan. *Science of Learning*. Paper presented at the Global Chinese Conference on Computers in Education, Taiwan.
- Chernobilsky, E., DaCosta, M. C., & Hmelo-Silver, C. E. (2004). Learning to talk the educational psychology talk through a problem-based course. *Instructional Science*, 32(4), 319-356.
- Flyvbjerg, B. (2001) *Making social science matter: Why social science inquiry has failed and how it can succeed again*. Cambridge: Cambridge University Press.
- Flyvbjerg, B., Landman, T., & Schram, S. (2012). *Real social science: Applied phronesis*. Cambridge Press.
- Hong, H.Y., Chen, F.C., Chai, C.S., & Chan, W.C. (2011). Teacher-education students’ views about knowledge building theory and practice. *Instructional Science*, 39, (4), 467-482. doi: 10.1007/s11251-010-9143-4
- Lam, I. C. K., & Chan, C. K. K. (2010, August). Scaffolding reflective assessment for conceptual and epistemic changes among chemistry students in Hong Kong. *Summer Institute on Knowledge Building*. Paper presented at the 14th Knowledge Building Summer Institute, University of Toronto.
- Law, N., & Wong., E. (2003). Developmental trajectory in knowledge building: An investigation. In B. Wason, S. Ludvigsen & Hoppe, U. (Ed.), *Designing for Change in Networked Learning Environments* (pp. 57-66). Dordrecht: Kluwer Academic Publishers.
- Lonchamp, J. (2012). Computational analysis and mapping of ijCSCL content. *International Journal of Computer-Supported Collaborative Learning*, 7(4), 475-497.
- Moss, J., & Beatty, R. (2010) Knowledge building and mathematics learning: Shifting the responsibility for learning and engagement, *Canadian Journal of Learning Technology*, 36, (1).
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal education in a knowledge society* (pp. 67-98). Chicago: Open Court.
- Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In J. W. Guthrie (Ed.), *Encyclopedia of education* (2nd ed., pp. 1370-1373). New York, NY: Macmillan.
- Scardamalia, M., & Bereiter, Carl (1991). Higher levels of agency for children in knowledge-Building: A challenge for the design of new knowledge media. *The Journal of the Learning Sciences*, 1(1), 37-68.
- Sternberg, R. J. (2003). The development of creativity as a decision-making process. In R. K. Sawyer, V. John-Steiner, S. Moran, R. J. Sternberg, D. H. Feldman, H. Gardner .M. Csikszentmihalyi (Eds.), *Creativity and development* (pp. 91-138). New York, NY: Oxford University Press.
- Yeo, J. & Tan, S.C. (2010). Constructive use of authoritative sources in science meaning making. *International Journal of Science Education*, 32, (13), 1739-1754.
- Zhang, J., Hong, H. Y., Scardamalia, M., Teo, C. L., & Morley, E. A. (2011). Sustaining knowledge building as a principle-based innovation at an elementary school. *The Journal of the Learning Sciences*, 20(2), 262-307.
- Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of 9-and 10-year-olds. *Educational Technology Research and Development*, 55(2), 117-145.

Acknowledgments

This work was supported, in part, by National Science Counsel Grants NSC 100-2511-S-008-015-MY3, NSC 100-2511-S-008-016-MY3, NSC 101-2631-S-008-003.

Making Collective Progress Visible for Sustained Knowledge Building

Mei-Hwa Chen, Jianwei Zhang, Jiyeon Lee
University at Albany

Email: mchen@albany.edu, jzhang1@albany.edu, edujyl@gmail.com

Abstract: This paper presents the design of Idea Thread Mapper (ITM), a timeline-based collective knowledge-mapping tool that interoperates with Knowledge Forum and potentially other learning platforms. Using ITM, students engage in collaborative, metacognitive conversations to identify “juicy” topics that have emerged from their knowledge-building discourse, as their communal focuses, and review ideas addressing each focus as a line of inquiry—an idea thread. A study was conducted in a grade 3 classroom to foster sustained knowledge building through ITM-aided collaborative reflection. Analyses of online discourse, classroom videos, ITM data, and student interviews elaborate the processes and benefits of ITM-aided reflection to foster student awareness of their collective knowledge and collaborative, sustained efforts to advance it.

Introduction

Schools need to cultivate collaborative, inquiry-based practices, by which knowledge-creating communities expand society’s knowledge. Knowledge productivity in such communities is achieved through a sustained trajectory of inquiry, by which ideas are continually generated, refined, and further built upon by peers to formulate more advanced ideas and problems. This process expands the community’s collective knowledge that continually informs further initiatives (Bereiter, 2002; Dunbar, 1997; Sawyer, 2007). However, current inquiry learning programs tend to focus on relatively short inquiry activities, carried out by fixed small-groups following tasks and procedures set up by the teacher. Further research needs to test designs that foster sustained, collective trajectories of inquiry in knowledge-building communities driven by student interactive discourse and ideas. To address this need, we created a timeline-based, collective knowledge mapping tool to make the collective trajectory of inquiry visible to students. Using this tool, students engage in reflective processes and conversations to co-monitor their collective knowledge and continually advance it.

Collaborative online environments have the potential to foster continual build-on and advancement of ideas (Scardamalia & Bereiter, 2006; Stahl et al., 2006). Integrating such online spaces into classroom interaction helps to give student ideas an extended public life beyond segmented activities, so the ideas can be continually revisited, improved, and taken up by community members as input to further cognitive operations (Zhang et al., 2010; cf. Dunbar, 1997). Knowledge that grows in this shared discourse space represents a product of the community as a whole—their community knowledge (Scardamalia & Bereiter, 2006) or group cognition (Stahl, 2006)—that leverages collaborative work and personal learning. Despite the above potential support, current online environments lack effective means to represent community knowledge in extended discourse, making it difficult for students to enact collective responsibility for monitoring and advancing it. In threaded discussions, chatting, and messaging, student ideas are distributed across individual postings over time (Suthers et al., 2008; Zhang, 2009). It is difficult for students to understand the conceptual landscape of their collective work, to identify knowledge advances, and to reflect on gaps and problems. Consequently, student online discourse is often disconnected, short-threaded, ill-grounded (Guzdial et al., 2001), lacks deepening moves, and focuses on addressing teacher-assigned questions (Zhang, 2009).

Therefore, fostering reflective awareness and monitoring of community knowledge has become a focal challenge (Engelmann et al., 2009; Suthers, 2001). The literature suggests a promising strategy that focuses on engaging students in creating synthetic knowledge representations during online discourse in forms of textual summaries or concept maps (Bell, 1997; Hewitt & Woodruff, 2010; Janssen et al., 2010; Suthers et al., 2008; van Aalst & Chan, 2007; Zhang, 2010). For example, in the work of Suthers and colleagues (2008), students created evidential maps to synthesize and review theories and evidence contributed to their argumentative discourse. These maps serve as collaborative “representation guidance” to foster reflective and coherent conversations (see also van der Pol et al., 2006). However, evidential/concept maps focus on small-group discussions of specific issues and may not be applicable to discourse of larger groups over an extended period, since they soon grow too large and complex to manage (Hewitt & Woodruff, 2010; Suthers et al., 2008).

New tools need to be created to represent collective knowledge progress in extended, online discourse. Such tools need to go beyond short discourse in small-groups to capture social and cognitive interactions over longer terms and at higher social levels within a whole classroom community and even beyond. Based on our previous research (Zhang et al., 2007), we created a timeline-based collective knowledge-mapping tool: Idea Threads Mapper (ITM). An idea thread represents a line of inquiry composed of a series of conceptually related

discourse entries that address a shared principal problem, extending from the first to the last discourse entry (Zhang et al., 2007). Interoperating with Knowledge Forum (Scardamalia & Bereiter, 2006), and potentially other collaborative learning platforms, ITM helps students to review shared focal themes of inquiry, as their communal focuses, that have emerged from interactive discourse, and identify important ideas contributed over time. This paper first describes the design of the ITM software and then presents a classroom-based study that uses ITM to support student collaborative reflection for sustained knowledge building.

Instrumentation: The Design of Idea Thread Mapper

The design of ITM is rooted in a view of knowledge-building discourse as multi-level, interactive, emergent systems (Zhang et al., 2009; Zhang, 2012). Through knowledge-building discourse, members in a community continually advance their collective knowledge: the state-of-the-art understanding of their community (Bereiter, 2002). Viewed from a complex systems perspective, community knowledge represents a community-level, macro structure that emerges from micro-level interactions focusing on diverse ideas contributed by members over time. Each idea, contributed by a member to the discourse, identifies certain things/issues to be investigated, as the focus and presents thoughts and work (a theory, question, experiment) to understand the focal issue. Peer members then respond to and build on the ideas to clarify and expand the focal issue to be investigated and examine and refine the concepts generated. Such interactions among the members give rise to the formulation of core and deep issues to be explored, as the community's focuses. Ideas contributed to address each major focus form into an unfolding line of inquiry, which often intersects with other lines of inquiry that investigate the related issues in the community's knowledge space. Online tools to represent collective knowledge thus need to capture and index the core focal issues to be investigated along with the concepts developed; both are progressively expanded and deepened as students engage in sustained knowledge-building discourse.

Working as a knowledge-building community, members engage in ongoing reflection to co-monitor how their diverse ideas relate, respond to, and build on one another to form unfolding conceptual streams and address important, focal goals. Additionally, they monitor how the different lines of inquiry evolve and relate to one another, advancing their collective knowledge. Such ongoing co-monitoring of ideas across the micro (individuals, small groups focusing on specific topics) and macro levels (the community as a whole unit, and beyond) will inform continual build-on and advancement of ideas across timescales. Students connect current inquiries to past inquiries and continually identify challenges and opportunities emerging from their current work to inform future inquiry.

ITM integrates three levels (or units) of ideas in knowledge-building discourse: an *idea* contributed in a discourse entry that conveys thoughts (question, conjecture) about a focal thing to be investigated, an *idea thread*, consisting of multiple discourse entries addressing a shared focus and its sub-focuses over a time period, and a *map/network of idea threads* for a whole inquiry initiative that may build on previous initiatives from the same or other communities. Constructing idea threads based on ideas contributed, synthesizing "Journeys of Thinking" in different idea threads, and mapping out a network of idea threads for whole-class reflection help students to see the larger picture of their collective knowledge space and, more importantly, it helps them rise above individual idea contributions to build coherent, high-level conceptualizations. Displaying idea threads on a timeline, with options to zoom in and out, helps students see idea build-on connections over time across multiple months (or years). A community can publish its idea threads and maps to share their progress of knowledge building with other classroom communities from around the world. Thus, metacognitive reflection and discourse supported by ITM enables multi-level, interactive and emergent knowledge building across the boundaries of time, discourse spaces, and communities.

ITM currently interoperates with Knowledge Forum, a collaborative online knowledge-building environment (Scardamalia & Bereiter, 2006). In Knowledge Forum, students contribute ideas, in the form of notes, to different views (workspaces) and build on one another's notes to engage in knowledge-building discourse. ITM was designed using a multi-tiered web architecture and implemented in JSP/Java programming language and MySQL database management system. A hierarchical view of ITM is shown in Figure 1.

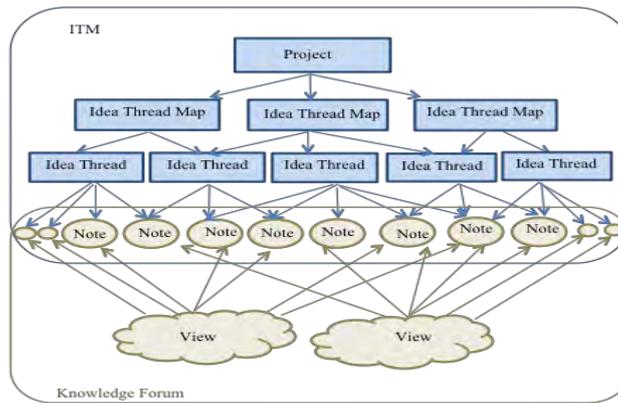


Figure 1. A hierarchical view of ITM.

ITM communicates with the servers of Knowledge Forum. On the basis of data (notes, views, users) retrieved from Knowledge Forum, ITM supports user actions to generate a map of idea threads for each inquiry initiative, as a project of inquiry. A teacher or student user can set up a project of inquiry by entering the topic of study, grade level, teacher’s name, and school’s name, and user groups involved. They then co-define communal focuses—core and deep issues to be investigated by their community in a domain area—and then search and select important Knowledge Forum notes for each focus to construct idea threads. A graphical chart is then displayed, rendering the distribution of the notes on a timeline from the first to the last note created (see Figure 2), with the options to zoom in to a specific time period (a day, a week) for a more detailed view. Each idea thread can be updated and edited by adding notes, removing notes, highlighting important notes, or renaming the focus of the thread. After reviewing the notes in an idea thread, students can summarize their “Journey of Thinking” in this line of inquiry, aided by a set of scaffold supports (e.g., *We want to understand, We used to think, we now understand, We need to read more about*). Such thread-based “Journeys of Thinking” (see Figure 3) are co-editable by all members of the classroom, with each version recorded for later review and analysis. To review collective progress in a whole inquiry project, the user can map out all the idea threads on the same timeline (see Figure 4) to examine idea progress and connection, identify productive advances, and decide on areas that need deeper work by the community.

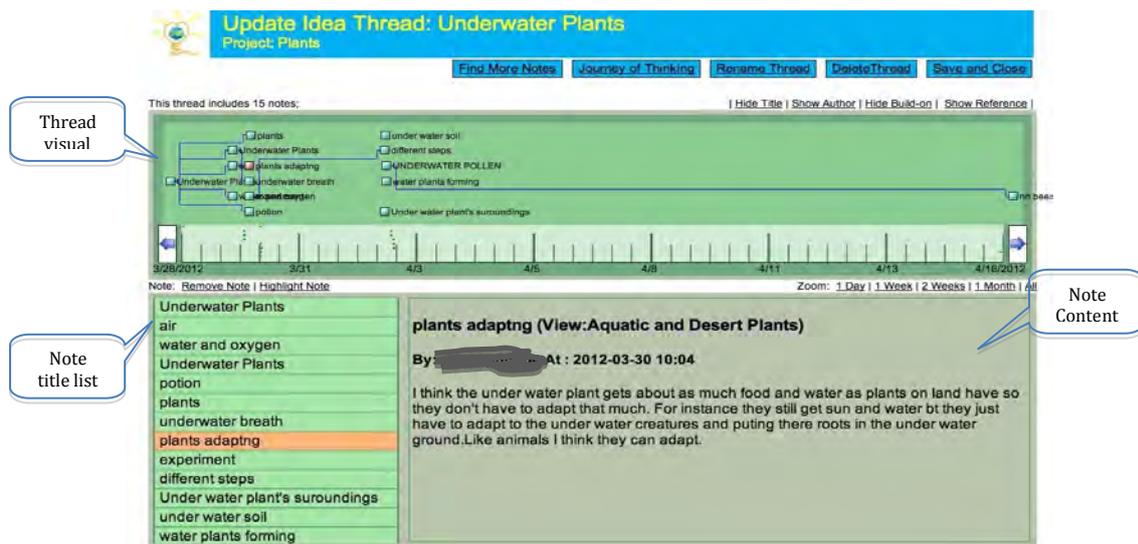


Figure 2. An idea thread on how underwater plants grow created by a grade 3 classroom studying plants. There were 15 Knowledge Forum notes addressing this problem, from March 28 to April 16. Each square in the thread visual represents a note, and a line between two notes represents a build-on link. The lower part shows the list of note titles and the content of a selected note. The user can choose to show/hide titles, authors, and build-on links and zoom into a specific time period (by day, by week, by month).

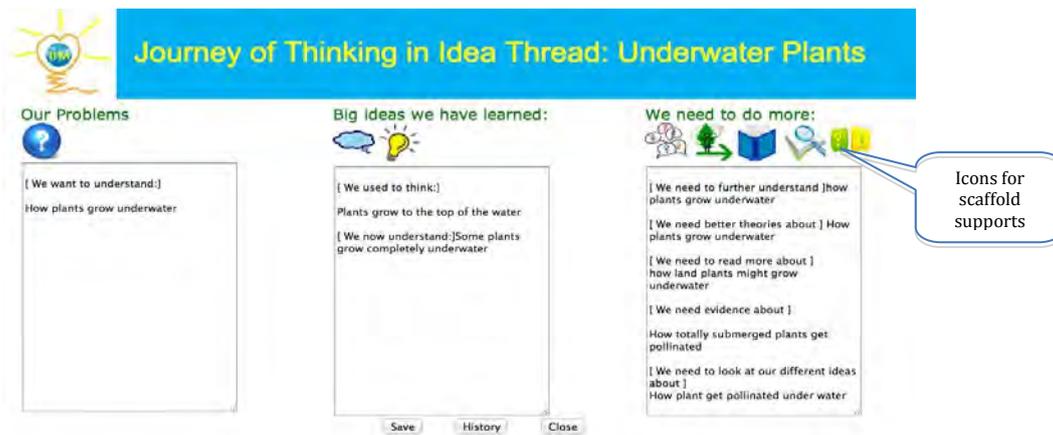


Figure 3. The “Journey of Thinking” summarized by a group of third-graders in an idea thread on underwater plants. It includes three sections: “Our problems,” “Big ideas we have learned,” and “We need to do more.” Texts in brackets (e.g., We want to understand) are scaffold supports added by clicking the corresponding icons.

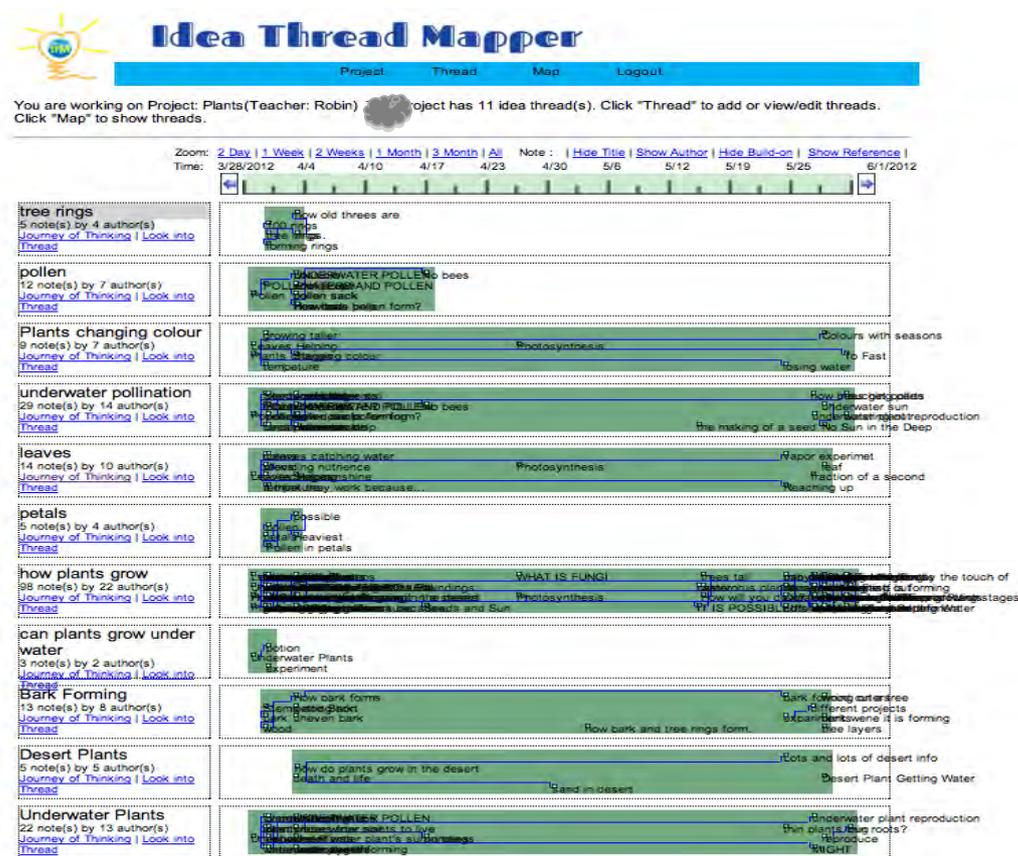


Figure 4. A map of idea threads created by a grade 3 classroom. Each stripe represents an idea thread. Each square in the threads represents a note, and a line between two notes represents a build-on link. Navigational links on the left allow the user to select an idea thread to view its details, make updates, create a “Journey of Thinking”

Classroom Research

To explore ITM-aided classroom designs to support collective, sustained knowledge building, a set of design-based studies was conducted in a grade 3 and two grade 5/6 classrooms (Zhang et al., 2013). The classroom designs focused on engaging student collaborative reflection on collective progress in their knowledge-building discourse. This paper presents our preliminary analysis of the grade 3 data. Our research question asks: In what

ways can young students benefit from ITM-aided collaborative reflection to better monitor their community's knowledge and make collaborative and sustained efforts to advance it?

Classroom Context

This design experiment was conducted in a grade 3 classroom with 22 students, taught by an experienced teacher. The students investigated plants over a two-month period. Their work integrated knowledge-building conversations, individual and group-based reading, student-designed experiments and observations, and online interactions in Knowledge Forum. Major ideas, questions, and findings generated in the classroom activities were contributed to Knowledge Forum for continual knowledge-building discourse. This design experiment involved two phases, marked by the first session of ITM-aided collaborative reflection around the middle of the unit. In the ITM session, which lasted about three lesson hours, students first worked in small-groups to identify "big ideas"—or "juicy" topics—from their knowledge-building discourse. The topics proposed were shared and discussed, resulting in a list of 11 topics. Focusing on one of the topics, underwater plants, the whole class used ITM to identify important Knowledge Forum notes and display the notes as an idea thread (see Figure 2). Temporary small-groups were then formed to construct idea threads for the rest of the topics and write a "Journey of Thinking" for each thread to reflect on the focal problems, progress of understanding, and deeper issues for further inquiry (Figure 3). The session concluded with a whole class conversation, with all the idea threads mapped out (Figure 4) and projected on a screen, to identify important knowledge advances as well as areas that required substantial deeper efforts. Focusing on the focal areas, deeper inquiry and discussions were conducted in the following month. The students then conducted another collaborative reflection session using ITM to revisit and update the idea threads. They reviewed Knowledge Forum notes created after the first ITM session, updated each idea thread to include selected important contributions, and further edited the "Journeys of Thinking" in reflection of their new insights and deeper questions.

Data Sources and Analysis

We video-recorded the ITM sessions to examine how the students engaged in collaborative reflection on their collective progress. The videos were transcribed and analyzed qualitatively to document the reflective processes. We interviewed five students before the first ITM session and five different students after the session, focusing on their awareness of themes explored by their community, knowledge advances and problems, and experience with ITM use. Complementing the video and interview data, we collected the notes taken by the small-groups that identified "juicy" topics of inquiry before the first ITM reflection. The interviews and notebooks were analyzed using content analysis (Chi, 1997) to categorize themes of knowledge advances and problems for deeper inquiry. Knowledge Forum recorded time-marked data about student online discourse, so we could analyze discourse patterns through social network analysis (Wasserman & Faust, 1994) and content analysis (Chi, 1997) to examine student collaborative deepening efforts.

Results

Student Awareness of Their Community's Knowledge

Prior to the first ITM reflection, five students were interviewed to identify important things they had investigated as a whole class. Each student identified 2.6 major themes on average (ranging from 1 to 5). The idea threads constructed by the whole class reviewed 11 themes as the community's focuses (see Figure 4), highlighting important lines of inquiry to the attention of all community members. As a student (JS) said in the interview: "it (ITM)... gives us a lot of information we didn't know." Reviewing notes in different threads and co-summarizing the "Journeys of Thinking" further helped students to understand the focal problems and advances. As student NK commented: "you can see all the notes in different areas and then ... you can look into inside all the notes to see what people put in. Then it's really cool because you can see that, wow, these people put a lot into this."

Collaborative Build-on of Ideas

As the video analysis reveals, the ITM reflection engaged students in metacognitive processes to co-construct "juicy" topics of inquiry in small-groups and as a whole class, search and review Knowledge Forum notes addressing each theme, select relevant and important notes to construct idea threads, and review all the idea threads to identify areas with major advances as well as weak areas for the community to further investigate. Doing so helped students realize how their own ideas connected to the contributions of their peers beyond those that they had explicitly built onto, to envision deeper challenges and goals for the community, and to formulate connected efforts. To gauge student collaborative efforts, we did social network analysis of the online discourse focusing on who had built on whose notes. The social network formed in the second half of the inquiry after the

first ITM session involved more intensive build-on ties, with the network density increasing from 9.74% in the first phase to 14.81% in the second phase.

Sustaining Idea Improvement

As the interview data suggest, ITM helped the students to see extended idea connection and progress through its temporal display of notes and build-on links, as well as its features to co-summarize the “Journey of Thinking” in each idea thread. When summarizing “Journeys of Thinking,” these young students could effectively identify problems of understanding as the focal goals and highlight conceptual advances achieved in their knowledge-building discourse. For example, reflecting on the development of understanding in the idea thread about petals, students wrote: “[we used to think] petals only make flowers look good. [Now we know] they help attract pollinators.” In the “Journey of Thinking” about underwater plants, it is said: “[We used to think] plants grow to the top of the water. [We now understand] some plants grow completely underwater.” In the “Journey of Thinking” regarding tree rings, students wrote “[We used to think] that tree rings help trees. [We now understand] that tree rings do not help trees but they do tell how old the tree is when you cut it down.” Focusing on the advances summarized, we traced and analyzed the notes in each idea thread and found that all the advances synthesized in the “Journeys of Thinking” were reflected in the actual Knowledge Forum discourse, often achieved through a series of discourse entries. At the end of the first ITM session, the community co-reviewed the map of idea threads to identify areas that involved productive discussions and advances as well as those that needed deeper inquiry (desert plants, underwater pollination, leaves, how plants grow in different conditions). In the last section in the “Journeys of Thinking,” students further proposed plans to investigate deeper questions in these idea threads using the scaffold supports provided by ITM. For example, the students wrote: “[We need to further understand] how plants grow.” “[We need to look at our different ideas about] how do the leaves help a plant grow.” “[We need to further understand] how underwater plants get pollinated?”

Focusing on the focal issues identified, students carried out intensive inquiry activities in the following month (May). These included problem-driven reading, experiments with underwater plants, a field trip to a wetland park, and extensive knowledge building conversations in the classroom and in Knowledge Forum. During the first week of June, the community conducted its second ITM reflection session to review new advances. Students identified 67 new notes that contributed to six idea threads, addressing issues related to how plants grow, underwater plants, desert plants, tree bark forming, leaves, and plants changing colors. As the build-on links in Figure 4 suggest, many of these new contributions in May built onto ideas contributed earlier, resulting in extended and sustained discourse. These new notes contributed more elaborate ideas than the notes created before the first ITM reflection, with the average note length increasing from 18.77 to 35.53 words per note.

For example, among the 11 idea threads, students carried out the most intensive discourse on how plants grow, a central topic that interconnects with other specific topics, such as leaves, pollen, and how plants grow underwater and in deserts. Students conducted reading, observation and dissection of plants, and rich discussions in the classroom and online. They developed initial ideas about how plants grow, as reflected in the following note:

Title: growing taller (by AF, 2012-03-30 09:43)

(My theory) the seed first attach the small roots into the ground then grows as it pushes out into the sky and drops the [actual] seed onto the ground. Then very tired it will eat the sunshine and store it into the leaves while the roots get the water and move it up the stem into the leave while the leave mixes it and turns it into food. Using the food the plant eats it and uses the energy of the food to grow taller.

Deepening their initial thoughts, students incorporated scientific concepts, such as photosynthesis, to refine their explanations. In the first ITM session, a total of 49 notes (1086 words) were selected by the students for the idea thread on how plants grow. Students further synthesized progress and identified deeper issues to be understood about how plants grow (e.g. how plants grow in different environments). The inquiry and discourse continued in this thread after the first ITM session, with the total number of notes increasing to 98 (2898 words) by the beginning of June when the second ITM session concluded. The contents of the new notes were found to be more focused and elaborated, with important progress explaining how different parts (leaves, pollen, seeds) of a plant help it grow in different environments (in deserts, underwater).

Title: underwater plant reproduction (by NK and AA, May 29 2012, 16:29:54)

Many plants are born from seeds that form after pollination, which can occur with the aid of the wind, with help from insects, and even in the water. When spring, many water plants cover themselves in flowers of striking [colors] and from each of [their] fruit more than 1,000 seeds may be freed (this is... from a book).

Title: Desert Plant Getting Water (by AD and MM, May 29 2012, 16:29:06)

A plant in the desert called [W]elwitschia mirabilis gets water from its massive leaves (2 meters.) The leaves will collect water from morning fog and channel it through the plant and into the round where the water is collected by the plants huge root. The reason that the water must go into the ground from the leaf is because it cannot process the other way by it coming through the plants leaves.

Discussion

To support collective knowledge representation and progress in extended online discourse, we created ITM as a timeline-based collective knowledge-mapping tool to make the trajectory of inquiry visible for ongoing reflection and advancement. The classroom designs, enabled by the ITM tool, engaged students in collaborative, metacognitive conversations—metadiscourse—to identify “juicy” topics and ideas from their knowledge-building discourse, review contributions, synthesize progress, and plan for deeper inquiry. Students experienced multiple levels of “rise-above” to identify major focuses from their extended discourse, to share and consolidate these focuses into a collective list of deep and “juicy” themes as the community’s focal goals, and then to create reflective syntheses of conceptual advancement, based on incremental improvements made by different members over time in different lines of inquiry.

As the results suggest, such reflective processes helped to increase student awareness of their collective knowledge, including the focal issues to be investigated and idea development achieved by the community over the two-month period. Understanding and monitoring ideas across the community’s knowledge space becomes essential in knowledge-building classrooms that encourage diverse participation, distributed expertise (Brown et al., 1993) and student-driven deepening discourse, such as the classroom in this study. Through monitoring their collective knowledge space, students learn from knowledge advances of the whole community, beyond their personal focus, and build on important thoughts of others to advance the community’s understanding (Engle & Conant, 2002; Palincsar et al., 1993; Resnick & Hall, 2001; Roschelle & Teasley, 1995). After the first ITM reflection, the third-graders in this study engaged in more extended, focused, and elaborated dialogues to address the deeper problems that they had identified, resulting in a denser network of build-on connections and refined understanding of a diverse range of interrelated issues about how plants grow. Such collaborative, reflective, and deepening efforts, leveraged by the ITM tool, are essential to fostering a sustained, collective trajectory of inquiry in knowledge-building communities.

Building on the preliminary research results presented in this paper, we are conducting design experiments to test the impact of ITM-aided reflection on sustained knowledge building within each community and further develop designs to enable cross-community build-on of ideas mediated through student-created inquiry threads and summaries. The ITM tool will be upgraded to better visualize idea build-on within different idea threads and integrate automated analysis to help students identify idea thread focuses, select important contributions, and reflect on progress.

References

- Bell, P. (1997). Using argument representations to make thinking visible for individuals and groups. In: *Proceedings of the 2nd International Conference on Computer Supported Collaborative Learning (CSCL 1997)* (pp. 10-19). Toronto: University of Toronto.
- Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Erlbaum.
- Brown, A.L., Ash, D., Rutherford, M., Nakagawa, K., Gordon, A., & Campione, J.C. (1993). Distributed expertise in the classroom. In G. Salomon (Eds.), *Distributed cognitions* (pp.188-228). New York, NY: Cambridge University Press.
- Chi, M. T. H. (1997). Quantifying qualitative analysis of verbal data: A practical guide. *Journal of the Learning Sciences*, 6, 271-315.
- Dunbar, K. (1997). How scientists think: Online creativity and conceptual change in science. In T. B. Ward, S. M. Smith & S. Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery and change* (pp. 461-493). Washington, DC: APA Press.
- Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior*, 25, 949-960.
- 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.
- Guzdial, M., Rick, J., & Kehoe, C. (2001). Beyond adoption to invention: Teacher-created collaborative activities in higher education. *Journal of the Learning Sciences*, 10, 265-279.
- Hewitt, J. & Woodruff, E. (2010, August). Bringing coherence to knowledge building discussions. Paper presented at the Knowledge Building/Knowledge Forum Summer Institute of the University of Toronto. Toronto, ON.

- Janssen, J., Gijsbert, E., Kirschner, P.A., & Kanselaar, G. (2010). Effects of representational guidance during computer-supported collaborative learning. *Instructional Science*, 38(1), 59-88.
- Palincsar, A.S., Anderson, C.W., & David, Y. (1993). Pursuing scientific literacy in the middle grades through collaborative problem solving. *Elementary School Journal*, 5, 643-658.
- Resnick, L.B., & Hall, M.W. (2001). *The principles of learning: Study tools for educators* (version 2.0) [CD-ROM]. Pittsburgh, PA: Institute for Learning, LRDC, University of Pittsburgh.
- Roschelle, J., & Teasley, S.D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69-97). New York: Springer-Verlag.
- Sawyer, R. K. (2007). *Group genius: The creative power of collaboration*. New York: Basic Books.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-115). New York: Cambridge University Press.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 409-426). New York: Cambridge University Press.
- Suthers, D. (2001). Collaborative Representations: Supporting Face to Face and Online Knowledge-building Discourse. In *Proceedings of the 34th Hawaii International Conference on the System Sciences (HICSS-34)*, January 3-6, 2001, Maui, Hawaii (CD-ROM): Institute of Electrical and Electronics Engineers, Inc. (IEEE).
- Suthers, D., Vatrupu, R., Medina, R., Joseph, S., & Dwyer, N. (2008). Beyond threaded discussion: Representational guidance in asynchronous collaborative learning environments. *Computers and Education*, 50, 1103-1127.
- 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.
- van der Pol J., Admiraal, W., & Simons, P. R. J. (2006). The affordance of anchored discussion for the collaborative processing of academic texts. *International Journal of Computer-Supported Collaborative Learning*, 1 (3), 339-357.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, Eng. and New York: Cambridge University Press.
- Zhang, J. (2009). Toward a creative social Web for learners and teachers. *Educational Researcher*, 38, 274-279.
- Zhang, J. (2010). Enhancing creative knowledge work with collaborative technologies. Presented in a symposium: "Complex Ecologies: A Panel Discussion on Educational Research and Web 2.0 Technologies." The Annual Meeting of American Educational Research Association (AERA 2010), Denver, CO.
- Zhang, J., Chen, M.-H., Chen, J., & Mico, J. (2013). Computer-supported metadiscourse to foster collective progress in knowledge-building communities. *Proceedings of the International Conference of Computer-supported Collaborative Learning (CSCL)*. Madison, WI: ISLS.
- 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.

Acknowledgments

The work presented herein was supported by a Cyberlearning grant (1122573) from the National Science Foundation (NSF), although the views expressed here do not reflect necessarily the views of NSF. We would like to thank the teachers (Robin Shaw, Ben Peebles, Julia Murray) and students at the Dr. Eric Jackman Institute of Child Study in Toronto for their valuable contributions to this research.

Designing Reference Points in Animated Classroom Stories to Support Teacher Learners' Online Discussions

Vu Minh Chieu and Patricio Herbst, University of Michigan – School of Education
610 E. University Ave., Ann Arbor – MI 48109-1259
Email: vmchieu@umich.edu, pgherbst@umich.edu

Abstract: This paper investigates how critical events or reference points in animated classroom stories can support teacher learners' online discussions about their professional practice. Research has indicated positive impact of shared artifacts or reference objects such as video records of teaching practice on the quality of teachers' conversations in both face-to-face and online discussions. Our earlier studies also show that embedding animated classroom episodes as reference objects into virtual discussion spaces can help teachers produce highly meaningful and in-depth conversations about teaching practice. Yet, all moments in a video or animation are not created equal and it is important to understand whether particular moments (reference points) included in a single reference object attract more attention than others as subjects of conversation. While this issue has been studied in the context of analyzing face-to-face conversations among teachers, we have not come across studies that examine the connection between particular reference points and the quality of online postings referring to those. This paper reports on a preliminary study that indicates promising results of how reference points can help improve the quality of teachers' online discussions. For example, teachers made more evaluative comments and proposed more alternative moves of teaching when they referred to reference points than when they did not refer to reference points. This kind of studies are important because they can help the course designer better design or select shared artifacts to facilitate and stimulate group conversations.

Introduction

Online communication tools such as chat and forum have been heavily used to sustain collaborative learning in the last 20 years, both in academic teaching (Stahl, Koschmann, & Suthers, 2006) and in professional learning (Falk & Drayton, 2009; Fishman & Davis, 2006). Obviously, communication tools themselves are not sufficient to create viable conditions that can help students learn from each other productively. Without appropriate support, learners often go off-track during their online sessions, making their discussions hardly meaningful and developed (Collison et al., 2000; Gunawardena, Lowe, & Anderson, 1997; Larson & Keiper, 2002). For example, in many non-moderated online chats and forums in which shared artifacts of the learning subject are absent, it is very difficult for members of a group to understand one another's messages, to specify their ideas, and thus to be able to develop negotiation and construction of meaning (Carroll et al., 2005; Neale, Carroll, & Rosson, 2004). Researchers of online teacher education have pointed out the important role of shared artifacts such as video records of teaching practice, as *reference objects* (Wise, Padmanabhan, & Duffy, 2009), in stimulating and facilitating learners to notice, evaluate, and reflect on critical events of classroom interactions collaboratively (Rich & Hannafin, 2009; Zhang et al., 2011). Our prior work concurs, Chieu and colleagues (2011, in review) have shown that embedding animated classroom stories into chat and forum spaces can improve teachers' online discussions about instructional practice. The presence of animations in the forum space seems to enable participants to better engage in noticing, interpreting, and evaluating important features of teaching, in proposing alternative moves of teaching, and in reflecting on their professional practice.

By reference objects we refer to the whole artifact about the learning subject that is shared and referred among members of a group during their discussion (Chieu, Aaron, & Herbst, 2013; Chieu, Kosko, & Herbst, in review). Wise, Padmanabhan, and Duffy (2009) had referred to those as reference points, but we propose to reserve this term to define another crucial concept in online interactions with continuous media artifacts: Each reference object may contain many different *points* (e.g., specific moments in a video timeline or specific locations in a picture). We believe that not all points of a reference object are equally important but rather that reference objects sometimes include key reference points that are expected to support discussion and learning. Our research aims at examining how *reference points* (e.g., critical events in a video) can help improve the quality of participants' discussions. In a preliminary study (Chieu, Aaron, & Herbst, 2013) we found a positive effect of critical events of an animated classroom story on the quality of teachers' comments when they viewed and annotated an animation individually. In this paper, we further investigate how critical events of animated classroom episodes affect the nature of teacher learners' discussions in online forums. These critical events were key in the design of the animations: As we have explained elsewhere (Herbst & Chazan, 2003; Herbst & Miyakawa, 2008) animations were designed not only to represent classroom scenarios but also to breach some instructional norms. For instance, in an American high school geometry lesson the teacher is seen asking

students to construct the proof of a claim but not providing students with statements that clearly identify the givens and the conclusion to prove, as is customary in geometry classrooms—we call this a breach of a norm of how proof tasks are assigned. Animations were designed so as to include breaches of norms since we expected this feature might prompt conversations about practice. This expectation has been met in face-to-face encounters among experienced practitioners (Herbst, Nachlieli, & Chazan, 2011). The present study explores how the quality of contributions to online discussions depended on whether contributions referred to those critical events. The study of the effectiveness of reference points thus can help validate the inclusion of breaches in the design of artifacts to support collaborative learning.

Theoretical Framework

Our research is grounded in ideas from CSCL and teacher education. Firstly, meaning negotiation and construction in social contexts are important conditions to sustain collaborative learning (Vygotsky, 1978; Engeström, 1999). Indeed, collaboration between facilitators and students and among students may enhance intellectual development of each individual as well as of the group (Cunningham, Duffy, & Knuth, 1993; Knuth & Cunningham, 1993). More specifically, students are engaged in active understanding and use of knowledge, and in learning from each other's work and errors, in particular, from feedback of other members on each individual's work and errors.

Secondly, communication and collaboration technologies have been crucial means to support teacher learning in groups (Fishman & Davis, 2006; Rich & Hannafin, 2009). For instance, Tapped In (Farooq et al., 2007) and The Math Forum (Renninger & Shumar, 2004) have used different communications tools such as email, chat, and forum to build highly interactive environments that can help teachers ask questions, discuss problems, share ideas and solutions with peers, receive feedback from mentors, and so on, in growing and active communities. Although those communities have been successful mainly thanks to features such as moderating, mentoring, and coaching, they can be improved by embedding shared artifacts or reference objects into members' virtual discussion spaces (Chieu, Herbst, & Weiss, 2011; Chieu, Kosko, & Herbst, in review).

Thirdly, representations of practice such as video records of classroom interactions have been widely and successfully used as reference objects to sustain face-to-face group discussions in teacher education and development (Rich & Hannafin, 2009; Sherin, Jacobs, & Philipp, 2011). Teachers are engaged in viewing and discussing video-based artifacts of teaching, with assistance of facilitators, so as to notice important features of teaching practice and to evaluate, interpret, and reflect on those features collaboratively (e.g., van Es & Sherin, 2008). Recently, a number of studies (e.g., Wise, Padmanabhan, & Duffy, 2009; Zhang et al., 2011) have also indicated the effectiveness of reference objects in online group discussions. For example, researchers of STELLAR (Derry et al., 2005) and eStep (Hmelo-Silver et al., 2005) have found a positive impact of a collaborative white-board that teachers can use to co-create reference objects, link resources to those objects, and discuss those objects, hence helping them share and develop their understanding collaboratively.

Our research extends the literature described earlier in two ways (Chieu, Aaron, & Herbst, 2013): (a) we use animations, instead of videos and other forms of representations of practice, as reference objects, and (b) we examine how reference points, in addition to reference objects, can help improve the quality of group discussions. The use of animations to support teacher learning is relatively new (Herbst & Chazan, 2006; Herbst & Miyakawa, 2008; Herbst et al., 2011; Moreno & Ortegado-Layne, 2008; Tettegah et al., 2008). Yet, the use of stylized cartoon characters can support the creation of representations of instructional practice very flexibly (Chieu, Kosko, & Herbst, in review). For example, we have been able to create a number of critical events in which teaching norms (Herbst & Chazan, 2003) are breached. Also, to the best of our knowledge, it seems that researchers in CSCL have not conducted studies about how reference points help improve the quality of group discussions systematically. In this paper, we examine how important events of animated classroom stories affect teachers' evaluation and reflection on their professional practice when they are engaged in a virtual setting for mathematics teacher learning that we describe below.

LessonSketch: An Interactive Rich-Media Environment for Teacher Learning

LessonSketch (www.lessonsketch.org) is a web-based learning environment that supports practice-based learning for mathematics teachers (Herbst, Aaron, & Chieu, in press). Both teachers and teacher educators can use Depict, an authoring tool in LessonSketch, to create cartoon-based representations of teaching in the form of slideshows. Teacher educators can use Plan, another authoring tool in LessonSketch, to build online experiences for their teacher learners. Online *experience* is a key notion in LessonSketch. It refers to a set of consecutive activities that engage teachers in viewing, examining, creating, or discussing rich-media representations of teaching (Chieu, Herbst, & Weiss, 2011). Teacher educators may use a library of more than 50 animated versions of 18 classroom stories in secondary geometry and algebra teaching, which our research group has produced, as well as any other streamed video material, to build their own online experiences. A key feature of Plan is the capacity to create advanced discussion spaces for end users. For example, Figure 1 shows a forum space in which an animation with full playback control is embedded to stimulate and facilitate users' discussion.

We have applied a design-based research approach (Brown, 1992; Collins, 1992) for the design, development, and evaluation of LessonSketch. That iterative and self-correcting process has helped us identify and validate a number of operational design principles and improve the tools as well as the user interface of LessonSketch (Chieu, Herbst, & Weiss, 2011; Chieu, Kosko, & Herbst, in review). The advanced forum tool presented in Figure 1 is the result of that process after several iterations. In those earlier studies, we have examined the effectiveness of reference objects and validated a key design principle: embedding shared artifacts (i.e., animated classroom episodes) directly into discussion space to support users' exchange of ideas and development of understanding. In the next sections, we report on another study to provide more evidence for that design principle. In this study, however, we emphasize the role of reference points.

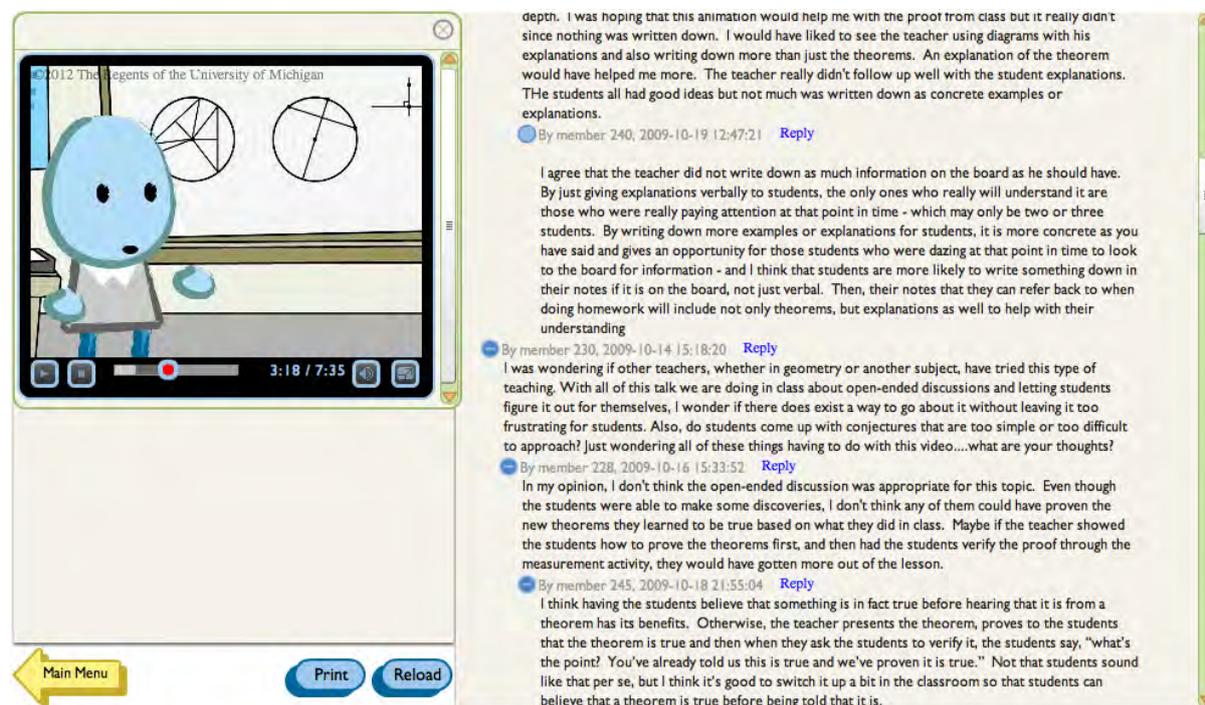


Figure 1. Advanced discussion space in which a version of “Chords and Distances” is embedded (users’ name is replaced with a system-generated id, in order to protect users’ identity).

Method

Research Questions

We use a method reported in earlier studies (Chieu et al., 2011, in review) to examine the correlations between whether and how users refer to embedded artifacts and what the qualities of their postings are. We use *Reference* (how a posting referred to the embedded animation) as an independent variable and *Evaluation* (whether a posting contains evaluative comments), *Reflection* (whether a posting contains a reflective comment), and *Alternativity* (whether a posting proposes alternative moves of teaching presented in shared artifacts) as three dependent variables. Verifying the presence of those dependent variables in users’ comments is crucial in collaborative and professional learning (Rich & Hannafin, 2009; Schön, 1983; Zhang et al., 2011).

Settings, Participants, and Procedure

We created eight online experiences for a one-semester class on geometry instruction in teacher education at a university in the East of the United States. Eleven pre-service teachers and ten beginning teachers (16 females and 5 males) enrolled in the blended course, making 21 participants in total (see Moore-Russo & Viglietti, 2011). In each experience, users first watched and commented on versions of animations individually, and then discussed those animations with peers in a non-moderated, asynchronous forum (see an example in Figure 1).

Data Collection and Analysis

We analyzed discussion logs of the stories “Chords and Distances” in Experience 3 and “The Square” in Experience 5 (watch those stories at www.lessonsketch.org). As mentioned previously, for each animated classroom episode we created a number of moments (i.e., reference points) in which one or more instructional norms are breached (see more details about how we created the animations in Herbst & Chazan, 2003). For example, Table 1 shows the main reference points of The Square.

Table 1: Critical events of The Square in which one or more instructional norms are breached.

Interval	Critical Event(s)	Breached Norm(s)
0:23 – 0:39	The teacher asks "what do you think happens with the angle bisectors of a quadrilateral?"	The teacher should pose a specific problem, which the provided one is not.
0:39 – 0:46	The teacher asks students to make conjecture and prove them.	The teacher should provide the givens and "prove" or the teacher should ask students to make conjectures only.
1:22 – 1:35	The teacher lets Alpha draw the diagram.	The teacher should provide the diagram (in doing proofs and in explorations).
1:41 – 1:46	The teacher criticizes Alpha's idea.	The teacher should encourage students who share their ideas.
1:50 – 2:01	The teacher asks the class to do a task based on Alpha's idea.	The teacher should provide the givens and prove. The teacher should choose the task that students work on.
2:05 – 2:11	The teacher repeats Alpha's words "cut the square in half."	Statements to be proven should be made using the diagrammatic register.
2:18 – 2:21	The teacher repeats Beta's words "the diagonals are also the angle bisectors."	Statements to be proven should be made using the diagrammatic register.
2:43 – 2:46	The teacher calls Gamma to the board.	The teacher should only invite a student to the board if s/he knows what the student is going to present.
2:50 – 3:17	The teacher lets Gamma generalize the square to the rectangle.	The teacher should keep the class on the task they were working on (the square).
3:27 – 3:32	The teacher repeats Gamma's point: "diagonals and angle bisectors are not the same thing."	Statements to be proven should be made using the diagrammatic register.
4:00 – 4:10	The teacher makes a strange inscription (in a \square the ang. bis \boxtimes) and asks students how would they prove something like that.	A proof problem should be stated in parsed form (Given-Prove) and using the diagrammatic register.
4:14 – 4:38	The teacher doesn't call Lambda to the board. The teacher doesn't remove one diagonal that Lambda requests.	Proofs are to be written in statements and reasons (the teacher is keeping Lambda away from any writing environment). Diagrams are supposed to have only the elements needed to do the proof.
4:50 – 4:56	The teacher doesn't enforce the diagrammatic register when Lambda uses conceptual language (isosceles triangle).	Statements in proofs are written using the diagrammatic register.
4:56 – 5:01	The teacher asks "Lambda, what are you trying to prove?"	The teacher should provide the givens and prove.
5:12 – 5:19	The teacher misunderstands what Lambda is talking about.	The teacher should enforce the students' use of the diagrammatic register when doing proofs.
5:29 – 5:45	The teacher reluctantly removes a diagonal according to Lambda's request.	The diagram should include only the elements needed for the proof.
5:56 – 6:02	The teacher doesn't correct Lambda's statements about "those triangles" and that "If you can prove congruent for one side you could prove it for the other."	Proof statements should be written in the diagrammatic register.
6:36 – 6:56	The teacher provides the givens and "prove" too late.	The teacher should provide the givens and "prove" at the beginning of the task.
6:56 – 7:10	The teacher asks for the proof of the statement after Lambda has spoken through it.	The teacher should have asked Lambda to write the proof in statement and reasons form earlier.

We consider that each forum post contains a single contribution to the discussion. So, we took the forum post as the unit of analysis. We used elements of Systemic Functional Linguistics (SFL; Halliday &

Matthiessen, 2004; Martin & Rose, 2007) to code forum posts. SFL provides the basis for an operational and valid approach with which researchers can analyze the content, context, and construction of a discourse. An important number of educational researchers have recently adopted SFL for content analysis (Schleppegrell, 2012). In our own work, SFL helped us identify, for instance, where and how users made evaluations of events in or reflections on an embedded animation (Chieu, Kosko, & Herbst, in review).

We reused the coding schemes of an earlier study (Chieu, Kosko, & Herbst, in review), which were validated by two independent coders, to assign values for *Reference*, *Evaluation*, *Reflection*, and *Alternativity*. Kappa statistics for all codes, which are greater than .45, suggest a moderate inter-rater reliability. For *Reference*, we distinguished three values, as follows: We assigned the value C when teachers made one or more references to critical events of the embedded animation (see examples in Table 1), the value S when teachers referred to one or more specific events of the animation but those events are not critical (e.g., teachers referred to a specific time code that is not presented in Table 1), and the value G otherwise (e.g., teachers referred to the animation in general). Table 2 shows an example of a reference to the critical event “The teacher doesn’t remove one diagonal that Lambda requests” at the interval 4:14-4:38 (see Table 1).

For *Evaluation*, we used markers of *affect* (indications of how teachers felt, e.g., like or dislike), *judgment* (indications of how teachers assessed characters of animations, e.g., nice or rude), and *appreciation* (indications of how teachers assessed actions in animations, e.g., engaging or boring), all of those founded on an SFL-based appraisal theory (Martin & White, 2005), to identify where teachers made evaluative comments on teaching practice. These codes helped us decide whether a post contained *Evaluation* (coded 1) or not (coded 0). The post presented in Table 2 contains *Evaluation*, because of the presence of an affect marker (like) and an appreciation marker (the teacher’s action of not changing the diagram confused some students).

Table 2: Analysis of a forum post in the discussion of The Square (we underlined pieces of the post and added codes in brackets and in capital letters right after those pieces of text).

I like [EVALUATION] your comment because [REFLECTION] I also wondered why the teacher never really made sure that the whole class understood the proof with one diagonal. At first the teacher did not even change the diagram while the student was explaining their proof [C REFERENCE]. This seemed to have confused some of the other students [EVALUATION] because [REFLECTION] the one student was talking about one diagonal and the drawing had two. I agree that the teacher should have had the students discuss why they were using one diagonal instead of two more [ALTERNATIVITY]. It may have also been useful if the teacher wrote down the steps of the proof that the one student was saying [ALTERNATIVITY]. It may have made it easier for the other students to follow.

For *Reflection*, we considered the presence of markers of manner and causal-conditional enhancements in comments on instructional practice as evidence of logical reasoning or thinking about thinking, and therefore evidence of reflection (coded by 1, otherwise by 0). Causal-conditional enhancement (e.g., if then, provided that, because) qualifies clauses through variations of logical connections. Manner enhancement (e.g., by means of, and thus) modifies clauses through comparisons or means. The post shown in Table 2 contains *Reflection* due to the presence of two markers of causal-conditional enhancement.

For *Alternativity*, we looked at participants’ proposal of teaching moves that should or could or would have been taken as alternative actions in the animation (coded by 1, otherwise by 0). We counted when participants used modals (e.g., should, could, would) in reporting actions, or potential mood (e.g., students would prefer another problem), or subjunctive mood (e.g., if the teacher had given another problem), or negative mood (e.g., the teacher did not provide a diagram to the students). We coded the previous post as 1 for *Alternativity* because of the use of modals (should) and subjunctive mood (if the teacher wrote down the steps of the proof).

An earlier study (Chieu, Kosko, & Herbst, in review) had shown that the use of different animated classroom episodes did not seem to affect the quality of teachers’ conversations over the semester. That is, the dependent variables of interest (*Alternativity*, *Reflection*) did not seem to depend on which animation was being discussed or when in the semester the discussion occurred. Thus, we aggregated the codes of all posts of the two discussions and used logistic regression models to examine the correlations between *Reference*, as an independent variable, and *Evaluation*, *Reflection*, and *Alternativity*, as three dependent variables.

Results and Discussions

Overall, we coded 89 forum postings of the two discussions. There were 36 postings referring to critical events, 11 postings referring to specific but not critical events, and 42 postings not referring to critical or specific events. Table 3 shows the difference between when participants made reference to critical events and when they did not make reference to critical events (i.e., they referred to specific or general events that are not critical). Results in Table 3 can be interpreted, for the case of evaluation, as follows: A post that did not contain reference

to critical events had a 60.3% probability of including evaluation. Yet, if a post did contain reference to critical events, then the chance of including evaluation improved to 94.4% ($p < 0.01$, effect size or odds ratio = 11.2). In other words, the odds of a post including evaluation if it contained reference to critical events was 11.2 times higher than the odds of a post including evaluation if it did not contain reference to critical events ($p < 0.01$). The other dependent variables also show similar trends.

Table 3 indicates evidence for a positive impact of reference points on the quality of teachers' discussions. In an earlier study (Chieu, Kosko, & Herbst, in review), however, we had found similar results of the impact of referring to specific (not necessarily critical) events in the embedded animation on the quality of participants' conversations. Results of the analysis of the two discussions described earlier (see Table 4) also reconfirm this claim. This could mean that making reference to specific events itself may be sufficient to improve the quality of teachers' discussions. The number of specific but not critical events was too low in these two discussions to warrant a comparison. Thus, we leave that for a future study. We note, however, that when participants made references to specific events they referred to critical events (77%) more frequently than non-critical events (23%), and the difference was statistically significant (two-tailed p -value < 0.05).

Table 3: Difference between referring to critical events and not referring to critical events.

Dependent Variables	Probability of a Post Containing the Dependent Variable When the Post Did Not Refer to Critical Events	Probability of a Post Containing the Dependent Variable When the Post Did Refer to Critical Events	P Value	Effect Size
<i>Evaluation</i>	60.3%	94.4%	$p < 0.01$	11.2
<i>Reflection</i>	60.3%	80.5%	$p < 0.05$	2.7
<i>Alternativity</i>	56.7%	86.1%	$p < 0.01$	4.8

Table 4: Difference between referring to critical or specific events and referring to general events.

Dependent Variables	Probability of a Post Containing the Dependent Variable When the Post Did Refer to General Events	Probability of a Post Containing the Dependent Variable When the Post Did Refer to Specific or Critical Events	P Value	Effect Size
<i>Evaluation</i>	52.5%	93.6%	$p < 0.01$	13.3
<i>Reflection</i>	54.7%	80.8%	$p = 0.01$	3.5
<i>Alternativity</i>	52.5%	82.9%	$p < 0.01$	4.4

Concluding Remarks

Reference objects such as video records of practice, animations, and pictures have played an important role in supporting collaborative learning. It would be even more crucial in online contexts in which it is very difficult to help group members stay focused and productive throughout a discussion session (Collison et al., 2000; Larson & Keiper, 2002; Gunawardena, Lowe, & Anderson, 1997). Obviously, not all reference objects have the same impact on the quality of group conversations. The course designer would still need to devise shared artifacts that are meaningful and useful to a group of learners. Sometimes the course designer would have to cut and select, for example, short and generative clips of long and unedited video records of practice for use in educational settings (van Es & Sherin, 2008; Zhang et al., 2011). We believe that there are important elements of a single reference object that could be subject of interest in learning sciences in general and in CSCL research in particular. Thus, we have introduced and examined another important construct: *Reference points*, as critical elements of a single reference object (Chieu, Aaron, & Herbst, 2013).

In this paper, we take critical events or moments, in which instructional norms are breached, of animated classroom episodes as reference points. We show preliminary evidence for a positive impact of critical events on the quality of teachers' conversations. For example, teachers made more evaluative or reflective comments and proposed more alternative moves of teaching when they referred to critical events than when they did not refer to critical events, which may help them produce highly meaningful and in-depth exchanges of ideas, and thus learn from one another about professional knowledge and skills. This finding would be beneficial for the course designer in the process of creating or editing shared artifacts for group discussions. In the future, it would be important to categorize reference points further (e.g., to identify different kinds of breaches) and to investigate the effectiveness of each category to inform of instructional design better. It may also help the moderator or facilitator of groups of learners better organize and sustain discussion and collaboration; for instance, during a discussion session s/he may frequently invite students to look at and talk about critical events, and sometimes ask them questions such as "What do you think of this event?" "What would you do if you were the teacher?" "Why would you do that?" so as to engage them in evaluating and

reflecting on their professional practice. Studying whether this type of intervention by the moderator or facilitator could improve the quality of group discussions further would be valuable as well.

By including in a video recording moments in which teaching norms are breached, designers can create reference points. There may have other ways to do so for the same reference object or for different kinds of reference objects, depending on the course designer's goals for the discussion session. For example, for animations or video records of practice one may include and emphasize moments in which student thinking is predominant as critical events if s/he wants teachers to develop the ability to notice and interpret student thinking. For the kind of picture media such as an image of student work, reference points could be specific locations in the image in which the student made an error or had an interesting idea.

Finally, cartoon-based artifacts such as animations and image sequences seem to have advantage over other types of media in representing practice, especially when considering the flexibility to create and edit reference points for use in educational settings. For example, it is much easier to produce animated videos that include moments in which teaching norms are breached than to find those breaches in video-recorded lessons.

References

- Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of Learning Sciences*, 2(2), 141-178.
- Carroll, J. M., Rosson, M. B., Dunlap, D., & Isenhour, P. (2005). Frameworks for sharing teaching practices. *Educational Technology & Society*, 8(3), 162-175.
- Chieu, V.M., Kosko, K.W., Herbst, P. (in review). Enhancing mathematics teachers' online discussions with animated classroom stories as reference objects. *IJCSCL*.
- Chieu, V.M., Aaron, W., Herbst, P. (2013). *Impact of critical events in an animated classroom story on teacher learners' online comments*. Paper presented at the Annual Meeting of AERA, San Francisco, CA.
- Chieu, V.M., Herbst, P., Weiss, M. (2011). Effect of an animated classroom story embedded in online discussion on helping mathematics teachers learn to notice. *Journal of the Learning Sciences*, 20(4), 589-624.
- Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15-22). Berlin, Germany: Springer-Verlag.
- Collison, G., Elbaum, B., Haavind, S., & Tinker, R. (2000). *Facilitating online learning: Effective strategies for moderators*. Madison, WI: Atwood Publishing.
- Cunningham, D.J., Duffy, T.M., & Knuth, R.A. (1993). Textbook of the future. In C. McKnight (Ed.), *Hypertext: A psychological perspective*. London, the United Kingdom: Ellis Horwood Publishing.
- Derry, S.J., Hmelo-Silver, C.E., Nagarajan, A., Chernobilsky, E., Feltovich, J., & Halfpap, B. (2005). Making a mesh of it: A STELLAR approach to teacher professional development. In T.T. May, B.M. McLaren, L. Bollen, E. Walker, A. Harrer, & J. Sewall (eds.), *Proceedings of the 5th conference on Computer-Supported Collaborative Learning* (pp. 105-114).
- Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Meittinen, & R. Punamaki (Eds.), *Perspectives on activity theory* (pp. 19-38). New York, NY: Cambridge University Press.
- Falk, J.K., & Drayton, B. (2009). *Creating and sustaining online professional learning communities*. New York, NY: Teachers College Press.
- Farooq, U., Schank, P., Harris, A., Fusco, J., & Schlager, M. (2007). Sustaining a community computing infrastructure for online teacher professional development: A case study of designing Tapped In. *Journal of Computer Supported Cooperative Work*, 16(4-5), 397-429.
- Fishman, B. J., & Davis, E. A. (2006). Teacher learning research and the learning sciences. In R. Keith Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 535-550). Cambridge, MA: Cambridge University Press.
- Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global on-line 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.
- Halliday, M.A.K., & Matthiessen, C. (2004). *An introduction to functional grammar*. London, England: Arnold-The Hodder Headline Group.
- Herbst, P., Aaron, W., & Chieu, V.M. (in press). LessonSketch: An environment for teachers to share and examine mathematical practices. In D. Polly (ed.), *Common Core Mathematics Standards and Implementing Digital Technologies*. IGI Global.
- Herbst, P., Nachlieli, T., & Chazan, D. (2011). Studying the practical rationality of mathematics teaching: What goes into "installing" a theorem in geometry? *Cognition and Instruction*, 29(2), 218-255.
- Herbst, P., & Miyakawa, T. (2008). When, how, and why prove theorems? A methodology for studying the perspective of geometry teachers. *ZDM—The International Journal of Mathematics Education*, 40, 469-486.

- Herbst, P. & Chazan, D. (2006). Producing a viable story of geometry instruction: What kind of representation calls forth teachers' practical rationality? In Alatorre, S., Cortina, J.L., Sáiz, M., and Méndez, A. (Eds.) *Proceedings of The 28th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol 2, 213-220). Mérida, México: Universidad Pedagógica Nacional.
- Herbst, P., & Chazan, D. (2003). Exploring the practical rationality of mathematics teaching through conversations about videotaped episodes. *For the Learning of Mathematics*, 23(1), 2-14.
- Hmelo-Silver, C.E., Derry, S.J., Woods, D., DelMarcelle, M., & Chernobilsky, E. (2005). From parallel play to meshed interaction: The evolution of the eSTEP system. In T.T. May, B.M. McLaren, L. Bollen, E. Walker, A. Harrer, & J. Sewall (eds.), *Proceedings of the 5th conference on Computer-Supported Collaborative Learning* (pp. 195-204).
- Knuth, R.A., & Cunningham, D.J. (1993). Tools for constructivism. In T. Duffy, J. Lowyck, & D. Jonnasen (Eds.), *Designing environments for constructive learning* (pp. 163-188). Berlin, Germany: Springer-Verlag.
- Larson, B.E., & Keiper, T.A. (2002). Classroom discussion and threaded electronic discussion: Learning in two arenas. *Contemporary Issues in Technology and Teacher Education*, 2(1). Retrieved from <http://www.citejournal.org/vol2/iss1/socialstudies/article1.cfm>.
- Martin, J.R., & Rose, D. (2007). *Working with discourse: Meaning beyond the clause* (2nd ed.). London, the United Kingdom: Continuum.
- Martin, J.R., & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. New York, NY: Palgrave MacMillan.
- Moore-Russo, D., & Viglietti, J.M. (2011). Teachers' reactions to animations as representations of geometry instruction. *ZDM—The International Journal of Mathematics Education*, 43(1), 161-173.
- Moreno, R., & Ortegado-Layne, L. (2008). Do classroom exemplars promote the application of principles in teacher education? A comparison of videos, animations and narratives. *Educational Technology Research and Development*, 56(4), 449-465.
- Neale, D. C., Carroll, J. M., & Rosson, M. B. (2004). Evaluating computer-supported cooperative work: Models and frameworks. In J. Herbsleb & G. Olson (Eds.), *Proceedings of Computer Supported Cooperative Work 2004* (pp. 112-121). New York, NY: ACM Press.
- Renninger, K. A., & Shumar, W. (2004). The centrality of culture and community to participant learning at and with The Math Forum. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 181-209). Cambridge, MA: Cambridge University Press.
- Rich, P.J., & Hannafin, M. (2009). Video annotation tools: Technologies to scaffold, structure, and transform teacher reflection. *Journal of Teacher Education*, 60(1), 52-67.
- Schleppegrell, M.J. (2012). Linguistic tools for exploring issues of equity. In B. Herbel-Eisenmann, J. Choppin, D. Wagner, & D. Pimm (Eds.), *Equity in discourse for mathematics education: Theories, practices, and policies* (pp. 109-124). New York, NY: Springer.
- Schön, D. (1983). *The reflective practitioner*. New York, NY: Basic Books.
- Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (Eds.). (2011). *Mathematics teacher noticing: Seeing through teachers' eyes*. New York, NY: Routledge.
- Stahl, G., Koschmann, T., Suthers, D.D. (2006). Computer-supported collaborative learning. In R. Keith Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 409-425). Cambridge, MA: Cambridge University Press.
- Tettegah, S., Whang, E., & Taylor, K. (2008). Narratives, virtual environments and identity semiotics: An exploration of pre-service teacher's cognitions. *Journal of E-Learning*, 5(1), 103-127.
- van Es, E.A., & Sherin, M.G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education*, 24(2), 244-276.
- Vygotsky, L.S. (1978). *Mind and society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wise, A.F., Padmanabhan, P., & Duffy, T.M. (2009). Connecting online learners with diverse local practices: The design of effective common reference points for conversation. *Distance Education*, 30(3), 317-338.
- Zhang, M., Lundeberg, M.A., Koehler, M.J., & Eberhardt, J. (2011). Understanding affordances and challenges of three types of video for teacher professional development. *Teaching and Teacher Education*, 27(2), 454-262.

Acknowledgments

Work supported by NSF grant DRL-0918425 to Patricio Herbst (PI) and Daniel Chazan (co-PI). All opinions are those of the authors and do not necessarily represent the views of the Foundation.

Identifying Gender Differences in CSCL Chat Conversations

Costin Chiru, Traian Rebedea, Stefan Trausan-Matu, “Politehnica” University of Bucharest, Department of Computer Science and Engineering, 313 Splaiul Independetei, Bucharest, Romania
Emails: costin.chiru@cs.pub.ro, traian.rebedea@cs.pub.ro, stefan.trausan@cs.pub.ro

Abstract: It is well known that there is a world-wide gender gap in most STEM domains. We propose a study of the participation of undergraduate students in computer science to a CSCL experiment in order to detect possible differences between female and male students. Moreover, we have tried to determine if the composition of the groups influences the value of the factors proposed for analyzing the activity of a student. The factors used for analysis are qualitative heuristics used to determine the activity of the users with regard to both involvement and content. Thus, we have been able to identify differences between the knowledge, innovation, involvement and vocabulary manifested by each gender group in several cases: chats only with males, with a majority of males or females and with an equal distribution of genders. The main conclusions of the research are that females are innovative and that equally distributed groups have higher scores than the others for several indicators.

Introduction

The gender gap in STEM (Science, Technology, Engineering, and Mathematics) studies and achievements has been widely debated during the last years and several measures for reducing this gap have been proposed both by governments and companies (Miyake et al., 2010). This is particularly true in computer science, where the number of females enlisted in undergraduate and graduate studies is far less than that of males especially in well-developed countries, but less pregnant in developing countries (Gharibyan & Gunsaulus, 2006). In order to reduce this gap, several educational strategies have been tried. For example, pair programming of women with other colleagues of the same gender has shown an increase in course completion rates (Werner, Hanks, & McDowell, 2004). However, there have been no initiatives that we know of to determine whether women involved in collaborative learning tasks enhance the performance of the groups they are part of or whether the group structure influences the performance of both women and men. These problems are tackled in the current paper as we have tried to identify the influence of the gender distribution for the individual and group outcomes in Computer Supported Collaborative Learning (CSCL) tasks involving online discussions.

Instant messenger (chat) is a main technological substrate for collaborative knowledge building in small groups (Stahl, 2006). Naturally, due to the massive extent of this type of applications, especially among young people, they have been used also for CSCL (Stahl, 2009). Probably one of the reasons for their success is their online character, implying the direct involvement of the participants. Another probable cause is their easy availability at any time and in any place. Eventually, the fact that participants do not have a face-to-face contact may induce easier communication, especially for shy or introvert persons.

However, the composition of groups and their diversity have an important impact on the collaboration accomplishment (Chidambaram & Carte, 2005). By considering various groups of different age, sex and work experience, they found out that groups that have either a perceived *or* an actual diversity had a “deeper and/or broader thinking about the problem” together with greater “idea exchange” (Chidambaram & Carte, 2005). However this was not observed for groups that had *both* an actual and perceived diversity. Other researches show that heterogeneous gender groups using computer-mediated communication exchange more messages to take a decision than homogenous groups (Savicki, Kelley, & Lingenfelter, 1996). An important aspect that goes beyond the current investigation is that groups with several leaders outperform groups with fewer leaders, but a correlation between leadership and gender has not been found (Carte, Chidambaram, & Becker, 2006).

In the case of chats and, in fact in the majority of CSCL applications, written language is the main communication media. Consequently, the success of CSCL sessions depends on a good and valuable written communication. If we take into account also the issue of group composition and diversity, we may conclude that we should investigate how different categories of participants use natural language in collaborative chats to determine possible differences between genders and caused by the gender distribution.

In this paper we focus on considering only one feature characterizing diversity in CSCL groups, the fact if the participant is a male or a female. The cognitive and affective particularities of each of the two groups of persons were studied in many researches in psychology and sociology (Eagly & Karau, 1991; Hyde, Fennema, & Lamon, 1990), but is almost absent in the case of CSCL.

Our analysis is starting from the hypothesis that males and females have different linguistic behaviors when talking in interaction and that the majority of one of the other category (or even the total absence of one) in the CSCL teams has important influences on the collaborative process and thus on the outcomes of the chat. In order to assess these outcomes, we have used a set of quantitative measures (Chiru et al., 2011) to measure knowledge, innovation, involvement and vocabulary for each participant and overall for each group.

We perform the analysis of collaboration in a corpus of CSCL chats starting from a theoretical framework that we developed in last years, based on discourse modeling in the case of online chats with multiple participants (Trausan-Matu & Rebedea, 2010). This conceptual basis was the starting point for several implemented systems (Rebedea et al., 2011), which combine Natural Language Processing (Jurafsky & Martin, 2009) with Social Network Analysis and considering also the Conversation Analysis basic ideas, such as adjacency pairs (Schegloff & Sacks, 1973). In the current case, we consider repetitions of words in subsequent utterances to be an importance factor of analysis. However, the analysis has been extended to also use lexical chains computed on the whole corpora of conversations using semantic distances computed on WordNet (Miller, 1995).

The paper continues with a description of the application developed for the comparison of chats with different gender distributions. The factors used to assess each individual and group outcomes are also presented in this section. The next section contains the main results of the comparative analysis that was performed on three levels: situation analysis for each type of group type, individual situation analysis and vocabulary analysis. The paper ends with the conclusions extracted from the previous analysis.

Designed Application

The purpose of the application developed for the research presented herein was to determine a model for the way different participants to CSCL conversations act and ‘talk’ in different conditions. We were interested especially in identifying if there are major differences between the participation of male and female subjects in this kind of online conversations and to decide whether the gender composition of the discussion groups may influence their outcomes.

Learning Scenario and Chat Corpus

The application was used to assess chat conversations created by senior year undergraduate students involved in a Human-Computer-Interaction (HCI) class. The students were asked to debate about different web-collaboration technologies (forums, blogs, chats, wikis, wave, etc.), highlighting the weaknesses and strengths of the existing tools and eventually devising a way to combine these tools in order to obtain an instrument that would be useful for sharing information and collaboration in a company. Each participant had to study individually the given technologies in order to identify their advantages and problems before the conversation started and afterwards they had to choose one of them and to support it in front of the other participants during the chat. The proposed scenario is similar to one of the most used macro-scripts in CSCL: the RSC (Research-Structure-Confront) script (Dillenbourg & Tchounikine, 2007). However, in our scenario the structuring stage has not been requested explicitly to the students.

The main purpose of these conversations was to ease the learning process about the considered platforms by providing each participant the possibility to critique the others’ platforms and in the same time to defend its own. This way, they were able to ‘see’ the platforms from different perspectives and had the chance to make a comparative analysis of all the considered platforms. Finally, they had to combine the existing technologies and to develop use case scenarios so that the advantages offered by each of them to be exploited to the maximum and in the same time to eliminate the identified problems related to each individual technology.

The corpus consisted of 21 different chat conversation, each of them ranging from 158 to 579 replies, for a total of 7346 utterances.

Participant Analysis

In order to obtain the corpus of chats, we started from a group of 114 students and gave them the possibility to group in smaller teams as they wanted. The distribution considering the participants’ gender was 74 males and 40 females. Five of the 74 males were involved in two different chats (were part in 2 different teams). At the same time, two females participated in two different teams. In the end, they grouped in 21 teams consisting of 4 to 11 participants, each group delivering one of the chat conversations that we analyzed (see Table 1). An important observation related to the size of the created teams is that there have been both smaller teams of 4-5 students, but also larger ones of up to 11 participants, depending on the advice the students received from their tutors regarding the advised team size: smaller teams of 4 (6 teams) or 5 (7 teams) students as compared to larger groups: 2 teams of 6, 7 or 8 participants, and 1 team of 9 or 11 participants.

The students were allowed to choose nicknames that could be used to identify them in order to be able to grade them or to have a completely anonym nicknames. Although they were assured that their (nick)names will not be made public and the usual privacy rules will be applied, some of them preferred to choose the anonymity. This is why in Table 1 there were 4 participants whose sex could not be identified based on their nicknames. These 4 cases were manually investigated by human experts and starting from their nicknames (Freaky-wiki/Wikilie, ThirdUser, Me2, BRIO), the distribution of sexes for the rest of the users, their words in the chats and their behavior, they were assumed to be males.

Table 1: Chats' length and participants' distribution

Chat No.	No. Utts.	No. Females	No. Males	No. Unknown	Total	Anonym Females	Anonym Males	Anonym + Unknown
1	203	1	3	0	4	1	1	2
2	342	3	2	0	5	3	2	5
3	289	2	3	0	5	2	3	5
4	340	1	3	1	5	1	3	5
5	311	2	2	0	4	2	2	4
6	211	1	2	2	5	1	2	5
7	397	1	2	1	4	1	0	2
8	549	1	6	0	7	0	1	1
9	347	2	7	0	9	1	1	2
10	254	3	2	0	5	1	1	2
11	158	3	3	0	6	0	1	1
12	493	0	8	0	8	0	0	0
13	541	4	4	0	8	2	0	2
14	467	2	4	0	6	1	1	2
15	203	0	5	0	5	0	0	0
16	502	7	4	0	11	2	0	2
17	463	2	5	0	7	0	0	0
18	196	1	3	0	4	1	1	2
19	579	1	3	0	4	1	3	4
20	181	3	1	0	4	3	1	4
21	320	2	3	0	5	2	3	5
All	7346	42	75	4	121	25	26	55

An interesting thing related to people's identification was that although most of the users provided enough information for establishing their gender (e.g. a first name), a lot of them chose not to provide their family name as well (this situation is defined as partially anonymised). Therefore, out of the 42 female nicknames that were identified in the chats, only 17 were providing their full names, while in the males' case 51 out of the 79 nicknames contained the participant's full name. Another interesting observation is that the most partially anonymised participants were found in small teams: 8 of the 13 teams of 4 or 5 participants consisted only on fully or partially anonymised participants, (representing 37 of the 55 anonymised participants), 4 of them had 2 partially anonymised participants and the last one had no anonymised participants. At the same time, the larger teams had at most 2 partially anonymised participants. This behavior might be explained by the need to be identified by the rest of the team: in the case of a small team, everyone knows very well the other 3 or 4 participants (so the family name is not necessary for in-group identification), while in larger teams is much more difficult to know all the members of the team well enough in order to avoid using family name. Another explanation might be given by the fact that in larger teams, there is a higher probability of having multiple persons with the same name, and therefore the family name is required for discriminate between them.

Heuristics used for participants' evaluation

In order to be able to evaluate if there are differences between the ways males and female participants act in different situations, first we had to be able to evaluate the contribution of each participant to the conversation. In order to perform this analysis, we have considered several heuristics that were previously suggested by Chiru et al. (2011):

- *Number of replies* indicating how interesting the conversation is for the considered participant;
- *Activity* of each user showing how complex one's replies are;
- *Absence* from the conversation of a participant (denoting listenership or lack of interest);
- *Persistence* of the user in the conversation, expressing the intention of a participant not to 'give the floor' to the others and computed as the number of consecutive replies issued by the participant;
- *Repetition* of other participants' concepts offering insights about how much attention a participant devoted to the content uttered by the others;
- *Usefulness* of the participant in the conversation stating how much the other participants benefited from this user's replies;
- *On topic* assessing the seriousness of the participants. It expresses how devoted the participant was to keep the conversation on the right track, considering the imposed topics for the discussion.

From the heuristics proposed by Chiru et al. (2011), we did not consider the topic rhythmicity because our purpose is somehow different from the ones of Chiru et al. (2011). We were not so interested in how often a specific topic came ‘on the floor’, but in building a model of how the participants were debating in different conditions. On the other hand, we investigated two different heuristic – *participant’s innovation*, expressing the number of concepts introduced in the conversation by each participant and *participant’s knowledge*, expressing the percent of the concepts introduced by the participant that are semantically connected with the ones imposed for debating (they are computed using a semi-automatic method of combing WordNet similarities with manual enrichment for words that have a high frequency in the analyzed corpus).

Analysis and Results

Since we were interested to see if there are differences between the ways males and females act in different situations, we considered five different scenarios: conversations with only male participants, conversations with fewer females than males, conversations with equal number of males and females, conversations where the majority of participants were females and conversations between females only. As it can be seen from Table 2, most of the chats consisted on both males and females, but with the majority of participants being males (12 chats). The lowest number of teams was observed for the conversations between participants from the same gender – only two chats between males and none involving only females.

Table 2: Distribution of chats and participants for the five classes of analyzed chats.

Distribution	Males Only	Majority Males	Equal share	Majority Females	Females Only	Total
No. Chats	2	12	3	4	0	21
Girls	0	1+2+1+1+1+1+2+2+2+1+1+2 = 17	2+3+4=9	3+3+7+3=16	0	42
Boys	8+5=13	3+3+4+4+3+6+7+4+5+3+3+3 = 48	2+3+4=9	2+2+4+1= 9	0	79

In order to analyze each class of conversations, the user has the possibility to choose one of the 5 classes or the overall statistics (see Figure 1 where the options are shown under the “Context” dropdown list).

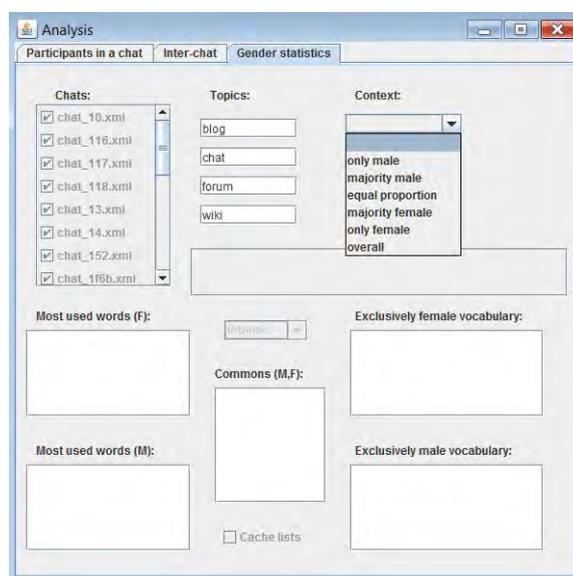


Figure 1. Front screen of the application.

Once the user has selected the desired situation, the statistics related to how male and female subjects acted in that situation are computed, and several diagrams are presented: involvement, knowledge, innovation, the percent of the vocabulary that was common and the percent of the vocabulary that is on topic. For example, see Figure 2 presenting the overall statistics. *Involvement* was computed as an average value of the scores received by the participants for the first five heuristics mentioned earlier (number of replies, activity, absence, persistence, repetition). The usefulness was used for computing how *innovative* each participant was (how many of the concepts introduced by that user were overtaken by the other participants). This heuristic, in combination with the on topic heuristic, was used to detect participants’ *knowledge*. Finally the on topic heuristic used alone showed us the participants’ *seriousness*. The last diagram - percentage of the vocabulary that was common – was used in order to see what share of females’ vocabulary is also used by males and the same thing in the case

of males' vocabulary. In the example from Figure 2, 62.21% of females' vocabulary was also used by males, while only 41.74% of the males' vocabulary was adopted by the females, the rest being used exclusively by males.

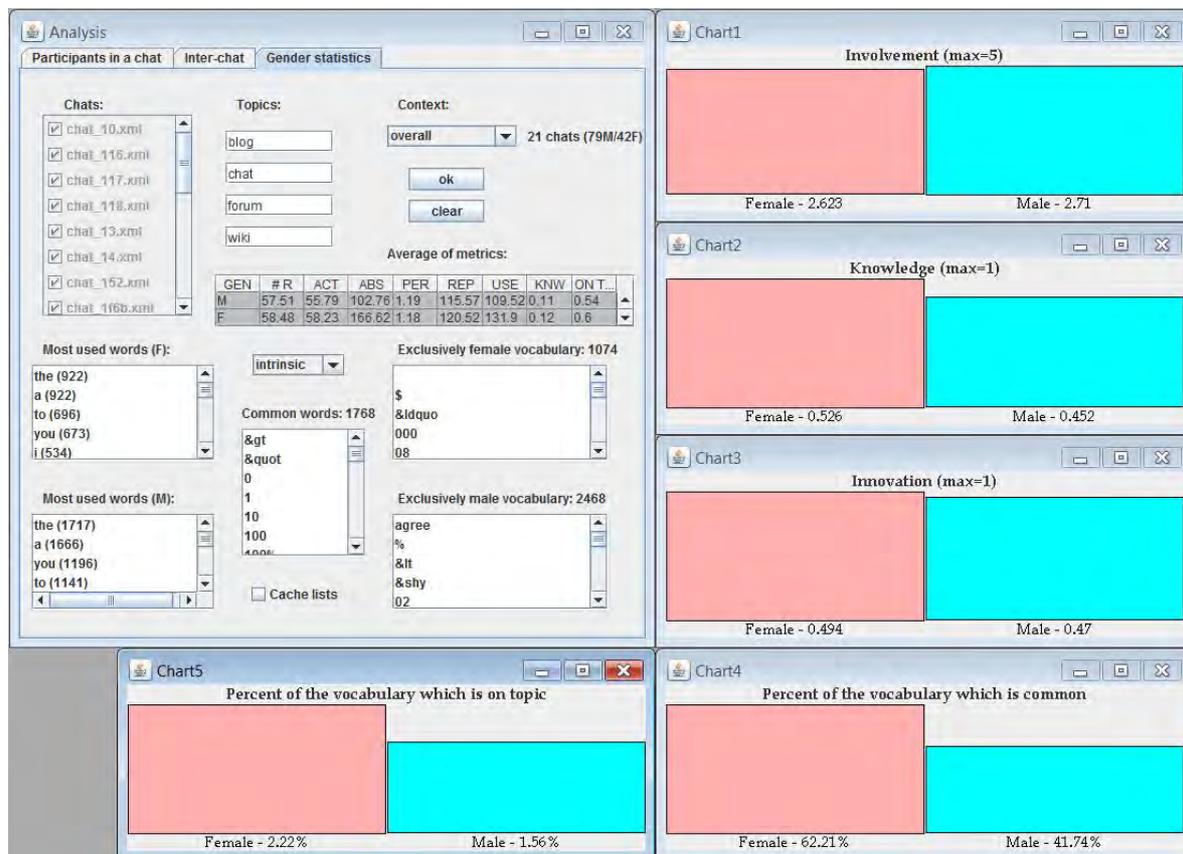


Figure 2. Application's output for the whole corpus of chats.

Since we were interested in comparing the participants acting in different situations, we have evaluated the conversations from each of the 4 different classes that we encountered (there was no chat consisting only of females) and presented the results in Table 3. Starting from these results, we can evaluate the activity of males and females in the same situations comparing to one another, or we can consider the activity of only males or females in different scenarios.

Table 3: Average scores received by the chats from the four classes identified and the overall statistics.

No chats	Situations	gen	#R	act	abs	per	rep	use	knw	on topic
21	Overall	M	57.51	55.79	102.76	1.19	115.57	109.52	0.11	0.54
		F	58.48	58.23	166.62	1.18	120.52	131.9	0.12	0.6
4	Majority Females	M	42	71.98	188.49	1.11	113.89	105.56	0.12	0.64
		F	51.69	67.27	258.13	1.13	134.94	139.62	0.14	0.65
3	Equal Share	M	53.11	50.83	108.94	1.12	98.89	84.89	0.09	0.5
		F	54.67	49.84	140.39	1.12	100.44	114.44	0.09	0.55
12	Majority Males	M	63.06	51.78	86.81	1.25	117.56	111.79	0.11	0.51
		F	66.88	54.16	94.37	1.26	117.59	133.88	0.12	0.57
2	Only Males	M	50.77	62.8	97.98	1.1	120.92	120.92	0.14	0.62

Situation Analysis

The first analysis that we did was to compare the activity of the same gender participants in different scenarios. Therefore, the males proved the least involvement in the teams formed exclusively by males (2.611) and were the most involved in the equal share teams (2.82). The males' average involvement was 2.71. Regarding the knowledge heuristic, the average was 0.452 and the most knowledgeable males were identified in the equal share teams (0.562) while the opposite was found in the case of teams formed by a female majority (0.209)

where they probably got distracted. The innovation was also influenced by the females' participation: when they were on their own, the males were the least innovative (0.43), while the most innovation from their side was seen in the case of the female majority teams (0.492), average being 0.47. Considering the seriousness, the males' words were most of the time on topic in the female majority teams (3.95%), and least of the time in chats where there were also females, but males represented the majority (1.87%).

In the case of females, they were involved more in the cases where they did not represent the participants' majority (2.77 – equal share, 2.76 – majority males, 2.38 – majority females, 2.62 – average). They showed the highest levels of knowledge when the genders distribution was not equal (0.554 – majority males, 0.512 – majority females, 0.499 – equal share and 0.526 – average). Considering their innovation, the more females were in the team, the more innovative they were: 0.545 – majority females, 0.465 – equal share, 0.461 – majority males and 0.494 – average. Finally, the females were the least serious when they represented the majority in the chat (2.93), being most serious in the case where gender distribution was equal (3.81).

Individual Situation Analysis

The second analysis regarded the comparative behavior of males and females in each of the situations. Since in the first case there are only males, we considered only the remaining 3 cases (plus the overall statistics): majority males, equal share and majority females.

The case with majority males was characterized by very small differences in the way males and females acted considering the first five heuristics representing the participants' involvement. The females uttered a little more content (both in terms of number of utterances and of number of characters/reply). On the other hand, they were missing longer periods of time from the conversation. In the same time, they were more persistent and used slightly more repetitions. Therefore, there is almost a balance between the males' and females' involvement, with a small advantage for the females. But once we get to the qualitative evaluation, we see that there is a big difference between males and females, the latter being more knowledgeable and more serious, trying much more to keep the conversation on the right track. The males stand out only for innovation.

The second situation considered was the one when the number of males and females was equal. In this case, the females uttered the most utterances, but the males' ones were more elaborate. Again, the females were absent more time than the males from the conversations and when they 'spoke', they used more repetitions than males. In this situation, the males and females seemed to have similar knowledge, but the latter proved to be more innovative. They were also more serious than the males.

The last situation analyzed was the case when the females represented the majority in the chat. In this case it can be observed the largest gap between the average number of utterances introduced by females (51.7) and males (42). On the other hand, the males introduced much more content, using utterances much more elaborated. As in the other cases, the females were more absent than the males and used more repetitions. In this case the females proved to be more knowledgeable and innovative than the males (this case having the largest gap in both statistics between the two genders), but this time the male and female participants had almost the same score in the case of seriousness.

Finally, the overall situation presents a general statistics of all the chats that have been evaluated. This statistic is somehow biased to favor males, since there was no chat having only females as participants, while the opposite situation was found in 2 chats. Still, the results show that most of the time females seemed to be more communicative than the males (both in terms of average number of utterances and number of characters), but they were also absent for larger periods of time from the chat. They used more repetitions and proved to be more useful in the conversation. Their uttered content showed they were more knowledgeable than males, more innovative and they also proved to be more serious, trying to use words that were on topic more often than the males. Still, the males seemed to be more involved, the only heuristic where they outperformed the females.

Vocabulary Analysis

Considering the vocabularies of the 4 individual classes, along with the overall statistics, we can extract several data: the vocabulary size of the whole corpus was 5310, 2468 of the words being used exclusively by males and 1074 words being exclusively used by females (see Table 4). Most of the vocabulary was generated in the chats with majority male participants, which is normal, since this category had the most chats. If we consider the normalized vocabulary with respect to the number of chats, we can observe a mean of 252.85 words, with the highest value obtained in the case of chats between males only – 619.5, the next value being obtained in the case of equal share of males and females – 524.66. If we consider the normalized vocabulary with respect to the number of participants, the average was 25.57 words exclusively used by females and 31.24 words exclusively used by males. The average number of words used by both males and females was 14.61, but it is influenced by the fact that in the case of only males chats, the size of common used words was 0 (since there were no female participants). Most common used words are observed in the case of equal share of males and females.

Table 4: Vocabulary sizes on each of the four classes identified and the overall statistics.

Situations	Females Voc.	Males Voc.	Common Vocabulary	Vocabulary size	Norm. Females	Norm. Males	Norm. Common	Norm. Voc.
Overall	1074	2468	1768	5310	25.57	31.24	14.61	252.85
Majority females	876	440	624	1940	54.75	48.89	24.96	485
Equal share	512	550	512	1574	56.88	61.11	28.44	524.67
Majority males	604	2111	1200	3915	35.52	43.98	18.46	326.25
Only males	0	1239	0	1239	-	95.3	-	619.5

Considering each of the possible situations, we can observe that in the case of conversations with male majority, 66.52% of the female vocabulary was also used by males, while only 36.24% of the male vocabulary was used by females. These values are very close to 50% in the case of chats with equal share of males and females (50% in the case of females and 48.21% in the males' case). Finally, in the last situation (chats with female majority), the percentage of the female vocabulary that is common is 41.6%, while in the case of males is 58.65%. If we are interested in the quantity of the males' and females' vocabulary that is on topic, these values are 3.15% in the case of males-only conversations, 2.94% for females and 1.87% for males in the case of males dominated chats, 3.81% for females and 3.48% for males in the case of equal share of males and females and 2.93% for females and 3.95% for males in the case of females dominated chats. The average values for the share of on topic words that was seen in males' and females' words were 1.56% for the males and 2.22% for the females. Next, we present the top 30 words used by the males and females in each of the situations in Table 5.

Table 5: Comparison of the most frequent words appearing in the four studied cases and overall.

Only males	F	-
	M	you (195), i (132), can (130), that (95), wiki (89), chat (87), be (75), but (68), if (64), this (61), think (50), information (50), forum (49), have (48), blog (46), we (44), blogs (43)
Mainly males	F	you (304), i (216), that (180), are (126), can (121), but (110), chat (104), have (99), this (86), wiki (67), be (66), all (61), one (60), forums (59)
	M	you (740), that (569), i (556), for (432), can (325), are (296), be (283), we (262), chat (253), have (243), wiki (195), all (171), think (170), forum (165)
Equal share	F	you (93), i (74), for (61), can (57), that (52), be (47), have (40), chat (38), are (37), we (35), blogs (34), all (30), people (30), they (29), forums (28)
	M	you (108), i (83), that (68), for (60), are (57), can (53), chat (44), be (42), all (39), have (39), this (34), we (27), forums (27), good (26), one (24)
Mainly females	F	you (250), i (230), that (160), can (147), for (118), be (117), are (112), blog (109), wiki (98), chat (91), forum (80), have (76), we (76), blogs (70), yes (68), information (68)
	M	you (107), i (87), that (75), can (61), are (54), be (48), chat (43), blogs (38), forums (34), blog (32), one (32), think (31), have (28), wiki (28), your (28)
Overall	F	you (673), i (534), that (406), can (356), for (328), are (287), be (245), chat (238), have (225), blog (209), wiki (183), we (179), blogs (170), all (162), forum (160), this (155)
	M	you (1196), i (881), that (835), for (665), can (599), are (508), be (461), chat (434), we (364), have (363), wiki (337), this (282), think (275), blogs (274), all (271)

As it was expected, most of the top frequency words are determiners, prepositions, pronouns, negations and auxiliary verbs (some of them have been excluded from Table 5). Still, the main topics that were imposed for debating can be found in these top frequency words. For example in the male-only chats, we find between these words concepts such wiki, chat, information, forum, blog, blogs. In the chats that were dominated by males we see the words chat, wiki, forums being amongst the 30 most frequent words by females and chat, wiki, forum in the top 30 most frequent words by males. The equal share chats proposes words such as chat, blogs, people, forums for females most frequent words, while for males we can see the words: chat, forum, good. The last category seems the one that was the best from the content point of view, with words such as blog, wiki, chat, forum, blogs, information for females and chat, wiki, blogs for males. A similar situation can be observed in the overall statistics where chat, blog, wiki, blogs, forum were amongst the most frequent words used by females, while chat, wiki, blogs were amongst the top frequent words used by males.

Conclusions

In conclusion, gender-unbalanced chat teams decrease participants' involvement – the more balanced the teams are, the more involved are the participants. In the case of knowledge showed by the participants, they seemed to be opposite sides: the males proved to be more knowledgeable in gender-balanced teams and in teams composed of only males. At the same time, the females acted worst in the case of equal shares of males and females and acted much better in the other two remaining cases (where males seemed to be the least knowledgeable).

Regarding innovation, it seems that the best situation is to have teams where females represent the majority, since in this situation both males and females are the most innovative. Considering seriousness, the males and females acted different again, males being the most serious when the females were the least (conversations with majority females). Still, a good tradeoff seems to be given by the situation when we have equal shares of males and females, since in this case the women seemed to be the most serious while for the men this was the second best case.

A definite overall solution to the best way for gender distribution in CSCL chat tasks for small groups is difficult to reach. However, the presented results suggest that women tend to be more innovative, while men appear to discuss more on topic when they are in heterogeneous groups (only males, majority females). For each individual CSCL task, teams should be composed taking into account these factors as, in most cases, gender distribution influences the overall performance of the participants to the task.

Acknowledgments

This research presented in this paper was supported by project No.264207, ERRIC-Empowering Romanian Research on Intelligent Information Technologies (FP7-REGPOT-2010-1).

References

- Carte, T., Chidambaram, L., & Becker, A. (2006). Emergent Leadership in Self-Managed Virtual Teams. *Group Decision and Negotiation*, 15(4), 323-343. doi: 10.1007/s10726-006-9045-7
- Chidambaram, L., & Carte, T. (2005). *Diversity: Is There More Than Meets The Eye? A Longitudinal Study of the Impact of Technology Support on Teams with Differing Diversity*. Paper presented at the Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 2005. HICSS '05.
- Chiru, C., Cojocaru, V., Trausan-Matu, S., Rebedea, T. & Mihaila, D. (2011). Repetition and Rhythmicity Based Assessment for Chat Conversation. *ISMIS 2011, LNCS 6804*, pp 513-523.
- Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro-scripts for CSCL. *Journal of Computer Assisted Learning*, 23(1), 1-13.
- Eagly, A. H., & Karau, S. J. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology*, 60(5), 685-710.
- Gharibyan, H., & Gunsaulus, S. (2006). *Gender gap in computer science does not exist in one former soviet republic: results of a study*. Paper presented at the Proceedings of the 11th annual SIGCSE conference on Innovation and technology in computer science education, Bologna, Italy.
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107(2), 139-155.
- Jurafsky, D., & Martin, J. H. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*: Pearson Prentice Hall.
- Miller, G. A. (1995). WordNet: a lexical database for English. *Commun. ACM*, 38(11), 39-41. doi: 10.1145/219717.219748
- Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the Gender Achievement Gap in College Science: A Classroom Study of Values Affirmation. *Science*, 330(6008), 1234-1237.
- Rebedea, T., Dascalu, M., Trausan-Matu, S., Armitt, G. & Chiru, C. (2011). Automatic Assessment of Collaborative Chat Conversations with PolyCAFe. *EC-TEL 2011, LNCS 6964*, pp. 299-312.
- Savicki, V., Kelley, M., & Lingenfelter, D. (1996). Gender and Group Composition in Small Task Groups using Computer-Mediated Communication. *Computers in Human Behavior*, 12(2), 209-224.
- Schegloff, E. A., & Sacks, H. (1973). Opening Up Closings. *Semiotica*, VIII(4), 289-327.
- Stahl, G. (2006). *Group cognition. Computer support for building collaborative knowledge*. Cambridge: MIT Press.
- Stahl, G. (2009). *Studying Virtual Math Teams*. New York: Springer.
- Trausan-Matu, S. & Rebedea, T. (2010). A Polyphonic Model and System for Inter-animation Analysis in Chat Conversations with Multiple Participants. *CICLing 2010, LNCS 6008*, Springer, 2010, pp. 354-363
- Werner, L. L., Hanks, B., & McDowell, C. (2004). Pair-programming helps female computer science students. *J. Educ. Resour. Comput.*, 4(1), 4. doi: 10.1145/1060071.1060075

The Impact of CSCL Beyond the Online Environment

Sherice N. Clarke, Gaowei Chen, Catherine Stainton, Sandra Katz, James G. Greeno, Lauren B. Resnick,
University of Pittsburgh, Pittsburgh, USA
sclarke@pitt.edu, gac28@pitt.edu, stainton@pitt.edu, katz@pitt.edu, jimgrno@pitt.edu, resnick@pitt.edu
Gregory Dyke, Université de Lyon, Lyon, France, gregory.dyke@ens-lyon.fr
Iris Howley, David Adamson, Carolyn P. Rosé, Carnegie Mellon University, Pittsburgh, USA
iris@cmu.edu, dadamson@cs.cmu.edu, cprose@cs.cmu.edu

Abstract: Accountable Talk is a form of classroom interaction that positions students as thinkers in interaction and encourages students to make their thinking visible for collaborative reasoning. This paper reports on a two-year teacher professional development program in which teachers were coached to use Accountable Talk practices in their classrooms. Online collaborative learning activities were used to prepare students for these whole-class, teacher lead discussions using the same paradigm. Findings from a series of studies embedded within the two year professional development program provides evidence that novel conversational agent designs, based on the Accountable Talk approach to discussion facilitation, improve learning during the online exercises and better prepare students to benefit from whole class discussions. In this paper we evaluate the effect on teacher uptake of Accountable Talk practices when their students have participated in these online small group activities.

Introduction

The history of CSCL has seen major advances in the theory of what makes collaborative learning beneficial, and the positive impact of technology for producing participation in those environments. New technologies enable monitoring of collaborative processes in real time and adapting support to the changing needs of groups of learners. With this emerging technology, CSCL technology now has the potential, as never before, to effect positive impact beyond the online context, and beyond the lessons in which it is used. In this paper we take one step in that direction by evaluating the use of CSCL as embedded within the context of an intervention targeted towards supporting teachers' development of what we refer to as Accountable Talk. We explore how using CSCL activities with students may facilitate the teacher's ability to take up productive discussion facilitation practices in whole group, face to face, teacher led discussions that follow the online experiences.

In this paper we first review the evidence on the use of Accountable Talk as an instructional tool to support student learning. Second, we illustrate the use of intelligent conversational agents designed based on this concept as dynamic support for collaborative learning in a CSCL environment. We review results from studies that demonstrate positive impact on student learning from participation in collaborative activities facilitated by intelligent conversational agents. Next, we introduce a two-year teacher professional development design study in an urban school district, in which CSCL activities supported by these Accountable Talk agents were integrated into the program. Finally, we evaluate the effect of inclusion of these CSCL activities on teacher uptake of the target discourse practices. We conclude with discussion and continued work.

Accountable Talk in the Classroom

Consistent with the literature on facilitation of collaborative learning groups (e.g., Hmelo-Silver & Barrows, 2006), a large body of work has shown that certain forms of classroom interaction, what we refer to here as Accountable Talk, are beneficial for student learning (Resnick, Asterhan & Clarke, in press). This literature shows that Accountable Talk is characterized by high demand tasks in which teachers help to scaffold student thinking and reasoning about subject-matter through talk. Well-structured teacher-lead discussions elicit student ideas through discourse moves that help to make student thinking visible and an object for whole class thinking and reasoning (e.g. Table 2). These discourse practices have been shown to increase student learning and reasoning, long-term retention and transfer across subject matter (Adey & Shayer, 1993; Bill, Leer, Reams & Resnick, 1992; Chapin & O'Connor, 2004, Resnick, Salmon, Zeitz, Wathen & Holowchak, 1993; Topping & Trickey, 2007a, 2007b; Wegerif, Mercer & Dawes, 1999).

Greeno argues that discourse structure positions individuals intellectually within the dialogue (Greeno, in press). Thus we make a distinction between structures where teacher questioning positions individuals as passive versus cognitively engaged with the subject matter, e.g. Table 1. One limitation of prior work has been that while there is strong evidence for the benefits of discursive teaching practices, the research has primarily been limited to special populations, including expert teachers. Other forms of dialogue have tended to dominate mainstream schooling, especially urban schools, which are the concern of this paper. In particular, discursive styles often position students to passively engage with the subject matter. For example, Table 1 helps to make

the distinction between teacher moves and the intellectual work these moves elicit. The style of teaching in those contexts is mainly monologic Initiation-Response-Evaluation (IRE).

Table 1. Comparison of discourse structure and intellectual positioning

Initiation-Response-Evaluation (Mehan, 1979)	Accountable Talk
<i>Jonathan, what is the answer to number 7... Correct</i>	<i>Jonathan, how might we arrive at an answer to number 7... so you mean... Elizabeth do you agree with Jonathan's reasoning?</i>

With rising student-to-teacher ratios in urban classrooms, the classroom discourse community has begun to question how discussions in classrooms can be academically productive, particularly if we wish to use such situations to develop reasoning skills. Efforts to support uptake of productive discussion facilitation practices in these environments has become a growing concern. The challenge for our research has been: *how do we make Accountable Talk* a widespread instructional practice across populations and contexts, in particular, non-expert teachers in urban schools.

We have been conducting a longitudinal design study of Accountable Talk, which we report on here, in an urban school district that has failed to meet national standards for achievement on standardized tests for 5+ years, and whose students were 63% below proficient in Reading and 56% below proficient in Math, a large percentage of which are African American students. The intervention has been targeted towards developing teachers' use of Accountable Talk discourse moves (Table 2) in the context of high school biology lessons to support student reasoning about biology. Table 2 illustrates the set of discussion facilitation moves that are the target of the professional development training, which are based on prior research (Resnick, Michaels & O'Connor, 2010). This same list of facilitation moves is used in the transcript analysis discussed later.

Table 2. Target Accountable Talk Teacher Moves

Speaker	Utterance	Accountable Talk Moves
TEACHER	Explain your thinking.	SAY MORE
TEACHER	They were all adopted?	REVOICE
TEACHER	What do you think Desmond? Can you repeat what she said?	RESTATE
TEACHER	What's it prove? Put it into words.	PRESS FOR REASONING
TEACHER	If capital 'G's dominant, wouldn't all babies be orange?	CHALLENGE
TEACHER	Kelly, are they right?	AGREE/DISAGREE
TEACHER	Help him out Stephen.	ADD MORE
TEACHER	So then put it in your own words. Explain why she's right or wrong.	EXPLAIN OTHER

Intelligent Agent Support for Accountable Talk

A key contribution of this paper is a demonstration of the impact of CSCL activities on teacher uptake of Accountable Talk practices. The CSCL activities used in this study included dynamic support provided by intelligent conversational agent computer programs. Conversational agents have a long history of successful support for individual learning with technology (Rosé et al., 2001). A series of results offer hope that they can be used productively to offer support for collaborative learning, especially in chat environments (Kumar & Rosé, 2011). The agents used in this prior work interacted with students through multi-step directed lines of reasoning focused on specific conceptual content rather than focusing on the style of interaction between the students. In our work, we have been developing conversational agents with a different style. In particular, our goal has been to pattern the behavior of the agents after Accountable Talk facilitation moves. In our work to date we have focused on two Accountable Talk facilitation moves in particular, namely Revoicing and Agree-Disagree. These moves in particular have two important functions. First, they mark a students' contribution, a social move that suggests that their contribution matters. Second, they prompt the contributor and other students to operate on the idea, an intellectual positioning move. These kinds of agent discourse moves might be followed by a series of student moves and reasoning behaviors such as reflecting on the idea, modifying, elaborating, or retracting it.

For both types of facilitation moves, the agent behavior in the discussions in the CSCL environment was implemented by programming the agent to compare student input against a list of correct statements drawn from the data collected during pilot runs of the system. If an entry in this list could be interpreted as a paraphrase of the student's input based on analysis of word overlap, it was offered by the agent as a "revoice" to the students, or alternatively, a prompt was offered for the other group members to discuss whether they agree

or disagree with the statement. Some examples are given in Table 3. The same statement was never offered more than once in the same session as a revoicing. When student statements were not close enough to match the revoicing list but contained the first mention of important lesson concepts (like "indicator" or "molecule size"), the agent would nudge the student or a peer to expand or restate their contribution.

Table 3: Examples of Revoicing Agent Behavior

Student Contribution	Revoicing Agent Response
basically the glucose will get inside	Maybe you could state that as "the cell membrane is permeable to glucose."
it changed because the tube absorbed the iodine,	So are you saying "the molecules diffused through the membrane?"
I predict that if the holes in the plastic are large enough, the glucose will go into the water solution..	Maybe you could state that as "both water and glucose molecules are able to move between the two environments."

Method

The primary goal of the study reported in this paper was to facilitate discursive change in the classroom in order to support student learning. This longitudinal study was conducted as a two year design study in 9th grade biology classrooms in an urban school district. The objectives of the study were to (a) develop teachers' capacity to facilitate Accountable Talk discussions in biology; (b) increase the amount of Accountable Talk in whole class discussions over time; and (c) increase student learning through participation in the Accountable Talk discussions.

Sample

As previously mentioned, much of the existing literature on discursive instruction has been limited to special populations, including expert teachers. Our intention was to make Accountable Talk widespread instructional practice. In particular, we sought to target populations of teachers that were not yet experts in Accountable Talk, and likewise, student populations that have not otherwise experienced rich discursive instruction. An urban school district in Pennsylvania agreed to participate in the study with a focus on high school biology. In Year 1 we trained 17 district biology teachers, from 8 high schools. The teachers had varying levels of experience, ranging from 1-15+ years experience. Of the 17 teachers, 7 consented to participate in the study, which reduced to 4 teachers in 2 schools after attrition. 108 9th grade biology students, from 12 classes (5 remedial) consented to participate in this first year of the study. In Year 2, we worked intensively with 3 consenting teachers (2 after attrition), in the same school, with 113 consenting 9th grade biology students from 12 classes (8 of which were remedial level), which represented 63% of 9th grade biology students. In this paper we report on analyses of one teacher's classes over the two-year period, who participated in the intervention both years, which enabled us to examine his growth in Accountable Talk facilitation across his classes over the 2-year period. 59 class sessions altogether were recorded and transcribed from this teacher.

Classes were audio recorded and live transcribed, focusing on recording utterances and attributing them to individual speakers. Audio recordings were used to fill in transcripts from live transcription. In the second year of the study, we developed software that enabled us to live transcribe and attribute turns with a timestamp, which helped to greatly reduce the transcription time of dynamic multiparty classroom talk.

Teacher-Level Intervention

The teacher-level intervention focused on developing teachers' use of Accountable Talk in whole class discussions. There were 3 core strands to the professional development intervention: (1) identifying leverage points in curriculum; (2) planning discussion lessons; (3) Accountable Talk simulations. In Year 1, the professional development was conducted in 6 half-day pull out sessions over the course of the academic year. Each session consisted of a series of tasks engaging these core strands. With respect to Strand 1, the coach worked with teachers to identify leverage points for Accountable Talk within the district-mandated curriculum. For Strand 2, the coach lead teachers in collaborative planning of discussions lessons, focusing on identifying overarching questions that could open up space for deep reasoning about the subject-matter. For Strand 3, the coach led Accountable Talk discussions, in which the teachers took on the role of students in the discussion. In the first iteration of the intervention, teachers voiced difficulties in translating their experience in simulated discussions with their teacher peers, to discussions that would be appropriate for their learners and their conceptual level. For example, Excerpt 1 includes comments from teachers during these post simulation reflective discussions: "...but my kids can't do this!", "...I won't be able to do this in my school!", "...We [teachers] know more, that's why WE can do AT [Accountable Talk]".

What we saw in the dialogue, as we tracked participating teachers' Accountable Talk discussions after the intervention, was a slow change in their discursive style (see Table 4).

Table 4: Excerpt of Year 1 discussion for the target teacher

Turn	Speaker	Utterance
1	Teacher	so if you had to make a prediction how would you-- what would you predict would happen in this situation just looking at it now?
2	Student 1	that uh--
3	Teacher	I'm asking Student 4
4	Student 4	um the sugar is going to get in the water
5	Teacher	ok so Student 4 thinks that some of the sugar is going to come out to the water. Ok. Alright what about the water? Is the water going to stay there?
6	Student 3	no it's going to go inside
7	Teacher	so we think the water is going to move inside ok. Alright. Why?
8	Student 1	Uh
9	Student 5	Because
10	Student 4	high concentration
11	Teacher	Right
12	Student 2	Permeability
13	Teacher	ok both, there's a concentration gradient that's what we call that right. So if you have a high concentration on this side and a low one on this side in between this is called a gradient. So things always move from high to low right.

What we can see in Table 4 is that the teacher is trying to elicit student reasoning in turn 1 by asking students to make a prediction, and in turn 7 which elicits justification for prediction, however in turn 13 he seems to be lecturing and returning to a focus on getting the answer right.

In the second iteration in Year 2, we focused on translation of professional learning to instructional practice. We refined the intervention so as to better support teachers in translating instructional learning into instructional practice. We adopted a coaching model (West & Staub, 2003), which included a tripartite pre-, during- and post-phase of professional development.

Table 5: Excerpt of Year 2 discussion for Teacher Nelson

Turn	Speaker	Utterance
1	Student 1	The strip will turn green and the pouch will weigh more because the glucose will enter the pouch
2	Teacher	Ok. You know what is important about what he said? He added "because" and then he put some reasoning there. That's what we want everybody to do. Ok, would you read that once more, louder, so everybody can hear it?
3	Student 1	The strip will turn green and the pouch will weigh more because the glucose will enter the pouch
4	Teacher	Ok. So he predicted that the glucose is going to enter the pouch and the strip will turn green, so he had a because. He told why. That's good. Student 2. Thanks.
5	Student 2	I said it continued at a steady rate in the glucose and the prediction is the glucose will always increase. And at the bottom, I said, "We think it will-" wait- "We think there will be glucose in the inside of the membrane eventually."
6	Teacher	We think that there will be glucose inside the membrane eventually.
7	Student 2	Yeah. The think on the scale

In the pre-instruction conference, coaches worked individually with teachers in their classrooms to plan lessons, again focusing on identifying leverage points for Accountable Talk within the curriculum and eliciting teachers to predict *their students'* ideas. In the midst of the during-instruction phase, the professional development coach observed lessons in person or via skype, attending to how teachers were using Accountable Talk to draw out student ideas and support reasoning. The post-instruction conferences elicited teachers' reflections on the observed lesson, again focusing on their facilitation of discussions using Accountable Talk to promote student reasoning. The teachers went through 7 iterations of the pre-, during- and post-conferences over the course of the academic year. The refinement of the intervention aimed to develop teachers' capacity to lead Accountable Talk discussions, in their classes, with their students. Teacher facilitation in Year 2 discussions (Table 5) can be characterized by greater attention to reasoning through dialogue, drawing out student ideas for joint reasoning, rather than searching for correct answers, characteristic of I-R-E.

Student-Level Interventions

In addition to the teacher-level intervention, we developed a series of student-level interventions designed to prepare students to engage better in the whole group teacher led discussions. Some of these activities were online CSCL activities focused on collaborative inquiry, while others were face-to-face activities focused on developing reading comprehension skills. We expected these activities to increase student responsiveness to attempts by the teacher to engage them in active discussion and therefore serve a reinforcing effect of the teacher-level intervention. Each of the CSCL interventions were themselves experimental studies in which students were randomly assigned to groups within their classes, and groups were randomly assigned to conditions. Results of these experimental studies are published elsewhere. In this paper, because we are evaluating the effect on teacher behavior, and because students within classes were assigned to different conditions within these studies, we are not evaluating the effect of those experimental manipulations, rather we are evaluating the gestalt effect of the class having participated in the study on the teacher's behavior. Regardless of condition, all students benefitted from the carefully designed enhancement activities. What varied was only the amount and style of support offered during the activities.

CSCL Activities

At one time point in Year 1 and at two time points in Year 2, the students participated in online small group activities facilitated by intelligent conversational agents, immediately prior to a teacher-led whole class discussion on the same topic. The Year 1 study and the first Year 2 study were both run during a unit on Diffusion. The second Year 2 study was run during a unit on Punnett Squares. A conversational agent-as-facilitator must be able to manage several differently-scoped supports and behaviors concurrently. Recent work has produced software architectures for conversational agents (Kumar et al., 2011) that can implement such orchestration within CSCL environments. In our implementation, the conversational agent acts as an instructor and facilitator, and presents a series of group exercises in a chat room with a shared whiteboard (Mühlpfordt & Wessner, 2005). In all three studies, students worked in groups of three to make predictions, discuss observations, and generate interpretations of their observations. In all cases, student groups were supported by intelligent facilitator agents that provided a macrolevel structuring of the task and some level of micro-level support such as encouragement to participate and positive reinforcement for contributions to the discussion as has been demonstrated to have a positive effect on group processes and learning in earlier studies with intelligent conversational agents as group discussion facilitators (Kumar et al., 2010). In all cases, students were instructed about Accountable Talk and were encouraged to use these practices in their discussions. What differed by condition was the nature of support targeting these behaviors specifically that were offered to students. The two target facilitation behaviors we have experimented with separately and in combination in our studies have been the Agree-Disagree move and the Revoicing move. In the first study, groups either received no additional support, support of an agent that directed them to provide facilitation behaviors to each other, or support of an agent that engaged directly in the facilitation behaviors. In the final two studies, we dropped the agent that directed the students to engage in the facilitation behaviors. Across the three studies, we found positive effects of the agent that engaged in Accountable Talk facilitation behaviors.

Support for Reading Comprehension

The second student-level intervention focused on developing students' readiness for a teacher-lead whole class Accountable Talk discussion through face to face collaborative learning. We conducted a collaborative reading experiment designed to scaffold students' reading and understanding of biology texts (O'Donnell & Dansereau, 1992) prior to an Accountable Talk discussion of the same topic. In Year 1 students in 5 classes were randomly assigned to either a collaborative learning condition or individual learning condition. Students in the collaborative learning condition were randomly assigned to groups to read the same text using a scripted collaboration protocol designed to structure peer learning and metacognitive learning strategies. Students in the individual learning conditions using the same scripted protocol for the text, but worked individually. In Year 2, 7 classes were randomly assigned to either the collaborative learning condition or the individual learning condition. Three classes were assigned to the collaborative learning condition and four classes were assigned to the individual learning condition. We revised the collaborative learning protocol so as to create authentic knowledge gaps in which collaboration was necessary to complete the task. In the collaborative learning condition, students were randomly assigned into groups of three for jigsaw reading of 1 of 3 texts (Aronson, 1978). The jigsaw activity included four phases, (1) reading and synthesizing text individually; (2) discussing and refining syntheses in same-text groups; (3) disseminating syntheses in groups with students that have read other texts; (4) discussing substantive themes across readings. The individual learning condition followed the same procedure for reading the three texts, thus received the same reading support, but did so individually in writing, allowing for a comparison between individual learning and collaborative learning to support readiness for Accountable Talk discussion participation.

Coding

Years 1 and 2 generated a corpus of 168 transcripts of Accountable Talk discussions, 59 of which we analyze in this paper as discussed above. Two raters hand-coded two full transcripts with the codes from Table 2, with an agreement of .8 Kappa. This amounted to approximately 500 teacher turns. In order to analyze change over time in teacher behavior, we needed the full set of transcripts used in our analysis to be coded. Thus, we made use of a machine learning tool called LightSIDE (Mayfield & Rosé, in press). Using LightSIDE, we first ran experiments over the two hand coded transcripts to verify that reliability of coding with automatically trained models would be at an acceptable level. We evaluated a model in LightSIDE utilizing the SMO classifier and a 10-fold cross validation evaluation methodology. Using this methodology, we divide the coded data into 10 equal segments, and on each of 10 iterations, we train a model on 9 of those segments and test on the tenth. We then average performance values across the 10 segments. We did this separately for each type of Accountable Talk move as a separate binary classification tasks as well as for the whole set taken together as a multi-way classification task. For the multi-way classification, performance was low, i.e., only .52 Kappa (Human agreement just on Agree-Disagree was kappa of .86). Performance also varied across the individual binary classification experiments. The performance for the Agree-Disagree move was the highest, with a kappa value of .56. Using a classifier trained over all of the hand coded data, we then automatically coded all of the other sessions on the Agree-Disagree dimension since it had the highest kappa. Then, we reassigned the codes over the whole corpus using cross-validation so that the codes on all segments would be automatically assigned, and therefore more consistent, so that we could avoid seeing a big difference in performance on the two sessions that were picked to hand code just because they were coded by hand rather than by computer. This is a standard practice in machine learning. Ideally, we would prefer to measure growth in uptake of all Accountable Talk moves over time. However, to the extent that teachers appropriate the facilitation moves as a set, as they are coached on all of them, we expect a high correlation between acquisition patterns. In this paper, we treat number of Agree-Disagree moves as a probe that is indicative of the teacher's uptake of Accountable Talk practices. And we leave for future work a more exhaustive analysis across the different facilitation moves, looking for different appropriation patterns over time.

Analysis of Change Over Time

In order to measure behavior trends and differences between years or between classes within years in these trajectories, we used standard growth modeling techniques. To compute these growth models, we used the Generalized Linear Latent and Mixed Models (GLLAMM) (Rabe-Hesketh, Skrondal, & Pickels, 2004) add-on to STATA (Rabe-Hesketh & Skrondal, 2012). Our growth models were three level models where time points were level one units, classes were level two units, and years were level three units. Thus, the structure of the model was time points nested within classes, which were in turn nested within years. Time was measured as days since the beginning of the school year.

Specifically we used what is referred to as a random intercept and slope model, which allows estimating a separate latent trajectory for the teacher's behavior in each class over time, where each trajectory is characterized by a regression with latent slope and intercept, relative to a slope and intercept per year. The slope and intercept per year is fit to a distribution around a slope and intercept for the whole model. The resulting set of slopes and intercepts from the classes within a year form a pair of distributions that can then be compared with the corresponding distributions for the other years in order to determine whether the teacher's behavior trajectory differs significantly based on intercept, slope, or both between years. A comparison can also be made between classes within years. A significant difference in intercept indicates that the teacher's trajectory differs from baseline with respect to the dependent measure at the initial time point. A significant difference in slope indicates that the teacher's growth in the dependent measure over time differs from baseline. Time points that correspond to months when there were no recorded lectures from a class were dropped from the analysis. Since a separate intercept and slope is estimated for each class, these time points can be dropped without biasing the model to the extent that would be the case if a single slope and intercept were estimated directly from the full unstructured collection of data points. Additional covariates can be added to the growth model to account for other factors that may have affected the level of the dependent variable.

Results

In order to evaluate teacher behavior over time, we used as a probe the automatically tallied number of Agree/Disagree facilitation moves used by teachers per class session. Number per class was used in our growth models as the dependent variable. Time point was the independent variable in the basic three level growth model, with time point nested within class period, which was in turn nested within years. In this basic model, we found a significant effect of time point ($F(1,57) = 4.27, R^2 = .07, p < .05$), but no difference between years or class periods either in terms of intercept or slope. The basic trajectory within each year is indicated as a regression line per year in Figure 1. Each dot, circle, or x in the scatter plots per year in Figure 1 represents one class period.

In a second analysis, we divided time points into 3 types. The first notable type were class periods that immediately followed a class period where the students participated in a CSCL activity. These are marked as Collaboration in Figure 1. The second notable type were class periods that immediately followed a class period where the reading comprehension enhancement activities occurred. These are marked as Other Group Prep in Figure 1. All other class periods are marked as Baseline. This 3-way factor was added to the model as a fixed effect within level 3 units. When we evaluated this more complex model, this factor was found to be significant, and it accounted for more variance in the dependent variable ($R^2 = .36$, $p < .001$). In a posthoc analysis, we determined that only the time points in the Collaboration group were significantly higher than baseline, and the effect size was 1.7 standard deviations.

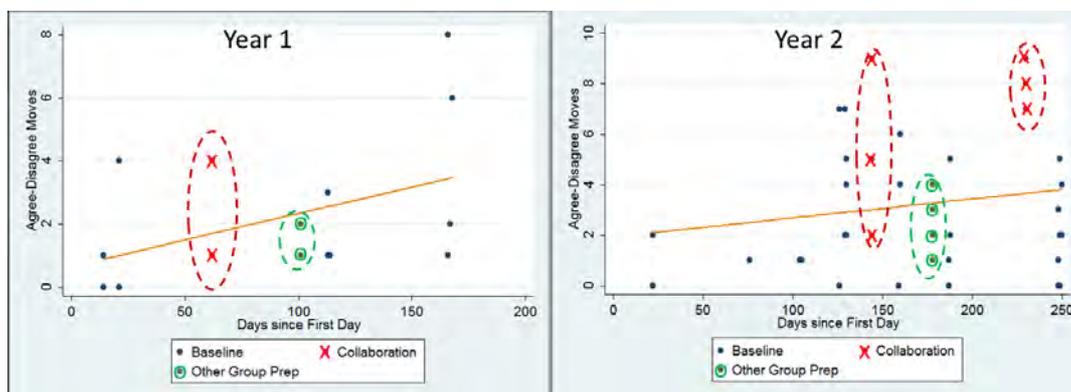


Figure 1 Results from two years of automatically coded class discussions

Thus, the finding is that having the students participate in a CSCL activity in the session immediately before a teacher-lead whole group discussion has a strong enabling effect on the instructor's facilitation behavior. We interpret the growth in the amount of teacher facilitated Agree-Disagree moves immediately following the CSCL activity that elicited these same behaviors helped to ready students for subsequent discussion. We do not find the same effect of the reading support activities, which were focused mainly on comprehension and not on inquiry skills, and were also not facilitated in the style of Accountable Talk.

Discussion

The results of our analysis complement the positive results reported in other publications about CSCL activities facilitated by intelligent conversational agents. As far as the authors are aware, this is the first evaluation of the impact of CSCL activities on discussion behavior in a larger group context outside of the online environment.

As we have reported, some of the challenges of the teacher-level intervention was their belief that this kind of instruction could be used with their students in their urban school district. As we interpret the finding of the CSCL activity in relation to shifting the discursive culture of 9th grade biology, one possible explanation for the results is that the teacher, knowing that the students had participated in preparatory activities, expected the students to be better prepared, and behaved differently because of that. While we cannot completely rule out that possibility, it is difficult to reconcile that interpretation with the fact that the reading enhancement activities did not lead to the same effect. In both cases the teacher knew the students were participating in enhancement activities meant to prepare them for the whole group discussion.

Conclusions and Current Directions

In this paper we have analyzed the results from a two-year teacher professional development program in an urban school district from the perspective of evaluating the impact of CSCL activities on teacher uptake of Accountable Talk facilitation moves in sessions immediately following the sessions in which the students participated in the CSCL activities. We find a strong reinforcing effect of these activities.

One limitation of the present analysis is that we do not analyze in detail the connection between student behavior within the small groups and their behavior in the large group discussions. Furthermore we do not analyze in detail the behavior of the students and how specific student behaviors that differed in the sessions immediately following the CSCL activities were responsible for the change in teacher behavior. One of the contributions of this paper is that it raises these important questions, which we leave for future work.

Another limitation of the study is that although it was conducted over 2 full school years, and although the effect size of the impact was very large, the total number of data points in the analysis is relatively small. Furthermore, the analysis only involves data from a single teacher. Thus while we have employed quantitative techniques in our analysis, we must consider the results we have obtained merely a case study. Nevertheless, the results show promise of an important role that CSCL can play in teacher professional development.

One of the big questions left for future work is why the impact on Accountable Talk uptake was local, and not sustained. One possible explanation is that what was driving the change in teacher behavior we observed was that students had thought deeply and critically about the ideas prior to that whole class discussion. If that readiness to engage at that level was not exhibited by the students at other time points because the preparation was not happening at those time points, the teachers may not have felt empowered to engage with the students at the same level. A possible solution is to make CSCL activities a more regular part of student involvement in the course.

References

- Adey, P., & Shayer, M. (1993). An Exploration of Long-Term Far-Transfer Effects Following an Extended Intervention Program in the High School Science Curriculum. *Cognition and Instruction*, 11(1), 1 - 29.
- Bill, V., Leer, M., Reams, L., & Resnick, L. (1992). From cupcakes to equations: The structure of discourse in a primary mathematics classroom. *Verbum*, 1, 2, 63-85.
- Chapin, S., & O'Connor, C. (2004). Project challenge: Identifying and developing talent in mathematics within low-income urban schools *Boston University School of Education Research Report No. 1*, pp. 1-6.
- Clarke, S. N., Spiegel, S., & Granger, R. (2012). *Supporting Student Engagement in Discussion through Collaborative Learning*. Technical Report. Learning Research and Development Center. University of Pittsburgh. Pittsburgh.
- Greeno, J. G. (in press). Classroom Talk Sequences and Learning. In L. B. Resnick, C. Asterhan & S. N. Clarke (Eds.), *Socializing Intelligence through Academic Talk and Dialogue*. Washington, DC: American Educational Research Association.
- Hmelo-Silver, C. E. & Barrows, H. S. (2006). Goals and Strategies of a Problem-based Learning Facilitator. *The Interdisciplinary Journal of Problem Based Learning*, 1(1), pp 21-39.
- Kumar, R., Rosé, C.P. (2011). Architecture for Building Conversational Agents that Support Collaborative Learning. *IEEE Transactions on Learning Technologies* 4(1), 1-1.
- Kumar, R., Ai, H., Beuth, J., Rosé, C. P. (2010). Socially-capable Conversational Tutors can be Effective in Collaborative Learning Situations, in *Proceedings of Intelligent Tutoring Systems*.
- Mayfield, E., & Rosé, C. P. (in press). LightSIDE: Open source machine learning for text accessible to non-experts *Handbook of Automated Essay Evaluation*.
- Mehan, H. (1979). *Learning lessons : social organization in the classroom*. Cambridge, Mass.: Harvard University Press.
- Mühlpfordt, M., Wessner, M. (2005). Explicit referencing in chat supports collaborative learning, Proc. Computer Support for Collaborative Learning (CSCL).
- O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. *Interaction in cooperative groups: The theoretical anatomy of group learning*, 120-141.
- Rabe-Hesketh, S. & Skrondal, A. (2012). *Multilevel and Longitudinal Modeling Using Stata*, Stata Press.
- Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). *GLLAMM Manual*. University of California, Berkely. U. C. Berkeley Division of Biostatistics Working Paper Series, Paper 160.
- Resnick, L. B., Asterhan, C. A., & Clarke, S. N. (in press). *Socializing Intelligence through Academic Talk and Dialogue*. Washington, DC: American Educational Reserach Association.
- Resnick, L. B., Salmon, M., Zeitz, C. M., Wathen, S. H., & Holowchak, M. (1993). Reasoning in conversation. *Cognition and Instruction*, 11(3-4), 347-364.
- Resnick, L. B., Michaels, S., & O'Connor, M. (2010). How (well-structured) talk builds the mind. In D. Preiss & R. J. Sternberg (Eds.), *Innovations in educational psychology: Perspectives on learning, teaching, and human development* (pp. 163-194). New York: Springer.
- Rosé, C. P., Jordan, P., Ringenberg, M., Siler, S., VanLehn, K., Weinstein, A. (2001). Interactive Conceptual Tutoring in Atlas-Andes, *Proceedings of AI in Education*
- Topping, K. J., & Trickey, S. (2007a). Collaborative philosophical enquiry for school children: Cognitive effects at 10-12 years. *British Journal of Educational Psychology*, 77(2), 271-288.
- Topping, K. J., & Trickey, S. (2007b). Collaborative philosophical inquiry for schoolchildren: Cognitive gains at 2-year follow-up. *British Journal of Educational Psychology*, 77(4), 787-796.
- Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: an empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction*, 9(6), 493-516.
- West, L., & Staub, F. C. (2003). *Content-focused coaching: Transforming mathematics lessons*: Heinemann Portsmouth, NH.

Acknowledgments

This work was supported in part by NSF grant SBE 0836012 to the Pittsburgh Science of Learning Center.

When Face-to-Face Fails: Opportunities for Social Media to Foster Collaborative Learning

Tamara Clegg, Jason C. Yip, June Ahn, Elizabeth Bonsignore, Michael Gubbels, Becky Lewittes, & Emily Rhodes
{tclegg, jasonyip, ahnjune, ebonsign, mgubbels, charley, eerhodes}@umd.edu
University of Maryland, 2117 Hornbake South Wing, College Park, MD 20742

Abstract: Productive collaboration is an integral component of socially constructed perspectives of learning. Yet effective collaboration is quite challenging and not without its own risks. Collaboration, both distributed and face-to-face, must be nurtured; technologies can support or undermine its positive growth in learning communities. We present an exploratory investigation of the types of social interactions that are both productive and non-productive in face-to-face informal science learning contexts. We include an analysis of the ways in which social media technologies can be designed to support more collaborative interactions.

Introduction

In our work, we aim to help learners see personally relevant aspects of science and begin to develop more scientific dispositions. Doing so involves helping learners see the social side of science and engage in scientific conversations with peers and adults who share their interests (Clegg & Kolodner, 2013). Productive collaboration is an integral component of such socially constructed perspectives of science learning. We seek to support collaborative dialogue among learners, their peers, and adults that involve the scientific inquiry practices of asking questions, designing experiments, collecting data, and developing claims (Chinn & Malhotra, 2002). However, effective collaboration is not an organic element of group-based scientific investigation (e.g., Barron, 2003). Collaboration, both distributed and face-to-face, must be nurtured; technologies can support or undermine its positive growth in communities of learners. While work has begun to address the needs of CSCL environments for supporting physically distributed groups, less is known about the potential of CSCL for supporting the social interactions of face-to-face groups. Alternatively, CSCW research has begun to look at the social interactions learners have in face-to-face environments when using collaborative technology. However, this research area has traditionally focused on the social difficulties learners encounter with collaborative technologies (e.g., virtually violating social norms) (Morris, Ryall, Shen, Forlines, & Vernier, 2004). In this paper, we look at the difficulties learners face when collaborating in face-to-face environments and ways in which CSCL technologies can help learners address these difficulties. Specifically, we ask: (1) *How can CSCL technology augment face-to-face environments to promote the productive social interactions necessary for collaborative learning?* and (2) *How do design features in CSCL technology facilitate productive social shifts?*

Background

In the context of scientific inquiry, the aspects of collaboration that we aim to promote include: sharing original insights and divergent ideas, providing critiques of one another, observing the strategies of others, building communal knowledge through conversation, and drawing on the expertise of others (Barron, 2003; Puntambekar, 2006). Yet effective collaboration is quite challenging and not without its own risks. When practiced in uninformed ways, collaboration can have negative effects, such as stigmatizing low achievers and creating dysfunctional interactions. Effectiveness often depends on such factors as how groups are organized, what the tasks are, and how the group is held accountable (Blumenfeld, Marx, Soloway, & Krajcik, 1996).

Cognitive Support for Collaboration

CSCL researchers have begun to address some of the challenges learners face with collaboration, and scientific inquiry more specifically, through the design of technology. For example, virtual spaces allow learners to connect across physical locations, share their experiences with one another, and build on one another's understanding (e.g., Scardamalia & Bereiter, 1996). This line of work has also shown that the structure of learning environments can often prevent or promote collaboration. For instance, the nature of the problems learners work on can play a role in fostering or prohibiting collaboration. Researchers have found that complex open-ended problems that utilize multiple elements of knowledge are best suited for collaborative learning because such learning situations impose more cognitive requirements on learners (Janssen, Kirschner, Erkens, Kirschner, & Paas, 2010). However, Janssen et al. (2010) also observed that working collaboratively imposes additional costs for learners, particularly the costs associated with coordinating group members and developing positive social relationships.

Social and contextual factors of collaborative learning come in many forms and impact learners'

cognitive processing. For example, competition in learning environments can prevent positive collaborative actions. Rick and Guzdial (2006) found that when university faculty did not have attitudes receptive of collaboration, the use of collaborative technology (i.e., wikis) in those classes was hampered. Likewise, when students expect teachers to drive the learning environment and be the director of discussions in CSCL environments, their motivation to collaborate is low (Kreijns, Kirschner, & Jochems, 2003; Puntambekar, 2006). These factors impact collaborative behaviors and affect the cognitive processing that learners can do together. For example, successful learning groups build discussion on their peers' previous contributions, attend to members' proposals and ideas, and pursue promising proposals jointly (Barron, 2003; Stahl, 2006). Positive group dynamics may enhance these collaborative learning behaviors, while negative group dynamics hinder them. This prior research, focusing on cognitive aspects of collaboration, illustrates that if we do not address the social aspects of collaboration, many learners will not experience the cognitive benefits of collaboration.

Social Support for Collaboration

Recent work has begun to focus on supporting learners' social interactions in CSCL environments. This work points to the importance of technology stimulating and promoting learners' social interactions (sociability) and prompting social presence, or the awareness of others in CSCL virtual environments. This research suggests that CSCL environments promote sociability by establishing two-way connections between distributed learners, enabling them to exchange diverse media forms (e.g., video, audio, text) as a means of stimulating interactions. Social presence can then be promoted non-technically in CSCL environments through such techniques as inclusion of moderators and training of participants on how to create social presence (Kreijns et al., 2003).

Many educational environments are termed blended learning, combining face-to-face interaction with distributed work using CSCL tools. Existing systems, such as WISE (Linn, Clark, & Slotta, 2002), have tools that support learners' scientific inquiry in such blended learning experiences. However, these technologies often specify what investigations learners will pursue and how they will go about their pursuit. Teachers or adult technology designers often take the collective cognitive responsibility (Scardamalia & Bereiter, 1996) for scientific inquiry processes. While collaboration and sharing across groups is supported in features of the system, collaboration is often a secondary focus to the individual or small group work of learners as they interact with teachers. On the other hand, systems like Knowledge Forum and wikis (Scardamalia, 2004) were designed to give collective cognitive responsibility to learners as they engage in inquiry. Collaborative cognitive work is foregrounded as learners strive collectively to build a shared knowledge base. While these knowledge-building systems seek to promote interaction with ideas, we aim to promote *engagement* in investigation, particularly to inspire young children to act, investigate, and experiment in their everyday lives. More work is needed to understand how CSCL technologies can support face-to-face interaction during learners' scientific inquiry experiences. Our work begins to address this gap in promoting social processes of blended scientific inquiry learning environments by investigating one face-to-face learning environment supported by technology in which both the environment and the technology were designed to promote scientific collaboration. We found the learners in this particular group had significant difficulties engaging in face-to-face, whole group conversations. Yet, working in an online community designed to support science inquiry seemed to shift the atmosphere of the physical setting. In this context, we aimed to understand the ways in which social media technologies supported more productive collaborative interactions.

Design of Social Media to Support CSCL

In this study, we used a CSCL tool called SINQ (for *Scientific INquiry*) (Ahn, Gubbels, Kim, & Wu, 2012), which leverages social media (SM) features to promote collaborative scientific inquiry. SM platforms, where individual members network, share information, and socialize are now a fundamental paradigm in computer-mediated communication (O'Reilly, 2007). SM technologies inherently tap into the sociable desires of individuals and thus may offer insight into ways of designing social support for collaborative learning. To date, researchers have focused on the literacy practices that young people enact in SM sites (e.g. Greenhow & Robelia, 2009), but little research has been conducted to examine how SM can provide scaffolded social support in areas such as science learning. Popular SM communities illuminate the different ways in which individuals want to engage with others, share knowledge, and incorporate online interaction into their daily lives (Shirky, 2011). For example, Wikipedia uses a wiki framework that allows thousands of individuals to contribute their knowledge, vet others' contributions, and in the process create a worldwide encyclopedia resource. In news sites such as Reddit, members vote up interesting posts, which then become a vetted source of information for the rest of the world.

We drew upon these SM approaches in the design of SINQ. SINQ is designed to enable distributed individuals to aggregate their micro-contributions into coherent science projects (called "challenges"). Members in SINQ do this in several ways. First, anyone in SINQ can contribute a question, hypothesis, or project idea to the system. In addition, any member can add a piece of the inquiry process to any peers' prior contribution. For example, Member 1 might contribute a question they wonder about in their everyday life; Member 2 might add

a hypothesis to test this question; and yet another Member 3 may devise a project to test the hypothesis. This framework taps into several features of SM platforms: (a) individuals can make micro-contributions to the inquiry process in ways that relate to them, or fit with their current skills and dispositions; and (b) the SINQ system aggregates these micro-contributions into coherent projects that members can try in any formal or informal setting. A second SM feature of SINQ is voting as modeled in Facebook and Reddit. These mechanisms were included in SINQ to be sociocultural tools for learning that allow collaborative groups to negotiate and vet norms of quality. In SINQ, the voting mechanisms ask members to reflect on what makes a good question, hypothesis, or project. For example, members can vote on whether questions are novel, whether the question is relatable, or if they spark curiosity and wonder. Votes on these criteria are incorporated into SINQ as the system aggregates pieces of inquiry into whole projects. In addition, it is possible that learners are implicitly guided to consider these criteria when posting new questions. We used SINQ for scientific inquiry activities within a particular science-learning context, *Kitchen Chemistry* (KC).

Design of Kitchen Chemistry (KC)

KC is an after-school program in which learners engage in scientific inquiry through cooking. In the first four sessions of KC, learners engage in *semi-structured activities*, becoming familiar with processes in measurement, data collection, and technology usage in the context of cooking experiments aimed to answer scientific questions (e.g., What do eggs do in brownies?). On *Choice Days*, learners are given opportunities to use what they have learned to develop questions, hypotheses, experimental procedures, and data collection techniques for their own food investigation. Learners make decisions on what recipes they want to modify, what variables they will control, what data to collect, and how to interpret their findings. In planning for Choice Day investigations, learners used SINQ to develop questions and hypotheses for their food investigations.

Methods

For this study, we employed the methods and standards of a comparative case study (Yin, 2003). The case is a single 12-week implementation of KC. Within this single implementation, we examined multiple units of analysis: the whole group discussions and three participant and facilitator interactions with SINQ. In this exploratory study, we examined the distractions and contributions that occurred during whole group discussion and the ways in which participants and facilitators interacted using SINQ.

Context and Data Collection

KC was implemented as a 12-week afterschool program that met once a week for roughly two hours in a local private school. Six learners between the ages of 8 to 11 participated in the program each week. The learners all attended the Montessori school that hosted KC. Anecdotally, parents and teachers told us that since KC provided a more hands-on approach to inquiry, the program appeared attractive to children with many difficulties in attention and interpersonal social dynamics. Each day we collected video recordings of all activities and discussions. As a part of a larger study, interviews were conducted with four of the learners and their parents at two intervals of the program. Lead facilitators also recorded post-observational field notes of their experiences each day in KC. Lastly, we collected analytics (e.g., time stamps, account logins) as participants logged onto SINQ and posted responses. The facilitators in KC in the case studies are Tammy, Jason, Mike, Beth, and Charley. All learners' names are pseudonyms.

Our goal was to understand learners' interactions across groups. We therefore analyze these interactions where they were most prevalent – in whole group conversations and in learners' interactions with SINQ. First, we characterized face-to-face whole group conversations. We aimed to understand aspects of the community that made conversations difficult and ways that SINQ mitigated some of these social issues. We transcribed three different whole group conversations in KC: Days 1, 4, and 5. These days represented a range of goals we had during whole group conversations, such as introductions (to learners, facilitators, and the program) and discussions about observations of experiments. We took a hybrid inductive and deductive coding approach, coding conversations by the productive aspects of learners' conversations and the points of social breakdowns. We used open coding to identify social breakdowns, which included episodes in which we experienced problems engaging learners in a community conversation (e.g., learner distractions, interruptions, and peer arguments). Clips in the category of learner productivity included those that were deductively coded as involving the types of scientific inquiry practices our program was designed to promote (e.g., asking scientific questions, making claims, generating evidence), based on existing frameworks for scientific practice (e.g., Chinn & Malhotra, 2002). Engagement in these practices in whole group conversation would then involve the types of productive collaboration others have called for (e.g., sharing ideas, understanding different perspectives) (e.g., Barron, 2003). To verify that our sample of whole group conversations from the first half of KC was representative of ways in which we observed participants interacting throughout the program, we also reviewed a subset of conversations on later days (Days 7 and 12). We confirmed that the codes we identified as occurring early in the project timeline retained similar frequency patterns throughout the program.

To examine learners' interactions with SINQ, we selected three video clips of learners interacting with the app, and coded the interactions using the same conversation breakdown and productive conversation codes. We also triangulated the clips with analytics from SINQ, interviews with learners, and our field notes. Once each of the cases was developed, we conducted a cross-case analysis of learner interactions using SINQ and whole group discussions to determine patterns in the data. Our coding of learners' interactions for productive and non-productive conversations offered key points of comparison between whole group discussion characteristics and learners' interactions in SINQ that then inform design implications.

Key Findings

In this section, we first summarize face-to-face breakdowns and productive conversations noted in the whole group discussions. Second, we present three cases, focusing our analysis on interactions with SINQ.

Whole Group Discussions

Four codes emerged from our data that characterized non-productive contributors to whole group discussions:

1. Competing foci: Learners were often focused on different aspects of the environment (e.g., the discussion, objects they could play with) or other ideas and topics that were unrelated to the learning context (e.g., songs, games). It was therefore hard for learners and facilitators to have a continuous large group conversation.
2. Talking out of turn: Learners often interrupted one another (and facilitators) with off topic musings such as loud noises, singing, or loud speech.
3. Learner derailing: As learners began to have outbursts, others would engage in the disruptive behavior as well. This behavior caused conversations to breakdown completely and often frustrated the facilitators.
4. Learner arguments: Some learners often got into social disagreements with one another that distracted them from the conversation. Two learners in the group, Donna and Anthony, were often in conflict. They would get into verbal arguments and small scuffles that demanded facilitators' attention during whole group conversations.

However, there were times when learners engaged in productive conversations. Across all of the clips coded as involving scientific contributions, we coded learners' generating evidence (3 instances), making claims (4 instances), and making observations (1 instance). Despite several instances of productive contributions, these clips were also interspersed with non-productive and interruptive behaviors. For example, on Day 4, while we discussed the "why" behind their cooking experiment, only one participant, Skylar, engaged in more than two turns of productive conversation. Overall, only two learners, Skylar and Arman, appeared to be on task. Anthony and Donna argued and had to be reprimanded while Freddie made noises, playing with the iPad and exacerbating Donna and Anthony's fights. Therefore, even during productive conversations, the environment was often still loud and facilitators often found it hard to manage the group to achieve our collaborative goals.

SINQ Interactions

In this section, we outline three cases of learner interactions with SINQ, to compare and contrast with our whole group discussion findings. In each case, we describe the interaction, followed by a brief analysis.

Case #1: Freddie and Eric

On Day 5 of KC, Freddie, a 10-year-old participant, worked with Tammy (facilitator) to develop his ideas for Choice Day in the SINQ platform. Although Freddie was initially distracted (i.e., playing with the research video cameras), he became intently focused on the activity once Tammy prompted him to consider an idea he had previously expressed interest in, making green brownies or "Greenies." As Tammy read Freddie's initial question in SINQ, "How should I make green brownies?" she prompted him to develop an experimental question to answer, as well as a hypothesis for the experiment. Tammy also reminded Freddie about a prior brownie experiment varying the number of eggs. Based on the past experiment, Freddie decided to increase the number of eggs in the new Greenies recipe because he thought that would make their brownies chewier. Although he was distracted again for short periods, he would return his focus to SINQ on his own.

Later, he became slightly distracted when the room lost Internet connectivity. As we regained Internet connectivity, SINQ was initialized back to the home page with all the learners' contributions. This incident prompted Freddie to read others' questions and ideas. Tammy asked him about some of the criteria SINQ uses to describe questions. Initially, Freddie told Tammy he was just voting for all of the questions. However, when they discussed other groups' questions he pointed to specific questions he wondered about and those to which he could not relate. Freddie was also able to get feedback from others on his own idea in SINQ. As he commented on others' questions, Freddie observed that someone had recently voted his question up. He also heard from a facilitator that another learner, Eric, had feedback for him. Interested to hear Eric's feedback, Freddie went over to Eric to learn more. However, Eric wanted to enter his hypothesis in SINQ first. Freddie continued to ask Eric if he had posted his hypothesis yet. Once he was done, Freddie and Eric looked at the iPad™ together, while Freddie read Eric's hypothesis out loud. Before leaving for the day, Freddie made sure to get his username and password for logging onto SINQ from home. Later, from home, Freddie provided a

hypothesis for Eric's first question in SINQ: "If I use 1 cup of baking soda and 1 cup of vinegar, how much explosion will I get, and how big will it be?" Posting a link to a slow motion video of baking soda being mixed with vinegar and foaming, Freddie replied to Eric, "4 and answer see this video..." In subsequent Choice Days, the two learners devised two variations of the same experiment (one making Greenies: green brownies; the other Whities: white brownies).

Case #1 Analysis

Two aspects of this case warrant attention. First, Freddie was one of the learners who, in whole group discussions, had to be reprimanded often for non-productive contributions, such as disruptive comments and noises. However, as we worked on SINQ, Freddie was more focused on the goal of the activity than he was during whole group discussions. Although he became distracted at points, he was easily prompted back to the activity or he brought himself back to it. The second important aspect of this case is Freddie's interaction with Eric. Previously, the group often ostracized Eric for talking out of turn frequently and loudly. During whole group conversations, Freddie and Eric only interacted when Freddie reprimanded Eric. However, on the day learners used SINQ, Eric was able to read about Freddie's idea from afar and think of feedback to give him in a non-disruptive way. As Eric wrote his hypothesis for Freddie in SINQ, the social landscape of the group was altered. Instead of pushing Eric away, Freddie was able to see that they had a common interest in his idea and he began to seek Eric out for his input. Their common interest persisted beyond the session in KC and at home.

Case #2: Donna and Anthony

For this second SINQ session, Donna, an 11-year-old girl, worked together with Jason (facilitator). Sitting next to them were Anthony, a 10-year-old boy, and Tammy (facilitator). Donna and Anthony were close friends from school, but they often got into arguments. They were the primary learners who were distracted during whole group discussions by arguing with one another. In this exchange, the physical positions of the interlocutors are important to note. From left to right, the group was seated next to each other as follows: Donna (child), Jason (adult), Anthony (child), and Tammy (adult).

As they began planning for Choice Day, Anthony did not have an idea about where to start. Tammy prompted him to refer to other questions previously entered into SINQ for ideas. Meanwhile, Jason was working with Donna who immediately had an idea of what to make for Choice Day, exclaiming "PUFFLES!" Jason worked with Donna to refine her idea to create a question in SINQ. Donna replied, "Ok, so they are these little balls of, ah, that are like cupcakes and they, you cover them with like some kind of hard sugar." As Donna excitedly detailed her "Puffles" to Jason, Anthony was glancing in Donna's direction. When Tammy asked Anthony if any of the prior food questions inspired him, Anthony replied, "Let's see, what do I want to make to compare to Donna's?" Instead of referring to SINQ entries for a Choice Day idea, Anthony chose to refer to Donna's idea. Meanwhile, Donna stopped typing her question into SINQ to talk more with Jason. As she talked about the hard sugar coating, Anthony responded to Tammy, "I like candy." Tammy prompted him to transform his thoughts about candy into a testable question. At this same moment, Donna repeated to Jason that her question was about "a hard sugar." Hearing this, Anthony quickly came up with the question, "Why are most candies hard?" At this point, Donna's hard sugar inquiry and Anthony's hard candies question were distinct enough that the children did not notice. The SINQ analytics data show that Anthony's question, "Why are candies hard?" was recorded before Donna's "How do you make a hard sugar?"

Later, Jason and Donna discussed how the Puffles coating would be like "jawbreakers," with Anthony again listening in. Tammy, who did not hear Jason and Donna's conversation, asked Anthony, "Now what kind of project do we need to do to answer that question?" Anthony replied, "Well we could, we could make kind of like a *jawbreaker* thing and like put..." The moment Anthony said "jawbreaker", Donna immediately and angrily yelled to Anthony, "You just take it from me!" Almost instantly, Anthony retorted, "No, I'm not!" and continued to tell Tammy how he would make a candy in the same type of hard sugar coating. Donna, in complete frustration, told Jason, "I never should have said it out loud!" Jason tried to assuage her, but she irritably folded her arms inward and stated, "Puffles was my idea! Then why is he stealing it?" At this moment, Jason guided her away from the situation, leading her outside the classroom to devise a new question.

Case #2 analysis

Donna and Anthony, while close friends, often could not work together in close physical proximity. In interviews, Donna called herself a "designer" and often had many ideas she wanted to develop. However, Anthony was at a different stage of question development than Donna. He had never used SINQ before and had not developed an investigation question on his own. Because Anthony and Donna were coming at the SINQ activity from two different levels of understanding, their physical proximity may have conflicted with what each of them needed. In a later interview with Donna, she said that the "stealing" of her idea made her feel "mistreated, like, they (referring to Anthony) didn't think that it, it was mine." In this case, an interesting difference between the face-to-face and virtual environment is highlighted. Having an idea upvoted requires

others to see and like your idea; this can be a form of social currency. But, for Donna, being face-to-face with another learner meant giving access to her idea verbally before it was attributed to her. This situation was problematic and led to conflict over authorship. However, this does not suggest that Donna did not want to interact with others' ideas. In the beginning of this session, Donna was browsing prior questions that had been written in the first SINQ session. Because both learners were at different points of ideation, in addition to learning new norms for participation, physically separating them may have allowed Donna to have authorship over her own idea first and then refine her idea while considering others' SINQ contributions. Conversely, Anthony may have been less tempted to overhear Donna's idea so he could instead focus on existing questions in SINQ to help inspire his own questions.

Case #3: Arman

In the first SINQ session, Arman, a 10-year-old boy, worked with Mike (facilitator) to devise an investigation. Arman was often relatively quiet in whole group conversations. His experience with SINQ sheds light on the ways in which SINQ facilitated conversation between him and the adults working with him, as well as his development of questions and an experiment. Initially, Arman did not have an idea for his investigation, and instead asked Mike for an idea. Mike randomly told Arman, "cinnamon goo," an idea Arman latched onto. At this point, Mike generated a question for Arman, "how can they be even more cinnamon?" Here, Arman asked, "Isn't like cinnamon like that stick?" Mike explained how cinnamon sticks are ground into powder. During this explanation, Arman began to type his first question into SINQ, "How does cinnamon relate to cinnamon rolls?" As Arman worked, SINQ's interface prompted him: "Do you wonder about this?" Using this prompt, Arman vocalized his thoughts on cinnamon, "because cinnamon, actual cinnamon is a solid thing." Mike interpreted Arman's response as a comparison of cinnamon sticks to cinnamon powder. Arman and Mike then discussed the nature of cinnamon and where it comes from. Concurrently, Arman scanned some of the questions other learners posted. Arman typed in his second question, "How do they make cinnamon sticks?"

While Mike responded to technical SINQ issues, Jason and Charley (both facilitators) recommended to Arman that he use SINQ to vote on other children's questions. As Arman did this, Charley asked Arman about his question. Arman explained his transition from the first and second question. Charley began a series of prompts to learn more about why he chose to explore cinnamon sticks and cinnamon rolls. Arman explained that cinnamon rolls are very sweet and "cinnamony". Probing for more detail, Charley asked, "is there such a thing as too sweet?" Both of them discussed what they thought is sweet and their experiences with the rolls. At this point, Charley asked Arman to consider working on a recipe for the rolls that would determine the amount of sweetness in the roll. Arman entertained the idea and entered a third question into SINQ: "How do they add the cinnamon to a cinnamon roll?" As Arman added his question, Charley shared how cinnamon is sprinkled between layers of dough. Building on this, Charley also suggested different ways to think about how to add the cinnamon into the dough (e.g., different flours or sugars, in-between layers or outside the layers). This discussion inspired Arman to develop another question: "Wich would taste more cinnaminy: cinnamon in the doe or cinnamon in between the layers."

Case #3 analysis

Arman's question development process iterated through several stages. To start, Arman needed a seed of encouragement. In his interview, Arman expressed a sense of initial anxiety for the investigation development: "...I got really worried about what I should do." However, working together with Mike, Arman felt better and more confident for the task: "and then somebody brought up cinnamon and I brought up cinnamon rolls. And then I had my questions for that." Second, the process of seeing other learners' questions helped Arman to see how his friends were contributing. He was able to see the questions and guess who authored them. Third, the prompts from SINQ sparked learner and facilitator conversations about scientific questions and evidence. For instance, SINQ asks learners to vote up questions by asking them to consider, "do you wonder about this?", "is this a novel question?", and "can you relate to this question?" As Arman followed these prompts, Mike asked him if he ever considered questions about cinnamon, which prompted Arman to ask about cinnamon's composition, and how cinnamon sticks relate to cinnamon powder. From one prompt, Mike and Arman were able to elaborate on cinnamon's transition from stick to powder, a line of reasoning that led to more discussions about how cinnamon is added to the rolls.

Discussion

In looking across cases, we observed three ways in which SINQ augmented the face-to-face environment to promote social interactions. First, SINQ facilitated *changes in the social dynamic* between group members by fostering shifts in learner's relationships with one another. The technology affordances helped bring some who were initially separate together, and others who were initially inseparable further apart. However, in both cases learners needed to be separated so that they could begin to work together. The technology then afforded and enabled such distanced collaboration. In Freddie and Eric's case, SINQ helped two learners who were

previously socially distant to find common interest with its social question-sharing features. In contrast, Donna and Anthony's work in close proximity with SING prompted a heated argument. After the angry exchange, Donna began to sign all of her SING entries under an alias ("the DESTROYER"), so that Anthony would not recognize her contributions. In this way, a technical feature enabled distance between the two learners, which the facilitators had to impose to be productive distance, so that they could focus on scientific collaboration.

The second way the technology affordances helped augment the face-to-face environment was in *fostering collaboration across groups when learners were ready*. SING provided ways for learners to interact individually in the face-to-face environment, yet still collaborate virtually in a scaffolded inquiry process. This enabled learners to focus on the activity without getting distracted by one another. Learners could then individually choose opportune times to find out about what others were doing and sharing. In this way, many of the breakdowns we experienced in whole group conversations (e.g., learner derailing) were either no longer a problem, or addressed more easily and quietly individually. Once learners were able to focus on the activity and collaborate virtually, they used the technology to engage in the collaborative processes we aimed to promote (e.g., sharing original insights). In each case, learners browsed and voted on one another's science questions according to the standards put forth in SING. This process helped them to be aware of others' ideas and to consider them scientifically (i.e., in choosing whether or not and how to vote them up or down).

Finally, the technology afforded *enhanced collaboration with facilitators*. Across the cases and in other non-represented cases, the types of questions SING asked learners to vote on (e.g., is this a novel question, can you relate to this question), and the process of voting questions up and down fostered conversations between facilitators and learners about what it meant to ask and evaluate scientific questions. Facilitators often needed to help define or differentiate the questions SING asked for learners to understand and answer them. This helped to set standards in the group for good questions, as learners were applying those standards. In addition, observing or having access to what other groups were doing in real time helped facilitators to foster connections with learners. For example, Eric's facilitator Beth was able to connect Freddie and Eric because they had seen Freddie's question on SING. Moreover, Charley and Mike used the prompts and context to help Arman come up with his cinnamon rolls question.

Design Implications and Conclusions

Our analysis points to several affordances of CSCL technology to facilitate productive social shifts in face-to-face learning environments. First, our analysis highlights the importance of providing *support for facilitating scientific communication*. Specifically, CSCL technology should show other learners' contributions at relevant and opportune times. Our findings show that there are ways in which SING provided such affordances and ways that it did not. In particular, we came to appreciate the importance of documenting ownership of ideas to prevent communication breakdowns. Our analysis emphasizes the need for technology that provides support for giving learners credit and protection of their ideas as they collaborate with others. The concept of authorship is particularly salient in a platform like SING, where individuals are asked to contribute pieces to a larger product, and thus give up some of their sole authorship in the inquiry process. In facilitating scientific communication, our findings also point to the importance of a *community repository* for enabling learners to put their ideas together and quickly provide feedback on one another's contributions. Voting mechanisms can then work well for prompting learners' scientific reflection and helping them to consider common interests. Providing learners with *multiple entry points into science*, or multiple ways in which to make contributions, is also important for supporting the social aspects of learners' collaboration. SING facilitated these different entry points by allowing learners to contribute different pieces of the scientific inquiry process.

Finally, our analysis underscores the importance of *factoring the learning context into the design and implementation of CSCL technology*. Specifically, our program was situated in a Montessori school where learners are acclimated to a culture of individual work and learning pace. The context in which we used SING more closely resembled that culture than our whole group conversations. Yet, we still needed learners to work together, to hear one another's ideas, and learn from each other. SING provided a means to strive toward both goals, as learners could work individually in the face-to-face environment, yet still share and hear one another's ideas, provide feedback, and recognize common interests. Our work suggests that CSCL designs that embed social affordances can help to foster productive changes in collaboration styles when needed, and can be used to alter the culture of face-to-face learning environments that are more individually focused. Previous CSCL work (e.g., Rick & Guzdial, 2006) suggests that there must be a culture of collaboration established in the learning context for computer-supported collaboration to be effective. Our work, however, moves outside of the classroom and looks at the challenge of supporting learners' social processes and their modes of working and communicating with one another.

Much of the CSCL, CSCW, and online communities research focuses on how technology can bring people together. Our work shows that CSCL tools sometimes need to provide separation to help learners begin to internalize the social skills needed for effective group work. SING explicitly scaffolded a collaborative process where individuals are asked to contribute pieces of inquiry and build from others' contributions. As our

learners used SINQ, they were guided implicitly to work from this paradigm, which then facilitated their face-to-face interactions in productive ways. Based on these findings, we suggest that CSCL technology can actually bring learners closer together when face-to-face interactions are difficult. Perhaps next steps would be to explore how CSCW frameworks for identifying and addressing social conflicts (e.g., Morris et al., 2004) arising from use of collaborative technologies might apply to addressing conflicts arising from face-to-face interactions.

References

- Ahn, J., Gubbels, M., Kim, J., & Wu, J. (2012). SINQ: Scientific INquiry learning using social media. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*, Austin, TX (pp. 2081 - 2086). New York, NY, USA: ACM.
- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12(3), 307-359.
- Blumenfeld, P., Marx, R., Soloway, E., & Krajcik, J. (1996). Learning with peers: From small group cooperation to collaborative communities. *Educational Researcher*, 25(8), 37-40.
- Clegg, T., & Kolodner, J. (2013). The role of life-relevant learning environments in helping middle-school learners develop more scientific dispositions. Manuscript submitted for publication. Copy on file with authors.
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86(2), 175-218.
- Greenhow, C., & Robelia, E. (2009). Old communication, new literacies: Social network sites as social learning resources. *Journal of Computer-Mediated Communication*, 14(4), 1130-1161.
- Janssen, J., Kirschner, F., Erkens, G., Kirschner, P. A., & Paas, F. (2010). Making the black box of collaborative learning transparent: Combining process-oriented and cognitive load approaches. *Educational Psychology Review*, 22, 139-154.
- Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior*, 19, 335-353.
- Linn, M.C., Clark, D., & Slotta, J.D. (2002) WISE design for knowledge integration. *Science Education*, 87(4), 517-538.
- Morris, M., Ryall, K., Shen, C., Forlines, C., & Vernier, F. (2004). Beyond "social protocols": multi-user coordination policies for co-located groupware. *Proceedings of the 2004 ACM conference on Computer Supported Cooperative Work (CSCW '04)*, Chicago, IL (pp. 262-265), New York, NY, USA: ACM.
- O'Reilly, T. (2007). What is web 2.0: Design patterns and business models for the next generation of software. *Communications & Strategies*, 65, 17-37.
- Puntambekar, S. (2006). Analyzing collaborative interactions: divergence, shared understanding and construction of knowledge. *Computers & Education*, 47(3), 332-351.
- Rick, J., & Guzdial, M. (2006). Situating CoWeb: a scholarship of application. *International Journal of Computer-Supported Collaborative Learning*, 1(1), 89-115.
- Scardamalia, M. (2004). CSILE/Knowledge Forum®. In *Education and technology: An encyclopedia* (pp. 183-192). Santa Barbara, CA: ABC-CLIO
- Scardamalia, M. & Bereiter, C. (1996). Computer support for knowledge-building communities. In T. Koschmann (Eds.), *CSCL: Theory and practice of an emerging paradigm* (pp. 249-268). Mahwah, NJ: Lawrence Erlbaum Associates
- Shirky, C. (2011). *Cognitive surplus: How technology makes consumers into collaborators*. New York, NY: Penguin Press.
- Stahl, G. (2006). Sustaining group cognition in a math chat environment. *Research and Practice in Technology Enhanced Learning*, 01(02), 85-113.
- Yin, R. K. (2003). *Case study research: Design and methods*. Applied social research methods series (Vol. 5). Thousand Oaks, CA: SAG.

Acknowledgments

We would like to thank the participants in our Kitchen Chemistry program, as well as supporting faculty and parents at the school in which Kitchen Chemistry was held. Additionally, we would like to thank Kidsteam researchers Allison Druin, Mona Leigh Guha, Gregory Walsh, Elizabeth Foss, Evan Golub, and KidsTeam child designers for partnering with us to design the SINQ app. Furthermore, we acknowledge Janet Kolodner and Christina Gardner for their work in developing the Kitchen Chemistry program. Finally, we thank the Computing Innovations Fellows Program for supporting this research.

Aggregating Students' Observations in Support of Community Knowledge and Discourse

Rebecca Cober¹, Colin McCann¹, Tom Moher², Jim Slotta¹

¹Ontario Institute for Studies in Education, University of Toronto

²Department of Computer Science and Learning Science Research Institute, University of Illinois at Chicago
rebecca.cober@mail.utoronto.ca, colinmccann@gmail.com, moher@uic.edu, jslotta@oise.utoronto.ca

Abstract: We present two case studies of scientific inquiry with Embedded Phenomena, where two middle school science classes participate in whole-class investigations of phenomena that are embedded within their classroom. Students share observational data with their peers using networked handheld devices. Student-contributed data is collected, aggregated and re/presented in coherent visualizations, designed to guide students in their investigations. We examine the forms of these collective representations with a view towards understanding their efficacy in scaffolding learners and resolving their driving inquiry questions. We analyze the patterns of discourse and of use surrounding these aggregate visualizations during teacher-led, whole class discussions. For each case, we report three trends based on a visual analysis of the coded discourse, followed by a discussion of the patterns of use surrounding the display of aggregate screens. We conclude with a synthesizing discussion of the two cases.

Introduction

Inquiry-based instruction has been advocated as an approach that is well-suited for teaching and learning about science (Bransford et al., 2000). Inquiry has been defined as “*an educational activity in which students individually or collectively investigate a set of phenomena—virtual or real—and draw conclusions about it. Students direct their own investigatory activity, but they may be prompted to formulate questions, plan their activity, and draw and justify conclusions about what they have learned*” (Kuhn et al., 2000, pp. 496-7). In the present research, we describe two case studies of scientific inquiry concerning Embedded Phenomena (EP; Moher, 2006). EP are “ambient” media (i.e., persistent, passive, and embedded within the classroom environment) where a running simulation is mapped to — and then embedded within — the walls, floors, or ceiling of the classroom, providing a rich context of scientific inquiry for whole class investigations.

Collaborative learning is an important dimension of inquiry-oriented learning. It includes the exchange of ideas, data, or artifacts amongst peers and is often facilitated by networked scaffolding technologies (Edelson, Gordin, & Pea, 1999). A whole-class investigation approach requires that students cooperate and share data, making it likely that they will need to draw on multiple data sources, including personal- and peer-collected datasets for their analysis. Managing multiple streams of data, especially data that accumulates over several days and weeks, presents learners with a significant challenge. Computing and networking technologies have the potential to support students in such complex investigations, enabling them to efficiently collect, manage, and visualize data (Roschelle et al., 2007). Collective or aggregate representations, generated from the data gathered by individual students or small groups of students using networked devices, can take a variety of forms, including histograms such as ClassTalk (Dufresne & Gerace, 1996), annotated image maps (Tatar et al., 2003), and graphs generated from probes and sensors (e.g., Tinker, 2000).

The task of designing such representations is not simple. Although it may be relatively easy, from a technological point of view, to collect and aggregate student-contributed data, the real challenge is to present the resultant collection of data in such a way that the students' and teacher can derive meaning from it and determine a sense of progress in their inquiry. The following research questions guide our discussion of the two cases we present:

1. Do the representations of collective input make patterns in the data visible, allowing the knowledge community to make progress towards resolving inquiry questions?
2. What interaction patterns emerge when teachers use these collective visualizations during whole-class discussions?

Theoretical Foundations

The establishment of a knowledge community in a classroom has been the aim of research projects such as Fostering Communities of Learners (Brown & Campione, 1996) and Knowledge Building (Scardamalia & Bereiter, 2006). These projects have sought to advance our understanding of the role of social interaction in

learning, emphasizing collective epistemology (Palinscar, 1998). In a learning community approach, the emphasis is on constructing knowledge through the sharing of data, ideas, and theories, within a rich social context.

Our pedagogical model, known as Knowledge Community and Inquiry, (KCI) builds on the foundation of knowledge communities as described above, with an added major emphasis on scaffolded inquiry (Slotta & Linn, 2009). In KCI, collaborative inquiry activities are carefully designed and mapped to curriculum goals, with students co-constructing a collective knowledge base, and then using this knowledge base as a resource for subsequent inquiry (Slotta & Peters, 2008). The design of KCI curriculum requires careful development of collaborative knowledge construction activities that result in a knowledge base that is indexed to the major learning goals (e.g., science standards or expectations). In addition, collaborative inquiry activities are designed to make use of that knowledge base, advance the community's understandings, and lead to assessable outcomes.

To investigate the KCI framework, we have advanced the notion of a "smart classroom", which is concerned with the coordination of the flow of people, roles, groups, activities, and materials, using specified pedagogical structures and curriculum content (Slotta, 2010). In this model, students and teachers work together as members of a knowledge community, using a variety of technologies, including laptops, tablets computers, interactive tabletops, and large format displays. A suite of custom software applications allows for the delivery of all materials, data collection (e.g., of student reflections, observations, tags or any other interaction), and coordination of all collaborative inquiry conditions. The smart classroom infrastructure ensures that (1) all devices connect and communicate with each other via a wireless network, (2) all inputs are appropriately attributed, sorted and available for retrieval, and (3) the developing knowledge base is accessible as a resource to individual students and the community as a whole.

The present research was conducted within an instructional environment referred to as Embedded Phenomena for Inquiry Communities (EPIC; Moher & Slotta, 2012), where KCI was applied as a pedagogical model to develop a knowledge community for elementary students to investigate Embedded Phenomena. In EPIC classrooms, students work collaboratively (i.e., in small groups) and collectively, sharing information and solving problems. Interactions, including the exchange of data and theories, are carefully designed to support the growth of collective knowledge concerning the EP under investigation, as captured in various representational forms (i.e., "aggregate representations").

In the present paper, two EPIC curricula will be described, with synthesizing discussion. In both cases, our goal was to support the construction of a community knowledge base, aggregated from the observational data collected by individuals and small groups, and to make the results of individual inquiry actions, such as observations and hypotheses, persistent and searchable. Further, we sought to dynamically generate (i.e., in "real time") views or representations of this aggregate knowledge that would make important patterns visible to student and teachers. Using a co-design methodology (Roschelle, Penuel, & Shechtman, 2006), researchers, technologists, and teachers worked together to develop the technology-enhanced materials.

Case One: HelioRoom

Methodology

Research Setting and Design Considerations

Our first EPIC case employed the HelioRoom EP, which maps the orbital planetary system onto the four walls of the classroom. Students adopt a heliocentric perspective and observe representations of the planets (colored circles, see Figure 1) through four "windows" — monitors affixed to each wall of the classroom. As the "planets" orbit in a counter-clockwise direction, the circles appear to move off of one "window" and then reappear on the next after varying time intervals (i.e., depending on the velocity of the circle). Students track the occlusion relationships of the colored circles (i.e., when one colored circle passes in front of a different colored circle), and use data aggregated over all students' observations to advance and support their theories concerning the identity of each planet.

In a previous nine-day implementation of the HelioRoom EP, paper cards were used to capture student observations and hypotheses, which were displayed on a classroom bulletin board. Thompson and Moher (2006) report "the density and lack of organization of the observations on the Idea Wall made it extremely difficult for students to retrieve the evidence they needed to construct argument chains to support their theories" (p. 1001). In the present study, we applied the ideas of KCI to develop a 90-minute lesson where students worked collectively to observe all occlusion relations at all four monitors, adding their observations to a collective set using networked mobile devices (i.e., tablets) — effectively replacing the paper supports of previous HelioRoom studies — and then reflecting on the aggregate as it emerged in real time.

Participants

Two classes of grade six students and their teachers (our co-design partners) participated in the trial. These students and teachers belong to an elementary school that has a strong history of inquiry-based instruction, located in a multicultural urban centre in Canada. The present analysis is restricted to one of the two classes (n=12) in order to develop a rich case study of the discourse patterns that occurred around the re/representations of student- and peer-collected data.

Method

Students had recently completed a curriculum unit on Earth and Space and were familiar with the order of the planets in our Solar system. The teacher began by demonstrating how to use the tablet software to contribute observational data to the knowledge base. Students worked in pairs to contribute observations. Our design supported viewing of one aggregate representation at a time (i.e., all observations for one planet). By touching a colored disc at the top of the screen, the tally (frequency count) for that color could be viewed – see Figure 1. The teacher used the interactive white board (IWB), which displayed the aggregate data, to inform whole-class discussions. Students could also view these data visualizations on their own tablets at any time during the inquiry process, and all such representations were updated in “real time” with each new observation.

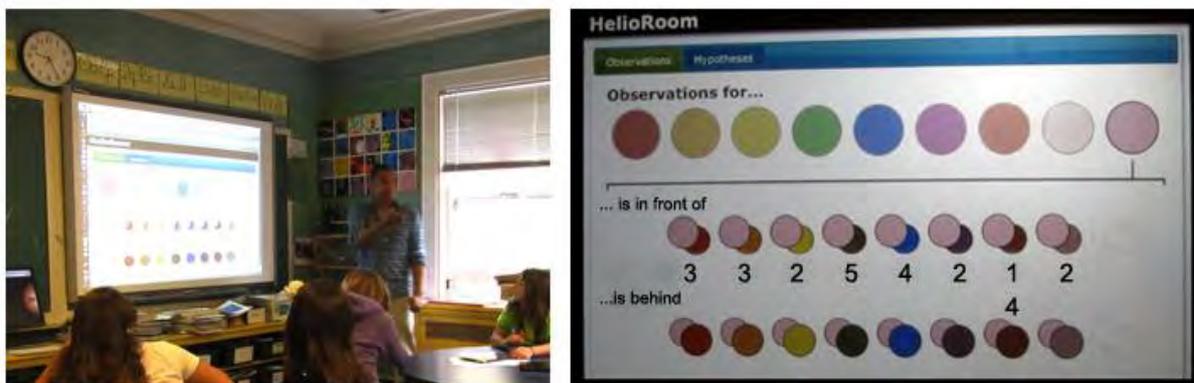


Figure 1. The image on the left shows the teacher and students using the aggregate tallies on the IWB to inform discussion. The image on the right shows data for “pink” planet, taken partway through the activity. At that time, the majority of tallies show “pink” passing in front of all other colors except for brown.

Data Sources

Data sources include data logs of student contributions, a transcribed video recording of the learning activity, researchers’ field notes, audio recordings of two debrief sessions with the teachers (one following each lesson), and focus group interviews with teachers and students following the lessons.

Analysis and findings

Students contributed 425 relational observations, with an 86% accuracy rate (i.e., they correctly observed an occlusion relationship for one planet pair).

This paper analyzes the whole-class discussions that were concerned with the aggregate representations of student-contributed observational data. This analysis informs an understanding of how the knowledge community understood the collective data, as well as its role in guiding them towards their inquiry goal.

A grounded theory approach (e.g., Creswell, 2008) was used to establish a coding scheme for the statements made during the whole-class discussions: summative, interpretive, directional, negotiation, and agreement, as shown in Table 1.

Table 1: Code name, description, and example for each of the five codes

Code Name and Description	Example from transcript
Summative statement regarding what the knowledge community has observed, giving words to the data	Teacher: “There has been one observation made that red is in front of orange.”
Interpretive statement concerning the numbers in the visualization, as well as best practices, efficacy, or accuracy of the aggregate representation	“Well, it depends how close the two numbers are...”
Directional statement guiding subsequent inquiry (i.e., for further observation focus) or interpretation of the aggregate (i.e., what we are looking for in the aggregate)	“So let’s actually look for those [areas of disagreement in the aggregate].” “Lets get some more observations of green.”

Negotiation statement - Working towards resolving a disagreement that has arisen	“But seven people are seeing it the other way around.”
Agreement statement - Reaching agreement regarding identity of a planet	“So do you think we’d be safe in saying that orange is Uranus then?”

Patterns in Coded Discourse

Our coding of discourse that occurred relating to the aggregate representations suggests that those summary views helped to make patterns visible and allowed the knowledge community to make progress towards their goal of deciphering the identity of the obfuscated planets. The codes were plotted on a timeline (Figure 2), which includes the duration of the whole class discussions and the points at which the knowledge community reached agreement about the identity of a planet.

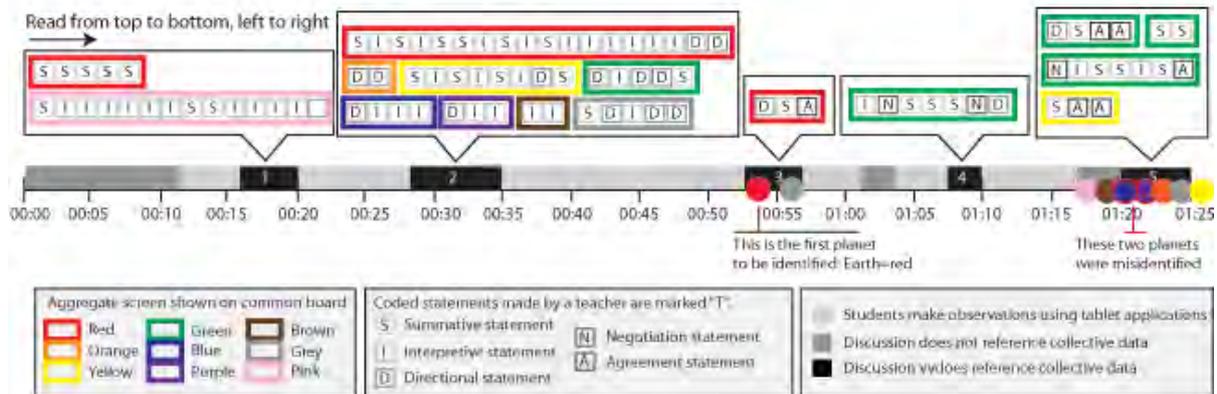


Figure 2. A timeline of whole-class discussions (the black rectangles on the timeline are exploded to reveal codes). Codes include a colored outline (representing the aggregate screen that displayed during comment).

This coding reveals three important trends: 1) In the earliest phases of the learning activity, discussions concentrated on *description* (code S) of the numbers displayed in the collective data sets. The teacher modeled for his students how to give words to the data: “One person has observed, or there has been one observation made that red is in front of orange.” 2) Following the descriptive phase, discussions focused on *interpretation* of the numbers (code I) in the observation aggregates. One student correctly interpreted the aggregate shown in Figure 2 as follows: “It means it [pink] is the second planet closest to the sun because the only one that [pink is] not in front of is brown.” In the middle phase of the learning activity, discussions often began or concluded with a statement that provided *direction* for ongoing inquiry (code D). For example, in discussion two, the teacher gave the following instructions: “I’ll click through them, and you see if you find any, okay? Real disagreement... Maybe that’s something we should look for.” By tapping on a color at the top of the IWB, the teacher displayed the tallies for that color. Students looked for patterns in the data — pairs of numbers that were “too close to call”. Students identified pairs that were only separated by two or three tallies (i.e., 5 observations that yellow passed in front of green vs. 3 observations that green passed in front of yellow). The teacher then directed students to focus their input of observations on problematic disc colors: “So go back to it and see if you can figure out any of those [colors]. These ones with disagreement to them [pointing to list on chalkboard].” The knowledge community determined their own set of best practices for acceptable levels of disagreement, and when pairs of numbers were deemed unacceptably close, the corresponding planets were targeted for further inquiry. 3) In the latter phase of the learning activity, discussions were resolved when the knowledge community reached agreement regarding the identity of a planet. A *summative* statement (code S) preceded all *agreement* statements (code A). For example, the teacher read the observation tally for green passing in front of orange from the aggregate: “Green faster than orange, 12-7”, followed by, “so do you think we’d be safe in saying that orange is Uranus then?”

Patterns of Interaction

In addition to the coding of discourse, we examined when the discussions of aggregate data occurred during whole class discussions, and how they proceeded. The timeline reveals that discussions (shown as black bars in Figure 3) were interspersed throughout the learning activity at regular intervals (beginning, middle, and end). In the first two discussions, the teacher worked systematically through the aggregates by tapping the colors at the top of the screen, from left to right (see Figure 2). This provided an overview of the student-contributed

observations, allowing the knowledge community to get a sense of which planets they had already identified and those that were still unidentified. In the latter discussions, the yellow and green datasets were the focus of discussions, possibly because these were the slower moving planets, and were thus more difficult to identify because nearly every color passed in front of them. In all discussions, the teacher actively chose which of the aggregate representations would be the most productive topic for discussion, and used them to help motivate and inform students' subsequent inquiry. *In summary, earlier discussions focused on the clear emerging patterns for the fast-moving planets (presumably because these were pedagogically advantageous for the teacher), and later ones addressed the problematic circles, whose evidentiary body was sparse.*

Case Two - Wallcology

Methodology

Research Setting and Design Considerations

The second EPIC case employed the WallCology EP as the setting for whole-class inquiry, targeting life sciences topics of biodiversity and population ecologies (Moher, Uphoff, Lopez-Silva, & Malcolm, 2008). Over several weeks, grade five/six students observed a digital ecosystem consisting of dynamic animations of insects and vegetation, visible through display monitors called "Wallsopes." (see Figure 3). The ecosystem comprised four differentiated but interconnected habitats, one on each wall of the classroom, which varied in terms of environmental conditions (temperature, light and humidity). In our EPIC activity, students made observations about the morphologies and behaviors of organisms to determine their life cycle relationships. Constructing a representation of the lifecycles of any species was a challenging task; it was not always clear which organism belonged to which species (e.g., does the adult form of the "green bug" hatch from the white egg or blue egg?). It required careful observation (and maybe a bit of luck) for students to actually "see" life events unfold (e.g., laying and hatching). Additionally, since each monitor displayed a different habitat, students at one monitor see something different than students at another monitor, necessitating the sharing of observational data.

In previous implementations of WallCology (see Moher et al., 2008) students had recorded their notes and sketches in paper-based field guides, which formed the basis of whole-class discussions. The goal of the EPIC design innovation was to create a more powerful means of sharing and working with these observations, by aggregating individual or group inquiry actions, encouraging teacher and students to attend to interesting patterns in the data, including where more work is needed. Ideally, our aggregate views would reveal the most constructive patterns. We designed several applications (e.g., a graphing tool, and a modeling tool), however the present paper analyzes the role of one tool called "Life Cycle Relationships" and a corresponding aggregate representation that tallied all the pair-wise relationships observed by students using the tool.

Participants

Forty-two students from two grade five/six classes participated, with the classroom teacher guiding each class. The present analysis is restricted to one of the two groups (n=21) and to one lesson within the nine-week unit in which students contributed observations concerning life cycles.

Method

Here, we analyze how the aggregate representations of students' observations of the lifecycle relationships of one species – the "blue bug" – helped them come to evidence-based agreements about its complete life cycle. Students had already formed two competing theories about the "blue bug's" life cycle in a previous class, and they were interested in pursuing this line of inquiry. We hoped that such disagreements could be resolved by teacher-led discussions of the aggregate view. In this example, we hoped the teacher could help students understand which additional observations would be needed about the "blue bug's" life cycle, and help the classroom community procure those observations and come to a consensus.

Six pairs of students and 8 single students used the tablet application to contribute life cycle observations. To contribute a life cycle observation, students selected two icons on the tablet interface — the first icon, to indicate an early life cycle stage (e.g., an egg) and the second icon, to represent the stage that immediately follows the first icon (e.g., the larva form), as shown in Figure 3. To stipulate that no life cycle stage immediately precedes the second icon (e.g., a vegetation icon), students placed the "X" icon in the first space. By tapping a life stage icon in the collective dataset view, a student could view tallies of observations concerning that organism. For each icon, there are eleven possible relational statements (e.g., "x organism precedes y organism" in a life cycle). By choosing one icon from the set of 11 icons, students and teachers view up-to-date relational tallies for the selected organism and the stage immediately preceding it, shown in Figure 3.



Figure 3. Left: Students observe WallCology ecosystem. Middle: Tablet interface used to contribute a life cycle observation. Right: Life cycle aggregate displays tallies of every observation made for the selected organism.

Data Sources

Data sources included data logs of student contributions, a video recording of the learning activity, researchers’ field notes, and focus group interviews with teachers and students following the sessions.

Analysis and Findings

Prior to this class, students had contributed 75 life cycle observations, with a 62% accuracy rate. By the end of this class period, 201 life cycle observations had been contributed, with a total accuracy of 72% -- more than doubling the observations in the evidentiary database and resulting in a 16% increase in accuracy. Observations concerning the life stages of the “blue bug” were entered at a higher rate, with 37 existing in the knowledge base before the class and 113 by the end of the period. The total number of correct observations about the life stages of the “blue bug” increased threefold, and the total number of incorrect observations about it increased by 69%.

We applied the coding scheme from case one (see Table 2 for definitions of codes) to the discourse of case two, as shown in Table 2. These codes were again plotted on a timeline (shown in Figure 4), which also shows where whole class discussions occurred within the context of the learning activity and the point at which the knowledge community reached agreement about the life cycle of the “blue bug”.

Table 2: The table provides a description of the codes with an example for each, taken from the transcript.

Code Name	Example from transcript
Summative statement	“And three people said the shrimp turned into the blue beetle.” - <i>Student</i>
Interpretive statement	“I don’t know if six people meant that the blue beetle hatches <i>from</i> the egg or <i>laid</i> the egg...” - <i>Student</i>
Directional statement	“Can we just check that that’s observed by other observers?” - <i>Teacher</i>
Negotiation statement	“So could we say for today that our current understanding of this is that [the life cycle begins with the blue egg?]” - <i>Teacher</i>
Agreement statement	YES!!! - <i>Students</i>

Patterns in Coded Discourse: Blue Bug discussion.

The timeline shows that the focus of the first whole class discussion was on the students’ disagreement about the “blue bugs” life cycle. The teacher said, “So what it seems like, if I’m reading the flavor of the room right, is that people have different understandings and have observed different things about [the cycle of the blue bug].” The timeline indicates that five minutes before the end of the class, the issue was resolved when the knowledge community came to agreement. The coded discourse reveals three interesting patterns: (1) A directive statement was always made by the teacher (pink) when the aggregate screen was first displayed. For example, “Can someone give words to this set of data - maybe even in words like ‘most’, ‘some’, like ‘most observations’, ‘some observations’. What would you say, [student]?” (2) Students made summative (yellow) and interpretive (orange) statements following the teacher’s directive statements. Viewing the tally for the “blue bug’s” larvae hatching from its egg (24 observations), a student said, “I think almost everybody thinks that the ‘shrimp

hatches from the egg”. Interpretative statements outnumbered summative statements, by nearly 50%. (3) The activity culminated with many negotiation statements followed by one final agreement statement: “YES”.

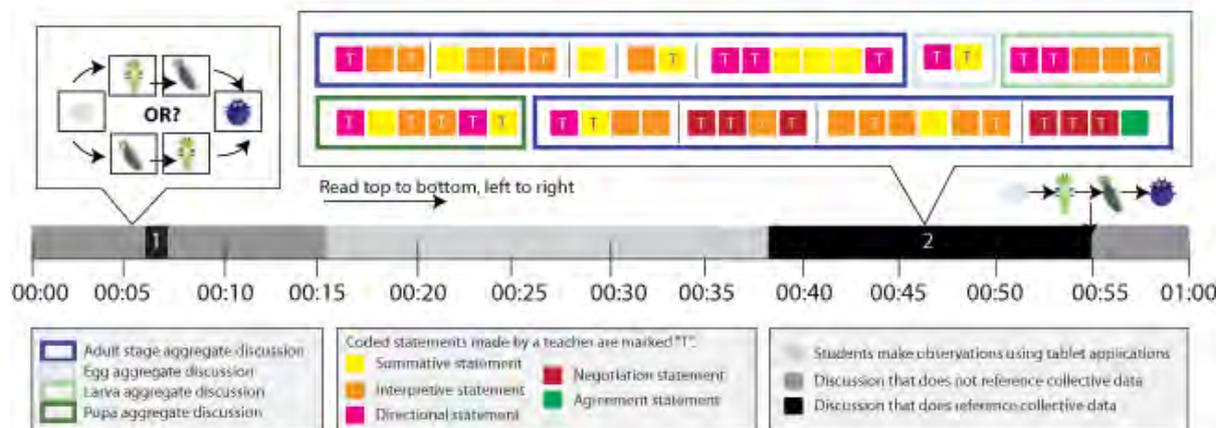


Figure 4. Analysis codes are contained within a colored outline (representing the aggregate screen that was on display, e.g., egg, larva, pupa, or adult). The vertical lines demarcate sub-discussions, regarding a particular pair of life stages shown within the aggregate.

Patterns of Interaction

The teacher began the discussion by presenting the aggregate view of the adult stage of the life cycle. This revealed that 15 observations were made that the pupa transformed into the adult form, while only three observations were made that the larva transforming into the adult form. The teacher put forward the following hypothesis: egg \neq larva \neq pupa \neq adult and said, “Can we just check whether that’s confirmed by other observers?” This statement provided direction for the presentation order of the aggregate screens: egg, larva, pupa, and adult. Student-contributed data was used to confirm the hypothesis, logically and systematically, resulting in resolution.

Discussion

In both cases, our designs provided the knowledge community with an overview of the “lay of the land” of their observations: a quantitative display of what the whole class had observed. Highlighting disagreement and gaps in the dataset made it possible for the community to pursue their inquiry goals, through productive discourse and by filling in obvious gaps through targeted observations.

There were also some differences that appeared between the discourse-analysis in the two cases. While the discourse codes applied equally well to both cases, there were some differences in their distribution. In the first case, many more summative statements were made, possibly because those students had never seen such a display (i.e., of binary data in table form) before. In the second case, students had already used the Life Cycle Relationships Tool in a previous class. Also, in the first case, agreement was reached several times (with negotiation and agreement statements interspersed throughout the timeline), whereas in the second case, only one issue was under discussion, and those codes appeared only at the end of the timeline.

In a follow-up interview with the teacher after the WallCology unit, he spoke of the power of aggregated data in this type of scientific inquiry. He said such representations really helped his students to consider data, as opposed to only pursuing theories (which is something his students are more comfortable with). He noted that without technology, it might be possible (although awkward) to construct such data sets, but that with networked technology it is possible to “[have] a set of data that people have contributed to, [which is] then accessible to the community in their work.” He said students felt a greater sense of accountability to each other, because they knew they were dependent on the aggregate for their information.

References

- Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Brown, A. L., & Campione, J. C. (1996). Psychological theory and 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: Lawrence Erlbaum Associates.

- Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications, Incorporated.
- Dufresne, R.J., Gerace, W.J., et al. 1996. Classtalk: A classroom communication system for active learning. *Journal of Computers in Higher Education*, 7, 3-47.
- Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences*, 8(3/4) 391-450.
- Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction*, 18(4), 495-523.
- Moher, T. (2006). Embedded Phenomena: Supporting Science Learning with Classroom-sized Distributed Simulations. *Proceedings ACM Conference on Human Factors in Computing Systems* (pp. 691-700). Montréal, QC.
- Moher, T., Uphoff, B., Bhatt, D., López Silva, B., and Malcolm, P. (2008). WallCology: designing interaction affordances for learner engagement in authentic science inquiry. *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems (CHI)*, (pp. 163-172). New York, NY: ACM.
- Moher, T., & Slotta, J. (Chairs). (2012). Embedded Phenomena for Knowledge Communities: Supporting complex practices and interactions within a community of inquiry in the elementary science classroom. *Proceedings of the International Conference of the Learning Sciences: Future of Learning*. (pp. 64-71). International Society of the Learning Sciences.
- Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, 49(1), 345-375. Roschelle, J., Patton, C., & Tatar, D. (2007). Designing Networked Handheld Devices to Enhance School Learning. In M. V. Zelkowitz, Ed. *Advances in Computers*, 70, 1-60.
- 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). International Society of the Learning Sciences.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York, NY: Cambridge University Press.
- Slotta, J. D. (2010). Evolving the classrooms of the future: The interplay of pedagogy, technology and community. In K. Mäkitalo-Siegl, F. Kaplan, J. Zottmann & F. Fischer (Eds.). *Classroom of the Future. Orchestrating collaborative spaces*. (215-242). Rotterdam: Sense.
- Slotta, J. D., & Peters, V. L. (2008). A blended model for knowledge communities: Embedding scaffolded inquiry. *Proceedings of the International Conference for the Learning Sciences* (pp. 343-350). International Society of the Learning Sciences.
- Slotta, J. D., & Linn, M. C. (2009). *WISE Science: Web-based Inquiry in the Classroom. Technology, Education — Connections*. New York, NY: Teachers College Press.
- Slotta, J. D., Tissenbaum, M., Lui, M., & Zukowski, M. (2012). Smart Classrooms for Knowledge Communities: EPIC Technology Environment. *Symposium conducted at the 10th International Conference for the Learning Sciences (ICLS) Conference*, Sydney, Australia.
- Tatar, D., Roschelle, J., Vahey, P., & Penuel, W. R. (2003). Handhelds to school: lessons learned. *Computer*, 36(9), 30-37.
- Tinker, R. (2000). A history of probeware. Retrieved May 30, 2012, from <http://makingsens.stanford.edu/pubs/AHistoryOfProbeware.pdf>
- Thompson, M. and Moher, T. (2006). HelioRoom: Problem-solving in a whole class visual simulation. *Proceedings of the International Conference of the Learning Sciences*, (pp. 1000-1001). International Society of the Learning Sciences.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Acknowledgements

We acknowledge members of the Learning Technologies Group at the University of Illinois at Chicago and members of the Encore Lab at the Ontario Institute for Studies in Education in Toronto, who were integral to these projects. We sincerely thank the teachers and students from the Dr. Eric Jackman Institute for Child Study who participated. The material presented here is based on work supported by the U.S. National Science Foundation under grant IIS-1065275 and Canadian Social Sciences and Humanities Research Council under grant 410-2011-0474.

Making use of collective knowledge – a cognitive approach

Ulrike Cress, Knowledge Media Research Center, Schleichstr. 6, 72076 Tübingen, Germany,
u.cress@iwm-kmrc.de

Abstract: From a cognitive perspective, *knowledge* resides in people's minds., and there is no conceptualization of 'collective knowledge'. In the socio-cultural approach the concept of collective knowledge is central. The Co-Evolution Model of Individual Learning and Collaborative Knowledge Building (Cress & Kimmerle, 2008; Moskaliuk, Kimmerle & Cress, 2009, 2012) combines both approaches and considers internal-individual and external-collaborative processes that take place when people work on a shared artifact. We apply this framework to social tagging and explain how tag clouds represent collective knowledge. Referring to the Information Foraging Theory (Pirolli, 2007; Pirolli & Card, 1999) we show how people make use of collective knowledge when navigating with tag clouds. We give an overview of several experimental studies that induce situations where individual and collective knowledge contradict each other. The results show that in such situations incidental learning takes place, and users' individual conceptual knowledge assimilates to the collective conceptual knowledge.

Introduction

As long as CSSL has existed as a research topic, there has been discourse about whether cognition and knowledge are bound to individual minds, or if they also describe group phenomena (Koshmann, 1996; Roschelle, 1996; Stahl, 2005). In this dialogue, the cognitive research tradition is based on the information-processing approach and states that internal mental presentations provide the 'substrate' of knowledge. Thus, this position denies that knowledge can exit outside a person's mind. Opposite to this point of view, the socio-cultural tradition sees the substrate of knowledge in situations, social interaction and cultural affordances. With this understanding, people do not have knowledge or acquire knowledge but instead participate in social practices which enable knowing (Sfard, 1998). The *Co-Evolution-Model of Individual Learning and Collective Knowledge Building* (Cress & Kimmerle, 2008), which is shown in Figure 1, integrates both perspectives and describes how individual and collective knowledge develop when people work with shared artifacts.

The co-evolution model combines a systemic and a cognitive perspective, conceptualizing users as cognitive systems and the collaborating group as social system. Each system has its specific mode of operation. A cognitive system operates by cognitive processes such as perceiving, thinking, or problem solving. These processes take place within the individuals. They are described in detail by the information-processing approach. The social system, which comes into existence whenever people behave and communicate in a stable and expected manner, operates according to rules or social norms (cf. Luhmann, 1984). When people collaboratively work with shared artifacts, the operations of the social system become manifest and observable. In Wikis, for example, text passages provided by users are deleted, revised or interlinked with other text passages. Over time in the Wiki a coherent text develops, where the single contributions of different users become indistinguishably interwoven (Kimmerle et al, 2012). These processes are social in nature, and they happen according to rules and norms of the group. In the Wikipedia, e.g. these rules are explicit (Oeberst, Halatchlyiski & Cress, resubm). In other communities they may be implicit, but nevertheless they determine how individuals deal with the contributions of others (Kimmerle, et al, in press).

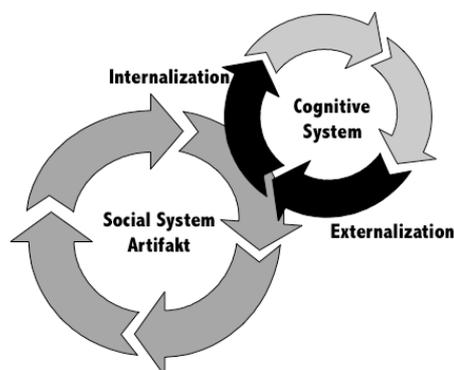


Figure 1: Co-Evolution Model

The Co-evolution model states that if people work on a shared artifact, the social and cognitive systems influence each and dynamically co-evolve. This takes place through processes of *externalization* and

internalization. An individual externalizes his or her own (i.e., internal) knowledge and conveys it into the shared artifact. There, it is processed according to the system's rules. If it is relevant for the system, it becomes part of the collective knowledge. This collective knowledge is an emerging phenomenon. It results from people's single activities and contributions, but it is not just an aggregation of people's individual knowledge. The individual contributions are further processed and integrated within the artifact. The resulting collective knowledge exists only in the artefact, and thus outside of people's minds. By working on the artifact a user may process and internalize it. Through these exchange processes, the cognitive systems as well as the social system develop.

The model states that incongruities between (individual) knowledge in the cognitive system and (collective) knowledge in the artifact trigger this co-evolution (Moskaliuk, Kimmerle & Cress, 2009, 2012). For a user, the incongruity leads to a cognitive conflict. One possibility to solve this conflict is that users work on the artifact and make it fit their own knowledge structure (equilibration through externalization). The other possibility is that users solve this conflict by adapting their own cognitive structures to the conceptual structure of the artifact (equilibration through internalization).

Up to now the model has been applied to small and large groups working with wikis (Moskaliuk, et al., 2011; Moskaliuk, Kimmerle & Cress, 2009; Kimmerle, Moskaliuk, Harrer & Cress, 2010), and to knowledge creation in Wikipedia (Oeberst, Halatchlyiski & Cress, 2012). In this paper we apply the Co-Evolution Model to social tagging systems. Social tagging systems are web 2.0 tools that enable users to annotate digital resources with individually chosen tags. The tags of all users are aggregated and can be visualized by tags clouds. As social tagging systems are highly analogous to human memories (both can be described with models of spreading activation), they are a good example to describe co-evolutional processes on a fine-grained and cognitive level.

The structure of the paper is as follows: First we show why tag clouds are external representations of collective knowledge, and why they have a structural similarity to a human's individual internal knowledge. We then focus on internal cognitive processes by referring to a prominent cognitive theory about web navigation (Information Foraging Theory: Pirolli, 2007; Pirolli & Card, 1999). This theory describes how individual knowledge determines which links people select when they navigate on the Web. We expand this model by considering the influence of the collective knowledge inherent in tags. We give an overview of several experimental studies that induced situations where individual and collective conceptual knowledge contradicted each other. All studies confirm the extended model and show that users make use of the collective knowledge and internalize it. Their own knowledge assimilates to the collective conceptual knowledge – just by navigation, and without any intention to learn. This confirms the assumption of the Co-Evolution Model.

In sum, this paper gives evidence that even from a cognitive point of view it make sense to conceptualize collective knowledge, to consider it as an emerging phenomenon that exists outside people's heads but influences people's individual knowledge.

Structure of Collective Knowledge created in Social Tagging Systems

Social tagging is an activity of annotating digital resources, for instance, bookmarks (e.g., delicious.com), pictures (e.g., flickr.com), blogs (e.g., Technorati), or products (e.g., on amazon.com) with 'tags' (cf. Golder & Huberman, 2006; Trant, 2009). In most applications, a user can choose individual tags for stored resources. So a tag reflects a user's internal association with a resource and represents the specific meaning or relevance for the respective user. On this individual level tags are metadata that help individuals to structure, organize, and re-find their own stored Web resources. If people tag resources, they externalize their individual associations.

Social tagging systems extend this individual level to a collective level. They aggregate the tags of all users and enable the creation of a folksonomy (Trant, 2009; Vander Wal, 2005). This folksonomy results from the tripartite network among users, tags, and resources and enables detecting similarities. For example, resources that are frequently annotated with the same tag are somewhat similar; and different tags that co-occur frequently across different resources or users indicate that they have something in common.

Figure 2 visualizes these processes: Figure 2a shows the tripartite network of resources and tags assigned by two users. Figure 2b shows the one-mode network tag-tag relation network derived from the two-mode resource-tag network. The nodes in the tag-tag one-mode network represent tags, the linking lines represent resources that the tags have in common.

User X tags resource 1 with the tags b,c and d, therefore these tags are connected in Figure 2a. Since both users X and Y annotate the resources 1 and 3 with the same tags c and d, the respective link has a higher weight, and the association strength between the tags c and d is higher.

The frequency of co-occurrence of two tags across all resources determines the *association strength* between these two tags (different weighting measures are discussed in Markines et al., 2009). So what results from the tagging activity is a semantic network. It shows how tags are semantically related to each other on the basis of a common set of resources. In Figure 2b it is obvious that the tags b,c and d have a strong semantic

relation, especially the tags c and d. This semantic knowledge just emerged from of the aggregation of tags. Users X and Y have different knowledge structures compared to the aggregated knowledge structure.

A common way of visualizing the association strengths between tags is the use of tag clouds (one is presented in Figure 6 later on in this paper). Tag clouds present those tags with the strongest associations to the search term: the stronger the association strength, the larger the font size of the tag. This means for our example in Figure 2: If one would search for c, a tag cloud would present a,b and d with d with the largest font size.

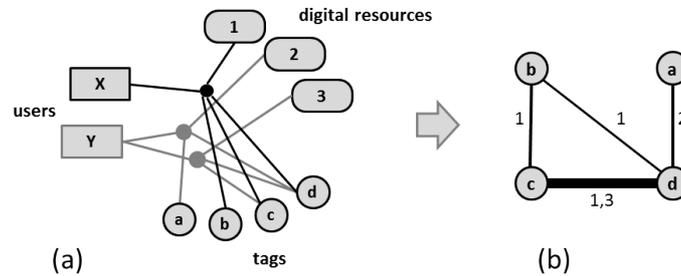


Figure 2: Example for social tagging: (a) Two users annotate tags a-d for resources 1-3. Social Tagging Systems make use of the information in this tripartite network. (b) They transform the tripartite graph in a one-mode network representing weighted tag-tag relations. Tags a-d are represented as nodes, linking lines represent resources that tags have in common. Numbers on the links indicate the number of the resource in (a).

The way social tagging systems create collective knowledge out of single contributions is highly analogous to processes in the semantic memory of individuals. We explain this conceptual correspondence between individual internal knowledge and collective external knowledge in the following section.

Structure of Individual Knowledge

A variety of cognitive models describe declarative knowledge in human memory as a network of *chunks* (e.g., Anderson, 1983; Collins & Loftus, 1975). Each of the chunks is connected to other chunks with a different strength of association. The strength of association derives from people’s past learning experiences. When two chunks frequently co-occur in a meaningful context, their association becomes stronger. If, e.g. “sun” often co-occurs with “Florida”, a strong association between these two chunks is established.

The strength of association determines the retrieval of chunks in semantic memory: in order to retrieve a chunk it has to be activated by other chunks. The activation spreads from one chunk to another. The stronger the association, the higher is the likelihood that a chunk will stimulate a certain level of activation. Figure 3 shows an example where somebody is asked where she would like to go on holiday. Through this question the chunks Florida and Himalaya are activated, and activation spreads to all linked chunks according to the association strengths.

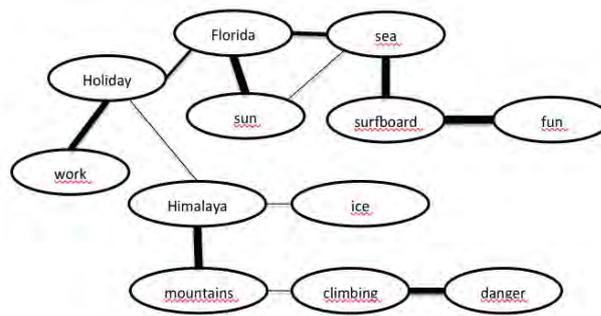


Figure 3: Activation in semantic memory

Similarity between Social Tagging and Individual Knowledge

Already the visualizations in Figure 2 and Figure 3 demonstrate the structural similarity of social tagging systems and human memory. Both systems have a network structure that is the basis of semantic knowledge. In the individual (Figure 3) case this knowledge is created through experiences made in the past, in the social case (Figure 2) it is created through people’s tagging of resources and the automated aggregation of these tags. And both systems develop analogously: An individual’s knowledge develops whenever two chunks are activated simultaneously: Then a relation is created, or an already existing one is strengthened. The collective knowledge develops whenever an individual tags a resource and thus creates a new relation or strengthens an already existing one.

Interaction of collective and individual systems

The Co-Evolution model states that collective and individual knowledge systems interact by two processes: Externalization and internalization. This paper focuses on internalization. It shows how people make use of the collective knowledge in information search and how this leads to individual learning processes.

We focus on a situation where people navigate through tagged resources, and we base our considerations on the most prominent model about Web search, the Information Foraging Theory (Pirolli, 2007; Pirolli & Card, 1999). This theory states that people's knowledge about the searched topic plays a crucial role in a user's navigation. When navigating through resources, a user has to decide which links or tags lead to the topic he is interested in. This decision is based on a user's association strength between chunks activated by the link and those activated by the search topic (topic of interest). This strength of association predicts that (subjective) probability that a link will lead to the desired topic. In the Information Foraging Theory this probability is called *information scent*. This information scent is the association strength between the terms used in a link and the searched topic (topic of interest in Figure 4). The stronger a link is associated with the topic of interest the higher is its information scent. This means that the individual knowledge determines which link a user selects. In Figure 4a the user would select tag *a* because it has the highest association strength with the desired topic.

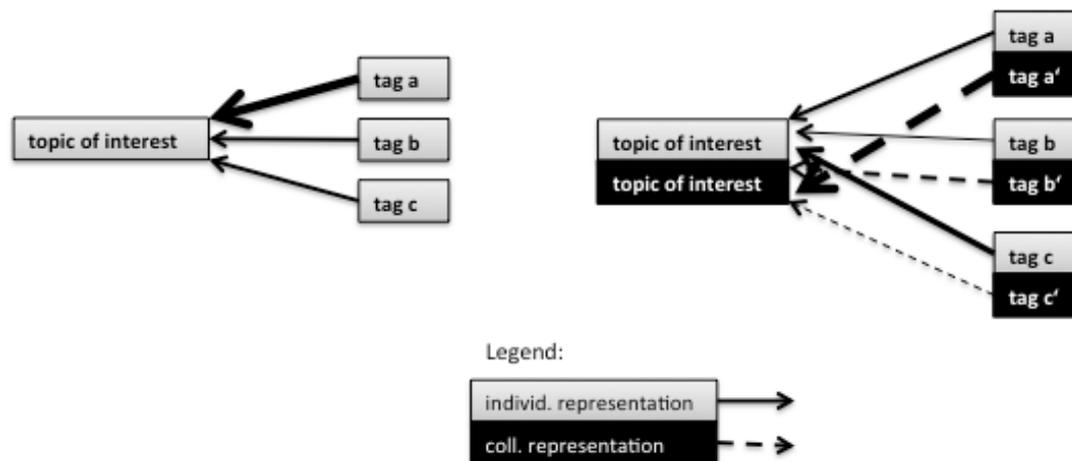


Figure 4a: Information-scent model (Pirolli, 2007)

Figure 4b: Collectively extended information scent (Cress, Held & Kimmerle, 2013)

Several studies have confirmed the information scent model and have shown that people's knowledge determines their Web navigation (e.g., Blackmon et al., 2002; Fu & Pirolli, 2007). We build on these results and ask whether people make use of the collective knowledge when they navigate in a tagging environment. Taking co-evolutional processes between individual knowledge and collective knowledge into account, we propose an *extended information scent model* as it is shown in Figure 4b. This model states that the information scent is a linear combination of a user's individual knowledge structure and the collective knowledge structure. The individual knowledge structure (displayed in light grey) is based on individual representations (chunks), the collective knowledge structure (dark grey) is based on tags and the folksonomy defining their interrelations. Individual and collective knowledge do not need to be consistent. The association strengths as well as the individual and collaborative representations may differ. The model states that the 'extended information scent' is a linear function of the individual and the collective association strengths. When users use tags for their navigation, individual learning processes take place. Then the internal representation of the topic of interest assimilates to the external and collective representation given by the folksonomy. The cognitive system in turn develops, and the internal knowledge structure assimilates to the collective knowledge structure.

Experimental Evidence for an Extended Information Scent

In the following sections we show three experiments we conducted in our Lab that give evidence for this model. We present here only the main results. Many more details about the materials, procedure and results with regard to other dependent variables are described in Cress and Held (in press), Held, Kimmerle and Cress (2012) and Cress, Held and Kimmerle (2013).

All three experiments induced some incongruity between the individual's knowledge strengths of associations and the collective strength of associations in topic domains previously unknown to the participants. The incongruity was established by (1) manipulating a user's individual knowledge through providing information *before* the navigation task and/or by (2) manipulating the collective knowledge *during* the

navigation task (through manipulation of the tag clouds). In the first experiment, people could navigate in the Web. In the other two experiments, an experimental setting was used where users navigated in a highly controlled but artificial tagging scenario. In all three experiments, we expected to confirm the model of the extended information scent shown in Figure 4. So we expected that the participants' prior knowledge as well as the collective knowledge inherent in the tag clouds would influence people's knowledge after they had interacted with the tag system.

1st Experiment: Navigation with the Tool Brower-extension "Search Cloudlet"

Materials and Procedure In the first experiment (Cress & Held, in press), participants could freely navigate in the Web. In order to be able to manipulate the participants' prior knowledge, we selected three topic domains that were unknown to the participants and for which we could easily induce incorrect prior knowledge. The domains were 'EMDR', 'Dannsa Biodag' and 'Manipogo'. The participants received, for example, the information that *Manipogo* is a Golf festival at the Manitoba Lake, whereas in fact it is a monster in the Manipogo Lake. The participants' task was to find more information about the three topics by navigating through the Web. The Firefox browser extension 'Search Cloudlet' was used for navigation. Cloudlet automatically creates tag clouds from the Google result list. It visualizes those words that are part of many search results with a larger font size. So the tag clouds represent a kind of aggregated knowledge of the Web with respect to the topics of interest.

During the task, the participants could not see the search results, they saw only the tag clouds. Figure 5 shows the tag cloud for the search term 'Manipogo Manitoba Lake'. It shows a strong association with *monster* and a smaller with *creature* or *Loch* and *Ness*, but only a weak association with *golf*.



Figure 5. Tag clouds for the search term 'Manipogo Manitoba Lake'.

Design The experiment implemented a 2x3 factorial design. The *individual's knowledge* was manipulated as within-factor: Before the navigation, each user received information that was either congruent with the collective knowledge (e.g. *Dannsa Biodag* is a North European war dance), or incongruent with the Web resources (*Dannsa Biodag* is a South American war dance), or they received no information. This within-factor was permuted. The second factor *collective knowledge* served as between-factor. The experimental condition could use the Cloudlet tool and could enter search terms. For each search term they received a tag cloud based on all information in the Web. So these tag clouds are representations of the Web knowledge. The control group did not receive any tag clouds. So the control group had no access to the collective knowledge and they could not navigate in any way. This group served as a treatment check for the manipulation of the individual knowledge, it had now access to any further knowledge.

At the end of the experiment all participants had to complete a post-test that tested their knowledge about the three domains.

Results The results of the study (n=54) confirmed the model of the extended information scent. Collective knowledge inherent in tag clouds and the individual knowledge people had built through the information provided before the navigation had a significant effect on people's knowledge scores in the post test (collective knowledge: $F(1, 52) = 22.94, p < .001$; individual knowledge $F(2, 104) = 23.45, p < .001$). If tag clouds gave users access to the collective knowledge, people internalized this knowledge and adapted their own cognitive structures to it. The two independent variables did not interact.

Discussion The results confirmed the model of the extended information scent. It shows that people make use of the collective knowledge in social tag clouds. Their individual association strengths adapted to the collective knowledge. As this first experiment was done in the real Web, the collective knowledge presented in tag clouds resulted from the majority of the search results. This means that in this experiment the association strengths could not be systematically manipulated for each single tag. The following two experiments were designed to go more into detail and measure the influence of single association strengths between a tag and a concept.

2nd Experiment: Navigation with weighted/unweighted tags

Materials and Procedure The second experiment (Held, Kimmerle & Cress, 2012) was done in a controlled setting where all participants were provided with tag clouds. With the domain 'Georgian Wine' (typical

Georgian wine regions, Georgian grapes and Georgian wine aromas), we chose a knowledge domain about which people had no prior knowledge.

The experiment was set up online. Participants were told that the overall goal of the task was to investigate how people search for products with the help of tags. We did not inform the participants that the experiment was intended to measure any kind of learning. The experiment started with a general introduction to tags, and the participants were informed that the tag clouds a user would encounter were based on tags of wine experts. The participants were instructed to find *typical Georgian wines* to build up a presentable wine cellar.

During the experiment the participants received tag clouds. The tag clouds showed representative tags related to, for example, specific wines or parts of Georgia. Figure 6 shows such a tag cloud. It shows a tag cloud for the search “Georgian wine regions”.



Figure 6: Tag cloud for the search ‘Georgian wine regions’.

For each presented tag cloud the participants had to click on that tag which seemed most appropriate for leading them to typical Georgian wines. After navigating through nine tag clouds, the participants had to complete a post-test. They had to indicate how strongly they associated specific tags with Georgian wine by rating their typicality (e.g., ‘How typical is the wine region ‘Kakheti’ for Georgian wine?’) on a rating scale of 1 to 7 (from very untypical to very typical.).

Design The experiment implemented a mixed 3x2 factorial design. Analogous to the first study we induced *individual knowledge* by providing some information before a participant could navigate. For it, we presented a Weblog where an anonymous user told participants that ‘Kakheti’ was a typical Georgian wine (which is congruent with the collective knowledge shown in Figure 6), or the information, that ‘Tsageri’ was a typical Georgian wine. In a third condition we presented no such information. Analogous to the first experiment we manipulated the factor *individual knowledge* as within-factor (permuted across the three domains (wine regions, wine grapes and wine aromas)). The between-factor *collective knowledge* had two levels: the tag clouds a user received had weighted tags (tag sizes varied according to their collective association with the search term), or the tags were not weighted. The weighted tag cloud made the collective association strengths visible, whereas the unweighted tag clouds did not deliver this information. As dependent variable we analysed people’s scores of the *knowledge test* provided after the experimental task.

Results. The study (n=207) revealed the expected main effects: Participants’ post-test scores were not just influenced by their *individual knowledge* $F(2, 352) = 40.34, p < .001$ but also by the *collective knowledge*, $F(1, 176) = 12.77, p < .001$. Also here we found no interaction.

So also the second experiment confirmed the model of the extended information scent: individual knowledge as well as collective knowledge influenced people’s knowledge test after the navigation.

3rd Experiment: Fine-grained Variation of Association Strengths

The third experiment (Cress, Held & Kimmerle, 2013; 1st experiment) varied individual knowledge and collective knowledge in a much more fine-grained manner than in the experiments before. It manipulated the association strengths of individual and collective knowledge in a linear way.

Materials and Procedure The experiment used the same domain as the second one, ‘Georgian Wine’. In the first phase of the experiment we manipulated the individual knowledge much more implicitly than in the second experiment. We provided participants with a wine list given from somebody ‘who loves Georgian wines’ and asked them to provide feedback on design features of this list. It was not mentioned in any way that participants should memorize any content of the wine list, nor were the participants informed that the content had any specific relevance for further steps of the task. The list was presented to the participants for 30 seconds, followed by five general questions about the design and information of the list (e.g., ‘Would it be helpful to provide further information on specific wine regions?’) in order to direct attention to the content of the list. The wine list was still available to the participants while they were answering the questions. The *second phase* of the experiment was a navigation task like the one in the second experiment. In this phase, participants had the task of collecting typical Georgian wine. After a basic introduction to social tags, participants were presented tag clouds and asked to click on that tag of each cloud which would lead them to a typical Georgian wine. After this task, the participants had to complete the same post-test we used in experiment 2.

Design A 5 x 4 between-subjects design was used. We manipulated a user's *individual strength of association* between the wine region 'Kakheti' and 'Georgian wine' by varying the content of the wine list that was presented in the first phase of the experiment. This wine list showed five Georgian wines. Wines from the region 'Kakheti' were part of the list (1) not at all, (2) once, (3) twice, (4) three times, or (5) four times.

We manipulated the collective strength of association by changing the size of the tag 'Imereti' in the tag clouds, which the participants encountered in the second phase of the experiment. The *size 'Imereti'* had four continuous levels: (1) the tag 'Imereti' had the same size as the tag 'Kakheti' (2) or was 33%, (3) 67%, or (4) 100% larger. No other tags varied in size.

As dependent variables we used ratings in the post-test, where we asked how typical 'Kakheti' and 'Imereti' are for Georgian wines.

Results The study was done with $n=596$ participants. In order to test the impact of the individual strength of association and the collective strength of association on people's resulting knowledge, we conducted multiple regression analyses. The predictor variables (*frequency Kakheti in the wine list* and *size Imereti in the tag cloud*) were centered, and the interaction term was computed by a multiplication of both variables. The regressions revealed that both independent variables had the expected effect on user's association strengths after the experiment (association for Kakheti: effect of individual knowledge $\beta = .25, p < .001$; effect of collective knowledge $\beta = -.10, p < .05$ and association for Imereti: effect of individual knowledge $\beta = -.14, p < .01$; effect of collective knowledge $\beta = .21, p < .001$). We found no significant interactions.

So in sum the results also confirm the proposed model. The linear regressions show that both the individual association strengths as well as the collective associations strengths have an additive linear effect on people's knowledge.

General Discussion

The use of tagging systems provides valuable insights into the exchange processes between individual and collective knowledge. The Co-Evolution Model states that within the social system, social processes occur that create a collective product out of the individual contributions. With regard to social tagging systems, this is primarily an automatic process. The only social rule is that the users are expected to tag resources with tags that are meaningful for them, and that have some relation to the resource. If users consider these rules then the tagging algorithm leads to a semantically meaningful network of tags, which represents the collective knowledge. A tag provided by an individual user then builds new associations within this collective knowledge or changes the association strengths of existing ones. Thus the externalization of an individual directly changes the knowledge within the social system.

What our experiments showed is the fact that users internalize this collective knowledge when they navigate with tag clouds. All three experiments confirmed the model of the extended information scent: When people navigate with tag clouds, their navigation is not only influenced by their individual knowledge, but also by the collective knowledge inherent in the tag clouds. This shows that people make use of the collective associations represented in the clouds. And this process happens incidentally during navigation, and without any explicit instruction to learn something about the topic. Across all three studies the effects were very congruent: we found main effects for individual and collective knowledge but no interaction. In all experiments the effect of people's individual knowledge was at least as strong as the effect of the collective knowledge. This is remarkable, because in our experimental setting, people's individual knowledge was just based on the information of an "anonymous blogger in a Weblog", whereas the collective knowledge was based on tags 'from wine experts'. Thus, objectively seen, the collective knowledge should have been much more credible than the individual knowledge. It seems, that once an association is established, it has a high impact. The next step in research will be to examine the variables that determine what influence collective knowledge has compared to individual knowledge.

In sum, the data strongly support the model of the collectively extended information scent. This describes the internalization processes at the cognitive level and thus provides a first fine-grained model to demonstrate co-evolutional processes as described by the co-evolution model. With regard to their formal structure, information in tagging systems can be considered as collective knowledge. Further experiments will have to prove, if users consider tag clouds as information coming from a collective, or if they just make use of it because they provide some additional information.

References

- Anderson, J.R. (1980). *Cognitive psychology and its implications*. San Francisco, CA: Freeman.
- Blackmon, M.H., Polson, P.G., Kitajima, M., & Lewis, C. (2002). Cognitive walkthrough for the web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 463-470). New York: ACM Press.

- Collins, A., & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Cress, U., & Held, C. (in press). Harnessing collective knowledge inherent in tag clouds. *Journal of Computer-Assisted Learning*.
- Cress, U., Held, C., & Kimmerle, J. (2013). The collective knowledge of social tags: Direct and indirect influences on navigation, learning, and information processing. *Computers & Education*, 60, 59-73.
- Cress, U., & Kimmerle, J., (2008). A systemic and cognitive view on collaborative knowledge building with wikis. *International Journal of Computer-Supported Collaborative Learning* 3 (2), 105-122.
- Fu, W., & Pirolli, P. (2007). SNIF-ACT: A cognitive model of user navigation on the World Wide Web. *Human-Computer Interaction*, 22(4), 355-412.
- Golder, S., & Huberman, B.A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2), 198-208.
- Held, C., Kimmerle, J., & Cress, U. (2012). Learning by foraging: The impact of individual knowledge and social tags on web navigation processes. *Computers in Human Behavior*, 28, 34-40.
- Kimmerle, J., Cress, U., & Held, C. (2010). The interplay between individual and collective knowledge: Technologies for organisational learning and knowledge building. *Knowledge Management Research & Practice*, 8, 33-44.
- Kimmerle, J., Thiel, A., Gerbing, K.-K., Bientzle, M., Halatchlyski, I., & Cress, U. (in press). Knowledge construction in an outsider community: Extending the communities of practice concept. *Computers in Human Behavior*.
- Kimmerle, J., Moskaliuk, J., Harrer, A., & Cress, U. (2010). Visualizing co-evolution of individual and collective knowledge. *Information, Communication and Society*, 13, 1099-1121.
- Kimmerle, J., Moskaliuk, J., Bientzle, M., Thiel, A., & Cress, U. (2012). Using controversies for knowledge construction: Thinking and writing about alternative medicine. *Proceedings of the 10th International Conference of the Learning Sciences* (pp. 1-8). Sydney: International Society of the Learning Sciences.
- Koschmann, T. (1996). Paradigm shifts and instructional technology. In T. Koschmann, CSCL: Theory and Practice of an Emergin Paradigm, pp. 1-23. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Luhmann, N. (1984): Soziale Systeme. Grundriß einer allgemeinen Theorie. Frankfurt/M.: Suhrkamp.
- Markines, B., Cattuto, C., Menczer, F., Benz, D., Hoho, A. & Stumme, G. (2009). Evaluating Similarity Measures for Emergent Semantics of Social Tagging. WWW '09 Proceedings of the 18th international conference on World wide web, 641-650. <http://www.ra.ethz.ch/CDstore/www2009/proc/docs/p641.pdf>
- Moskaliuk, J., Kimmerle, J., & Cress, U. (2009). Wiki-Supported learning and knowledge building: Effects of Incongruity between Knowledge and Information. *Journal of Computer Assisted Learning*, 25 (6), 549-561.
- Moskaliuk, J., Kimmerle, J., & Cress, U. (2012). Collaborative knowledge building with wikis: The impact of redundancy and polarity. *Computers and Education*, 58, 1049-1057.
- Moskaliuk, J., Rath, A., Devaurs, D., Weber, N., Lindstaedt, S., Kimmerle, J., (2011). Automatic detection of accommodation steps as an indicator of knowledge maturing. *Interacting with Computers*, 23, 247-255.
- Oeberst, A., Halatchlyski, J. & Cress, U. (2012). Knowledge: What is it? How is it acquired? Who creates it? Who possesses it? Answers from individual, social, and systemic perspectives: A review and a case study. Resubmitted for publication.
- Pirolli, P. (2007). *Information foraging theory: Adaptive interaction with information*. New York: Oxford University Press.
- Pirolli, P., & Card, S.K. (1999). Information foraging. *Psychological Review*, 106(4), 643-675.
- Roschelle, J. (1996). Learning by collaborating: convergent conceptual change. In T. Koschmann, CSCL: Theory and Practice of an Emergin Paradigm, pp. 209-248. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher* 27, 4, 4-13.
- Stahl G. (2005). Group cognition in computer-assisted collaborative learning. *Journal of Computer Assisted Learning*, 21, 79-90.
- Trant, J. (2009). Studying social tagging and folksonomy: A review and framework. *Journal of Digital Information*, 10 (1). Retrieved from <http://journals.tdl.org/jodi/article/view/269/278>.
- Vander Wal, T. (2005, January 18). Folksonomy Explanations. Retrieved from: <http://www.vanderwal.net/random/entrysel.php?blog=1622>.

The Benefits and Limitations of Distributing a Tangible Interface in a Classroom

Sébastien Cuendet and Pierre Dillenbourg
Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
{sebastien.cuendet, pierre.dillenbourg}@epfl.ch

Abstract: We report the design and testing of TapaCarp, a CSCL environment for carpenter apprentices. From the start, we designed TapaCarp for a classroom usage. This led to an interface distributed over several components and modalities. A first user study conducted in a classroom environment and involving 24 carpenter apprentices produced mixed conclusions about the distribution of the interface. On one hand, it proved suitable in terms of classroom integration and flexibility. On the other hand, it hurt usability, both at the group level and at the classroom level. Based on these results, we discuss the relevance of distributing a learning environment interface and introduce the concept of “over-Hutchins threshold”, a point after which the distribution of the interface becomes harmful to classroom orchestration.

1. Introduction

Classroom orchestration has been a rising topic in the CSCL community in the last few years (Dillenbourg, Järvelä, & Fischer, 2009). A core idea of classroom orchestration is that bringing learning technologies to the classroom should accommodate the many constraints of classroom management, including practical constraints (time, discipline, etc.). These constraints have often been somehow neglected in CSCL either because studies were conducted in labs or because technologies were designed to be used ‘anywhere’.

Our recent work has focused on designing technology for face-to-face learning in the classroom. To design for the classroom, we consider both the *intrinsic* pedagogical constraints that make for an effective learning tool (how people learn, who is the audience, what are the content features) and the *extrinsic* constraints that come with the deployment of the learning environment in the classroom (time, discipline, teacher’s energy, space, etc.). In this article, we describe how we created TapaCarp, an augmented reality environment for the specific context of a classroom of carpenter apprentices attending a vocational school.

With other tangible interfaces we tested in vocational schools and elementary schools, we found paper to be a specific type of interface that fits particularly well within classroom practices. Two decades ago, Hutchins (1995) provided an explanation of why paper was surprisingly useful in an aircraft cockpit, which is a supposedly high-tech environment. According to him, the two pilots and the various artifacts spread over the physical cockpit environment form a distributed system in which information flows across different media. A classroom is a more diverse environment than a cockpit, but Hutchins’ analysis is nonetheless relevant for the classroom. We therefore used it while designing TapaCarp and to interpret the data collected during the deployment of Tapacarp in a classroom.

TapaCarp is the result of an iterative design process conducted with a teacher and his students over two years. Our continuous attempt to integrate the environment into the classroom life led to a highly distributed interface. We could even call it a ‘scattered’ interface: the learning activities imply interactions across five different modalities: tangible wooden blocks, paper cards, paper sheets, digital augmentation and even sometimes a computer mouse. Conceptually, with such a scattered interface, the environment does more or less disappear as one environment; it somehow ‘molds’ into the classroom ecosystem. Practically however, we found that distributing the interface too much leads to usability issues, both for the teacher and for the learners.

We first describe the context of carpenters training and the interface that resulted from participatory design. Second, we present the results of a study conducted with three classes. Finally, we discuss the concept of distributed interface and the tensions that can appear between the individual usability and the ‘orchestrability’ of the classroom, which Dillenbourg defined as the first and third circles of usability (Dillenbourg et al, 2011).

2. The design of TapaCarp

Context

This work focuses on carpenter apprentices, (mostly male) students aged between 16 and 20. They follow a dual system, which means that they work four days a week in a company and go to school for the remaining day. Carpenter apprentices were chosen for 3 reasons: (1) CSCL work on young adults other than university students is rare; (2) carpentry involves hard 3D reasoning skills that are challenging for apprentices; and (3) we wanted to see if the tangible-paper distributed approach we successfully tested with logistics apprentices (Zufferey, Jermann, Lucchi, & Dillenbourg, 2009) would be relevant in a different professional domain. After visiting five companies and following the apprentices at their workplace as well as at school, we identified three main topics

in which carpenters need to be trained: the law of statics for building, spatial reasoning skills, and building physics (sound/heat/humidity/insulation). This study focuses on spatial reasoning skills, which carpenters typically develop through drawing classes. Learning the practice of drawing is controversial in the carpenter community. Although the school curriculum for carpenter apprentices allots 3 hours of drawing classes per week for 3 years, at their workplace it is their superiors that draw the roof structures and not the apprentices themselves. In fact, most of the time the plans are not drawn by hand anyways but instead with CAD software. Our research addresses this tension between the needs of the companies and the practices at school: Could apprentices develop the spatial reasoning skills they need at work in a more efficient way?

Blocks and Cards

Our learning environment, TapaCarp, runs on the Tinkerlamp, a camera-projector tabletop system (see Figure 1, left). The projection area has a dimension of 70 by 50 centimeters. The system detects objects equipped with fiducial markers and provides visual feedback through the projector. This environment is designed for teams of 2 or 3 students seated around the table on which the lamp sits.

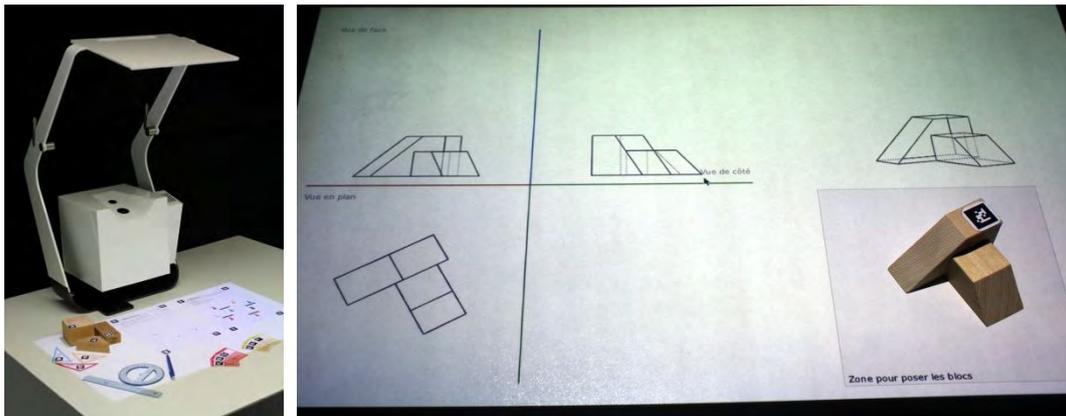


Figure 1. The Tinkerlamp (left) and the layout of TapaCarp (right): the wooden block and a perspective representation of it, as well as the three orthographic projections of the block.

TapaCarp is a collaborative learning environment that has been iteratively designed with the assistance of a carpentry teacher and his students. The teacher's main complaint about his students was that they did not make the link between the 2D representations of an object (its orthographic projections) and its 3D shape. This leads them to draw plans that are wrong, i.e. correspond to unbuildable roof structures. According to the teacher, students tend to follow descriptive geometry "recipes" to draw their plans, without understanding the link between the orthographic projections that they are drawing and the final object that they mean to represent.

The interface of TapaCarp is distributed over several components. The first component is a set of 3D wooden *blocks*. The blocks are equipped with fiducial markers (Figure 2, left), which allows TapaCarp to track their position and orientation accurately. Knowing the topology of a block, TapaCarp displays its orthographic projections and a perspective view of the object (Figure 1, right). The 2D and 3D representations are dynamically linked, allowing the users to explore the 2D-3D relationship by moving the blocks and seeing the effect of the movements on each view. The blocks serve both as an external representation of the drawing and as input for the system. Using tangible blocks that have the same geometrical shape as the digital object that they represent has been shown to be beneficial for learning (Cuendet, Bumbacher, & Dillenbourg, 2012).

While blocks are the core manipulation handles of the interface, they do not allow the users to trigger specific actions such as "launch activity 1", or "show feedback". We therefore introduced the second component of the interface: a set of papers *cards*. Cards are used to issue actions or to change options such as to launch an activity, to check the correctness of a solution, to ask for help, and to change features of the display. Each card has only one function and the number of functions provided by the system is therefore proportional to the number of cards available. This makes it easy to adapt the number of features to the students' level of expertise: the teacher simply gives them the appropriate set of cards. Cards were also chosen for practical reasons, such as their ease of distribution, storage, and sharing, all of which go in the direction of reducing the global orchestration load faced by the teacher. They are easy to manipulate and share between several users. This opens up possibilities for role taking (Burton, Brna, & Treasure-Jones, 1997).

A third component of TapaCarp is a standard computer mouse that was used to interact with the digital models. We do not have as a principle that modern interfaces should avoid using traditional computer input devices. Instead, we are looking for the most appropriate artifact for each type of interaction. In previous usability tests, which compared the mouse to tangible 'selectors', the mouse proved to be the fastest and most accurate tool to select a thin line on the views projected on the table.

Activity booklet and drawing tools

Based on the blocks and cards interface, we developed a series of learning activities to help apprentices learn to link the 2D and 3D representations. For instance, one activity was the following: given an edge shown on the 3D representation, the students had to identify it on the three orthographic projections. Another activity asked the student to place and orient one or several blocks based on two of the three orthographic projections. The activities could be completed in a short amount of time (less than 5 minutes). Their level of difficulty was adaptable by selecting simple versus complex blocks.

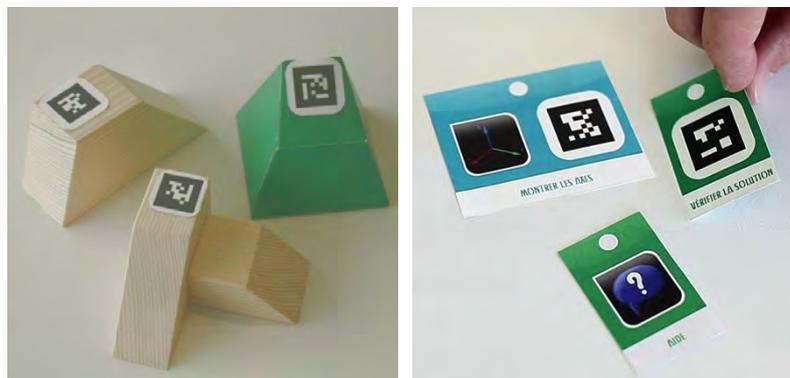


Figure 2. The blocks (left) and the cards (right) used to command the system.

Although the teacher had participated in their design and testing, he did not want to use them “for real”, i.e. as genuine classroom activities, because the activities were not part of the regular curriculum. This came as a surprise, since TapaCarp activities and the ones done routinely in the classroom had the same goal (improving the 2D-3D link), an observation that the teacher did not refute. The major issue was that the professional and school environments of a carpenter are deeply embedded with paper and drawing, and that TapaCarp used neither paper nor drawing tools. Drawing-based practices are the DNA of these classrooms.

We therefore added a new interface component a *paper activity booklet* (Figure 3, left), so that students could perform the act of drawing on paper. They drew with their *regular drawing tools* (Figure 3, right), which further satisfied the teacher – learning to use those tools properly is a curriculum requirement. The booklet was composed of A4 pages, each page being a separate activity and equipped with a fiducial marker so that the system could augment it with instructions and feedback. New activities were designed to make use of the paper and drawing tools. In the end, except for the presence of the block and the possibility to augment the paper with the projector, the activities were very similar to the ones done in the regular curriculum.

This booklet and the set of tools constitute respectively the fourth and fifth component of our distributed interface. In fact, the tools are not properly speaking part of the interface – they are not tracked by the camera and are hence not an input device – but the same fiducial markers could be used for instance, to check if the center of a protractor is accurately placed in the center of an angle to be measured.



Figure 3. The activity booklet (left) and the regular students' drawing tools (right).

While it met the teacher's requirements, introducing the activity booklet had a side effect: the use of TapaCarp became closer to the usual classroom pedagogical structure, and more scripted. Each activity was now designed as a step-by-step process through which the students were guided by dynamical instructions projected by the system, as can be seen in Figure 3 (left). It is this final system that was tested in a classroom environment. We were closer to the teacher's needs but further away from socio-constructivists principles.

3. User study

Evaluations of TapaCarp were conducted in a classroom in one vocational school over three days. The class was split into two halves: half of the students attending the normal class with the teacher, while the other half used TapaCarp under the supervision of one researcher acting as the teacher. Ideally we would have liked the teacher to give the class with Tapacarp, but practical constraints made it impossible (someone had to take care of the other half-class). The goal of the 1.5 hours lesson was to teach the apprentices how to find the true size of an object from its orthographic projections. This is one of the key tasks in the 2D-3D passage. There are various techniques to find the true size of an object from its orthographic projections, but carpenters mainly use the rabattement technique, introduced by Monge (1798) and graphically explained in Figure 4. This technique and descriptive geometry are important subjects not only for carpenters, but also for many other professions such as mechanical engineers, architectural draftsmen, and even dentists (Sheryl Sorby, 2009).

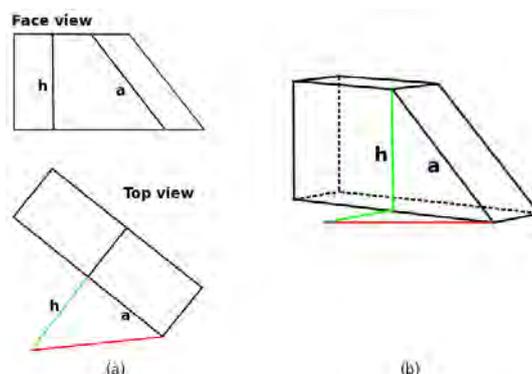


Figure 4. (a) The rabattement technique on 2D projections: to find the true size of the edge a by rabattement, one typically takes its height h on the face view and report it perpendicularly on the top view; the true size is then the red line. (b) A 3D representation of a rabattement.

3.1. Participants and procedure

Apprentices used TapaCarp in pairs. The activities had been designed together with the teacher. There were 11 activities presented in an increasing level of difficulty and grouped in three parts. All of them included exercises around the notion of rabattement: the first part (activities 1-3) was an introduction to the principle of rabattement; the second part (activities 4-6) dealt with finding the true size of an edge; the third one exercised finding the true size of a face (activities 7-11). The apprentices were all males aged between 17 and 31 with a mean age of 19. A total of 24 apprentices (12 pairs) used TapaCarp for 1.5 hours.

The blocks, the cards, the paper tools, and the activity booklet were provided to the apprentices. They were also asked to use their own regular tools: pencils, a ruler, a protractor, an eraser, and a compass. None of the activities required the mouse, which was hence not included for the study. In total, six blocks were given to each group. Each activity made use of one block. All the material was given to the apprentices at the very beginning of the class, except for three cards managing the animated feedback: These were given to them after they completed activities 3 and 4.

The data used for the analysis of the results were collected through the log files of the application, a questionnaire given to the students at the end of the experiment, and video recordings.

3.2. Results

One frequent concern with learning technologies is how fast students learn to use them. All apprentices were shown a short demonstration of TapaCarp (less than 5 minutes). To reduce the novelty effects that could distract them from the activities, the apprentices were allowed to play with the system for as long as they wanted before starting the activities. They typically tried the system for 2-3 minutes before starting the activities, and completed the activities for the remaining time within the 1.5 hour allotted.

Users feedback

We gathered both formal (written questionnaire) and informal (class discussion) feedback from the users. The questionnaire included 13 assertions to be assessed on a seven-point Likert scale, and 5 more open questions. A large majority of apprentices (18) had positive opinions, 2 were neutral and 4 were negative. From the 4 negative ones, most of the criticisms came from the lack of accuracy of the projection due to some calibration inaccuracies between the camera and the projector. The students were enthusiastic about the system, both in terms of perceived usefulness of TapaCarp for their training as well as in terms of its usability. For instance, they were interested in using TapaCarp more often (0.96 on the Likert scale) and said TapaCarp helped them

understand the rabattement better (1.25). Only three apprentices said TapaCarp did not improve their understanding of the rabattement, out of which two said that they had already understood it beforehand. The animations were deemed especially useful to better understand the rabattements (1.95).

Left-right differences

The manipulation of objects by the two users reflects the asymmetry and the modularity of the interface. Figure 5 shows the number of activities in which the user on the left or right of the workspace performed an action or manipulated an object. There was a significant difference in the usage of the modalities: the participant to the right manipulated the blocks more, while the left participant used the cards more. Part of this asymmetry can be explained by the physical placement of the manipulation zone of the blocks (on the right). However, the cards could be used anywhere, so the fact that they were used mostly by the left participant is more surprising. These differences in the usage of the interface did not lead to different learning outcomes ($F[1,22]=3.54$, $p = .07$), although in 9 groups out of 12, the post-test score of the apprentice sitting on the left was higher than his colleague's.

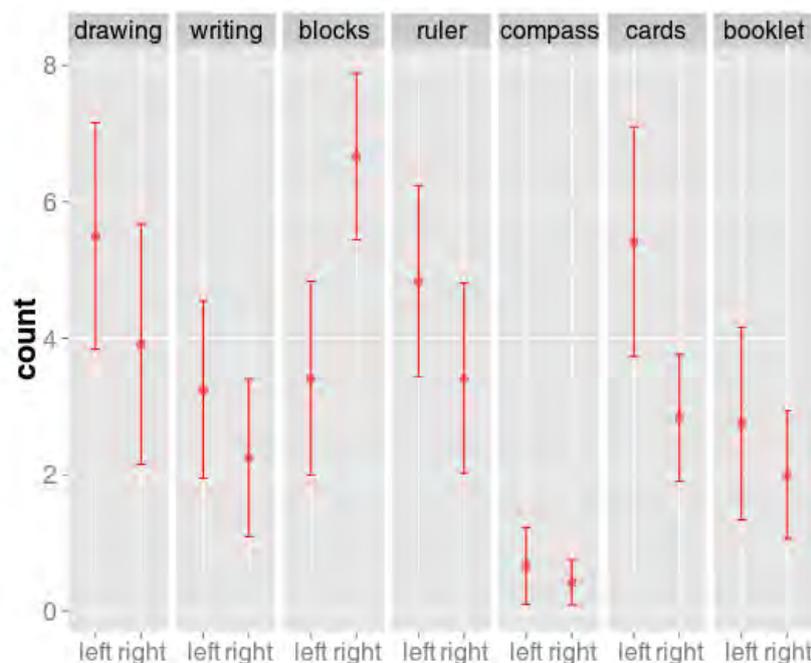


Figure 5. Average number of activities (with standard errors) in which each student on the left or right performed an action or manipulated the objects at least once.

Blocks and their interaction with the paper and tools

None of the groups tried to use the wrong block for an activity, most likely because a perspective view of the block was printed at the top of each activity page. The blocks serve both as manipulation handles and as an external representation. The apprentices' behavior showed that they understood this, and that they made the link between the 3D block, their drawing, and the projected representations. They used the blocks extensively to take measurements, check their solutions, or change the angle of the displayed projection. They measured dimensions both on the block and on the orthographic projections, and laid the block on their drawing to check that the length of an edge that they found by rabattement actually matched the real length. One could say that some of them even understood all too well how to use the blocks, since they sometimes used it not only to check their solutions but to find the solution by measuring the true size of an edge or face directly on the block instead of finding it by rabattement.

Usually, carpenter apprentices do not dispose of the physical model of what they are asked to draw on paper. Our activities forced them to link the actual block to the drawing. This link between the block and the drawing was done either directly by laying the block on the drawing, or indirectly by taking measurements on the block and reporting them on the drawing. Noteworthy is the fact that some groups did not make a direct link between the blocks and the paper and always used an additional tool – ruler or compass – to make this link. Others, on the contrary, used the block directly by laying it on the drawing.

Activity booklet

The activity booklet had a good ‘orchestrability’: easy to distribute and gather, no ordering problems or loss of activities. The apprentices are used to receiving exercises this way and did not question it. The navigation between the activities was not programmatically enforced, allowing students to browse through all the activities without any mandatory checkpoints. This resulted in some students not calling the teacher when the written instructions asked them to (so that the teacher could check their solution and give them more cards). A minor issue was that after completing several activities and flipping the pages, the stapled corner of the page was higher than the other ones, leading to some misdetection of the fiducial marker. This could easily be fixed by placing additional fiducial markers on the page.

Other observations

The augmentation of the paper with dynamic instructions and feedback generated a split attention effect. Some basic instructions had been printed on the paper to help students complete the sheet in a structured way. The dynamic instructions were projected on the top of the page. However, despite the flashing of a bright color the projected instructions were sometimes ignored. When they were stuck and asked for help, the teacher simply pointing to the instructions often solved the problem.

Each group received a total of ten cards and six blocks. These came in addition to the drawing tools and the activity booklet, and resulted in a large number of objects to manage on the tabletop. It was sometimes complicated for the students to find the card they were looking for; in some instances, a card was activated by mistake. Although students did not complain about that in their feedback, this appears to be a usability issue. From the orchestration point of view, the instructor had difficulties distinguishing what activity the students were working on from a distance, because the table was so cluttered with objects.

The experiment was not focused on learning outcomes. Actually, the cognitive activities we designed were not optimal and the learning gain in the context of this study was rather modest (4%). Experiments focusing on the evaluation of the learning gain from using TapaCarp have been reported elsewhere (Cuendet, Bumbacher, & Dillenbourg, 2012; Cuendet, Jermann, & Dillenbourg, 2012). The focus of this article is to investigate the usability of a distributed interface in a classroom environment.

4. Discussion

TapaCarp was designed with the constraints of the classroom in mind. The interface needed to be easy to use and robust to potential mishandling by new users. This led to a first version of the system with two modalities: the blocks as the main items of the interface, and cards to control the flow of the activities. Then the system evolved towards a more complex interface including three more modalities: a paper booklet, the mouse and drawing tools.

In terms of usability, the results of the field study were globally positive. With just a few minutes of introduction to TapaCarp, the students were able to use it to complete complex activities. Their feedback, although it must be taken with a grain of salt in light of the novelty effect, was positive both on the perceived usefulness and on the global usability of TapaCarp. The distributed interface allowed it to mold into the classroom ecosystem by using some of the media traditionally used in the classroom (paper and drawing tools). TapaCarp also allowed the students to work as usual on a tabletop and to keep their habit of working in pairs.

Another positive aspect of the distributed interface was that it fostered the emergence of roles within the pairs. Typically, while one student manipulated the cards, the other manipulated the blocks. On several occasions, one student prompted his partner with phrases such as “OK, we’re good, now you can use the card to go to the next step”. Role distribution in collaborative learning can be beneficial to learning (Burton et al., 1997) and has the advantage of engaging both students in the pair. In the case of tangible interfaces, where the “manipulation temptation” has been shown to be counter-productive to learning (Do-Lenh, 2012), paying attention and controlling who is manipulating what and when is especially important.

The positive aspects of a distributed interface in the classroom should not prevent us from seeing its limitations. We describe the limitations that we have observed in the next section.

Limitations

From its original orderly distributed configuration, the TapaCarp interface became over-distributed with its components being spread out all over the place without much order. This can be observed in Figure 6: on the left side of the lamp there is a ruler and a setsquare. On the right side, one can distinguish another setsquare, the six blocks, and an eraser. Close to the students’ arms are the cards, the compass, as well as some pencils and pens. With all the objects added across the four modalities, the final interface represented a total of twenty objects. We analyze the impact of the distributed interface on the usability by using the three levels of usability linked to the classroom orchestration theory (Dillenbourg et al., 2011).

- Individual level: as mentioned earlier, there were no major usability issues at the individual level.
- Group level: During the study, we observed that the large number of objects led to some usability issues at the *group level*, although the students did not explicitly complain about it. The issues arise mainly from the cards, which students either unintentionally activated or, in some case, had trouble locating. For example, the “next” card was unintentionally pushed under the projection surface – and therefore activated – which led to a rapid “completion” of the activity that was not planned by the students. On several occasions, the students wanted to ask the system for help and had to look for the “help” card because it was hidden under the activity booklet or another card, or behind a block.
- Class level: The number of objects also reduced the teacher’s *awareness* of learners work by making it difficult for him to see what group was working on what activities from a distance. This increases the teacher’s orchestration load (Dillenbourg, 2013), namely the effort to assess the progress of each group.



Figure 6. A group of students working with TapaCarp. One can see the various modalities: the drawing tools in the foreground, some cards close to the students' hands, and the blocks in the background. One block is active, and the students are reading off of the activity booklet.

While it had some positive impact, the distribution of the interface also led to some usability issues. The impact of the distributed interface on classroom orchestration can be better understood in the light of the four following design principles of classroom orchestration (Cuendet, Bonnard, Do-Lenh, & Dillenbourg, 2012): integration, flexibility, awareness, and minimalism. The distribution of the interface stemmed from a need to *integrate* into the classroom environment – the activity booklet and the drawing tools were included on the teacher’s request to fit in the classroom. The resulting *flexibility* is also increased: it is easy to add/remove part of the interface depending on the level of students, or to distribute the interface over three students rather than two if a class has an odd number of students. The usability issues described for the group level can be interpreted as a lack of *minimalism* in the design of the interface. As for the usability issues at the class level, they were directly linked to the lack of *awareness* resulting from the clutter of objects on the table.

Hutchins exposed how the distribution of information could help a cockpit remember its speed (Hutchins, 1995). Similarly to Hutchins’ approach, we developed a learning environment in which the interface is distributed across several modalities and media. Hutchins’ analysis of a cockpit assessed the distribution of information as positive. In our study, we witnessed that the distribution of the interface has potential benefits, but also that over-distributing the interface can lead to a usability reduction on some of the three usability levels. The question that begs an answer is: how many objects can a distributed interface have without hurting usability? In other words, could there be a point, which we could call the “*Hutchins threshold*”, beyond which distribution of the interface could be harmful? In reality, Hutchins neither claimed that an interface should be distributed, nor how much it should be distributed. As a tribute to his work, we simply use his name to discriminate the point where the advantages of distributed may be counterbalanced by the shortcomings.

We do not have an answer to this question. This threshold is not simply a number of objects, it also depends on the characteristics of the objects: how much space they occupy on the interaction surface, how easily they can be stacked, sorted and put away, how often they are moved unintentionally (e.g. because they are too light), etc. It would take many carefully designed experiments to answer it in the case of TapaCarp. What we observed in the user study is that some distribution of the interface increased collaboration and decreased the orchestration load, but that a higher degree of distribution hurt the usability of TapaCarp and increased the

orchestration load. While it is doubtful that there exists a general theory linking the number and type of objects in a distributed interface with the collaboration and orchestration loads, we believe that the degree of distribution of an interface is worth considering when designing a learning environment for the classroom.

For example, in TapaCarp, it is mainly the cards and the blocks that brought TapaCarp beyond the Hutchins threshold and caused the decreased of usability. It may be that reducing the number of cards and blocks needed simultaneously would solve the usability issues. This can be achieved by placing only the objects needed for the current activity on the tabletop. The cards could, for example, be bundled with the corresponding activity sheet, while the blocks could be stored in a corner of the classroom where students could go to pick up the block corresponding to their current activity.

5. Conclusion

The approach presented for the design of TapaCarp is distinctive in that we sought from the start to design TapaCarp for a classroom usage, and that it led to a distributed interface. Following the observations of the user study, we discussed the benefits and drawbacks of a distributed interface for a learning environment. On one hand, distributing the interface proved suitable in terms of classroom integration and flexibility. On the other hand, it hurt the usability of TapaCarp, both at the group level and at the classroom level.

We see great potential in distributed interfaces for learning environments. However, over-distribution may hurt usability, and to this end we introduced the concept of the “Hutchins threshold”, a point at which the distribution of the interface breaks one of the five principles of classroom orchestration design. We do not claim that the Hutchins threshold can be computed and will hold true invariant of the conditions. Rather, we see it as an important design concept to keep in mind when designing tabletops and tangible interfaces for learning.

Acknowledgements

The authors would like to thank Stéphane Testuz, Yongsung Kim and Sara Alsudairy for their contribution to the development of Tapacarp. We are grateful to Patrick Jermann for his many suggestions all throughout the project. We also thank the reviewers for their insightful comments. This work is supported by the Swiss State Secretariat for Education, Research and Innovation (SERI).

References

- Burton, M., Brna, P., & Treasure-Jones, T. (1997). Splitting the collaborative atom: How to support learning about collaboration. *AI in Education*.
- Cuendet, S., Bonnard, Q., Do-Lenh, S., & Dillenbourg, P. (2012). Designing Augmented Reality for the Classroom. *Computers and Education, Special Issue on Augmented Reality for Learning*.
- Cuendet, S., Bumbacher, E. W., & Dillenbourg, P. (2012). Tangible vs. Virtual Representations: when Tangibles Benefit the Training of Spatial Skills. *The Seventh Nordic Conference on Human-Computer Interaction (NordiCHI 2012)*.
- Cuendet, S., Jermann, P., & Dillenbourg, P. (2012). Tangible interfaces: when physical-virtual coupling may be detrimental to learning. *Proceedings of the 2012 British Computer Society Conference on Human-Computer Interaction*.
- Dillenbourg, P. (2013). Design for Classroom Orchestration. *Journal of Computers in Education*.
- Dillenbourg, P., Järvelä, 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). Springer Netherlands.
- Dillenbourg, P., Zufferey, G., Alavi, H. S., Jermann, P., Do, L. H. S., Bonnard, Q., Cuendet, S., et al. (2011). Classroom orchestration : The third circle of usability. *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings. Volume I — Long Papers* (Vol. 1, pp. 510–517). International Society of the Learning Sciences.
- Do-Lenh, H. S. (2012). *Supporting Reflection and Classroom Orchestration with Tangible Tabletops*. EPFL, Lausanne.
- Hutchins, E. (1995). How a Cockpit Remembers Its Speeds. *Cognitive Science*, 19(3), 265–288.
- Monge, G. (1798). *Géométrie Descriptive*. J. Klostermann fils, Paris.
- Sheryl Sorby. (2009). Educational Research in Developing 3-D Spatial Skills for Engineering Students, 31(3), 459–480.
- Zufferey, G., Jermann, P., Lucchi, A., & Dillenbourg, P. (2009). TinkerSheets: Using Paper Forms to Control and Visualize Tangible Simulations. *TEI'09, Third International Conference on Embedded and Tangible Interaction*.

Cohesion-based Analysis of CSCL Conversations: Holistic and Individual Perspectives

Mihai Dascalu, Stefan Trausan-Matu, University “Politehnica” of Bucharest,
313 Splaiul Independentei, 060042 Bucharest, Romania
Email: mihai.dascalu@cs.pub.ro, stefan.trausan@cs.pub.ro
Philippe Dessus, LSE, UPMF Grenoble-2 & IUFM–UJF Grenoble-1, France,
philippe.dessus@upmf-grenoble.fr

Abstract: Although Computer Supported Collaborative Learning (CSCL) technologies have gained an increasing role in educational environments, there are few automatic systems that address involvement, knowledge-building and collaboration in order to support tutors in the time consuming process of analyzing conversations. We propose a cohesion-based analysis model integrating multiple natural language techniques, an intervention scoring mechanism and a comprehensive collaboration assessment method, derived from social knowledge-building, reflected at utterance level through cohesion. Furthermore, by combining a holistic perspective of the entire conversation with a more fine grained view focused on each participant, we obtain a thorough evaluation of chat conversations with focus on topics modeling, participant interaction and collaboration. In order to sustain our model, we have performed a preliminary validation study that proves that our analysis is consistent with tutor evaluations.

Introduction

Learners collaborating through forums or chats use language as a mediator for building knowledge, more specifically for developing one’s view about the material, for negotiating meaning with each other and for defining new cultural artifacts (Stahl, 2006). The activity of reading is the cornerstone of every CSCL situation: learners are typically engaged in reading to acquire new pieces of knowledge; to be aware of the way their peers get acquainted about this knowledge; to understand what is required by their teachers or tutors. They can even carefully re-read their own utterances before and after posting.

Cohesion, i.e., relations of meaning within a text, is one of the most important features driving the understanding of read texts (Tapiero, 2007). Cohesion gaps within textual materials, e.g., online conversation, may yield comprehension difficulties, mostly for low-knowledge learners (McNamara et al., 1996), but also be viewed as topic change along discussion threads (Dascalu et al., 2010a). However, the cohesion-based assessment of learning materials or learners’ utterances is a very demanding task and systems that automatically analyze utterances during learning sessions are of particular interest (Fujita & Teplovs, 2009; Trausan-Matu et al., 2007). Such systems often make an intensive use of Latent Semantic Analysis (Landauer & Dumais, 1997), which has been proven to capture cohesion-based textual relationships (Graesser et al., 2004).

In this paper we introduce *ReaderBench*, a multi-purpose system that proposes several cohesion-based views on the learners’ contributions to the discussion and presents them upon two perspectives: *individual*, centered on each participant, and *holistic*, covering the entire discussion. The first perspective allows learners to individually understand the extent to which they contribute to a given discussion, while the latter is useful for teachers and tutors to assess its fruitfulness in terms of knowledge-building and collaboration.

Moreover, our aim is to provide a comprehensive approach for *assessing collaboration* throughout a multi-participant chat conversation, as a refinement of the initial information gain assessment of collaboration (Dascalu et al., 2010b; Trausan-Matu et al., 2012). Our implemented model automatically identifies intense collaboration zones and extrapolates the concepts of personal and social knowledge-building (Bereiter, 2002; Scardamalia, 2002; Stahl, 2006) to a finer grained dimension, more specifically to each utterance. In this particular context, cohesion is the bridge between utterances, enabling information transfer and cumulative knowledge-building from both *personal* (utterances from the same speaker) and *social* perspectives (utterances from different speakers). Additionally, our model of analysis goes one step further towards defining an implemented *polyphonic model* of a discussion (Trausan-Matu et al., 2007) by also considering an analogy to key dialogic concepts (Bakhtin, 1986). In order to provide a quick equivalence that will be detailed later on, *topics* emerging from the conversation and pertaining to different speakers are actual *voices*, whereas *cohesion* induces the *echo cancelation* effect of each utterance upon later interventions within the discussion.

The following section provides an insight into our analysis model, with emphasis on building the cohesion graph and on the multi-layered scoring mechanism. Afterwards, the individual assessment section is centered on the identification of topics and modeling the participants’ interaction throughout the conversation. Later on, we shift the point of interest towards the assessment and evolution of collaboration. Eventually, a preliminary validation experiment is presented and the relevance of the results is debated.

Cohesion-based Conversation Analysis

Cohesion, seen as the relationship at lexical and semantic levels between two segments of texts, is defined within our model in terms of *recency*, *lexical proximity* and *semantic distances*. Therefore, cohesion is a symmetric function that combines: 1/ the inverse distance between the analyzed textual elements within a fixed window of analysis that defines a local context, 2/ lexical proximity as a relaxed condition of having identical stems within both text fragments and 3/ semantic distances determined from three perspectives.

The first semantic dimension of cohesion is represented by semantic distances within lexicalized ontologies. As *ReaderBench* was designed to cover both English and French, we used besides *WordNet* (Miller, 1995), a transposed and serialized version of *WOLF* (*Wordnet Libre du Français*, <http://alpage.inria.fr/~sagot/wolf.html>) in order to measure similarity between concepts. Our experiments showed that Wu-Palmer semantic distance performs best as it scales correspondingly to the [0; 1] interval and it works properly for both English and French languages. Although *WOLF* is rather incomplete and mixes word glosses, expressing them partially in English and partially in French, it is currently the best alternative for having a multi-lingual analysis platform.

Going deeper within our cohesion estimation function, the other two semantic dimensions are represented by Latent Semantic Analysis (LSA) (Landauer, & Dumais, 1997) and Latent Dirichlet Allocation (LDA) (Blei et al., 2003). Whereas LSA is used for measuring similarity between discourse fragments through cosine similarity after projecting concepts in a pre-trained vector space, LDA exploits the Jensen-Shannon dissimilarity (Schütze & Manning, 1999) of posterior topic distributions in order to grasp inconsistencies or discrepancies between texts. Therefore, from a computational point of view, we have combined Information Retrieval specific techniques, mostly reflected in word repetitions and normalized number of occurrences, with semantic models (LSA and LDA) in order to provide a comprehensible and measurable value of cohesion between textual fragments.

After defining the cohesion function between different analysis units within the discourse, the next step consists of automatically building the cohesion graph, which plays a central role within our analysis. We propose a multi-layered graph consisting of the following hierarchy: *document* (entire chat conversation or reading material) > *block* (utterance or paragraph within the initial text) > *sentence*. Within the previous structure, we enforce hierarchical links determined by inclusion functions (sentences within a block, blocks within the document) and mandatory links obtained from: 1/ *explicit links* defined by participants within the user interface – in our case, *ConcertChat* – or 2/ *adjacent blocks* of text. Later on, cohesion is determined between all sentences of each block (intra-block cohesion) and between blocks within a predefined sliding window of analysis (inter-block cohesion measurements that within our experiments considered a window of 20 utterances). The resulting links were pruned after the sum of mean and standard deviation of cohesion values in order to keep only relevant links within the discourse model.

Our aim was to provide a generalized and customizable model for assessing different types of discourses in terms of cohesion: multi-participant chat conversations, on one hand, and texts in general, on the other. More specific to CSCS conversations, we opted for Dong's perspective (1995) of separating utterances based on turn-taking events between speakers. Although most participants' utterances consist of one sentence (in some cases, elliptical sentences are quite frequent), we preferred to create a scalable model, which can be later on easily adjusted to forums, for example, in which interventions entangle multiple sentences.

An important component in the evaluation process of each participant and of collaboration throughout the conversation is our bottom-up intervention scoring method. Firstly, specific natural language techniques are applied on each utterance: tokenizing, splitting, part of speech tagging, parsing, stop words elimination, dictionary-only words selection, stemming, lemmatizing, named entity recognizer and co-reference resolution (Manning & Schütze, 1999). Afterwards, each sentence is assigned an individual score of each remaining concept equal to its term frequency – inverse document frequency (Manning & Schütze, 1999) multiplied by its cosine similarity from the LSA vector space to the corresponding sentence, denoting the local significance of each word. Furthermore, cohesion with a generic title or with a list of predefined key concepts is also measured and added to the score of each sentence. Due to the nature of our experiments conducted in an academic environment, each conversation had a specific set of concepts, defined by the tutor, which had to be covered by the participants. In other words, we measure through cohesion the degree in which each participant addresses each concept and whether the conversation stays on-topic.

Secondly, at utterance or block level, individual sentence scores are weighted by cohesion measures and summed up in order to define the intra-block score. By going further into our discourse decomposition model (document > block > sentence), inter-block cohesion values are used to augment the previous intra-block scores by also considering all the links from the previously defined cohesion graph. In the end, all block scores are combined in order to determine the overall conversation score. Figure 1 depicts the scores for each intervention in brackets, immediately after each participant's intervention. The proposed scoring mechanism is effective because it considers the local importance of each intervention and it also combines these individual scores through cohesion to other semantically related elements within the discourse structure.

Individual Participant Assessment

Our aim regarding the individual assessment of participants is to focus on covered topics and interaction modeling. *Topics*, seen as key concepts from the conversation, play a leading role in obtaining a view of the overall discussion, but also in observing each participant's points of interest. Tightly connected to the cohesion graph and the previous scoring mechanism, topics are extracted per entire conversation and per participant. The *relevance* of each concept mentioned within the discussion is determined by combining a multitude of factors: 1/ individual *term frequency – inverse document frequency* of each word, 2/ *semantic similarity through LSA* to the entire conversation, corresponding utterances and sentences 3/ *cosine similarity* between the distribution of the word in the predetermined topics from *LDA* and the conversation's overall distribution over the same classes and 4/ the *scores* of each *sentence* and *utterance* in which the word is present. Within this evaluation, the initial natural language processing, with emphasis on lemmatization, played an important role in standardizing the inputs. Moreover, the scores of each concept are normalized by the length of the lexical chains in which they occur; the previous lexical chains are determined from the disambiguation graph modeled through semantic distances from *WordNet* and *WOLF* (Galley & McKeown, 2003). In addition, as an empirical improvement, filtering the previous list of topics by corresponding parts of speech and selecting solely nouns provided more accurate results in most cases, as nouns tend to better grasp the conceptualization of the discussion.

In this context, key topics together with their corresponding lexical chains can be considered voices that spread throughout the conversation, while cohesion simulates echoes of voices to other inter-linked utterances. Moreover, by changing the focus on a specific participant, we can observe the strength of a voice as being directly proportional to the relevance of previously identified topics.

Figure 1 presents the main user interface from *ReaderBench* in which a chat conversation has been loaded from the XML file exported from *ConcertChat*, the scores of each intervention are displayed in brackets and the automatically identified topics for a specific participant are presented in the right sidebar.



Figure 1. *ReaderBench* Main User Interface.

Following the transition from a global view of discourse to a user centered perspective, the visualization component of the conceptual space for each chat participant as a mind-map is based on cosine similarities computed within our LSA vector space. Terms central to a given discussion may not appear in any utterance but, nonetheless, be worth displaying for comprehension's sake. We thus enriched the previously identified topics list with *inferred concepts*, not mentioned within the text, for which three measures are computed and summed in the end: 1/ the similarity to the entire conversation vector, 2/ the similarity to the weighted topics vector determined from the previous list of concepts displayed within the user interface and 3/ the similarity to the closest concept within the text, because tight correlation to a concept already present within the text increases the proximity to the overall discussion. These three dimensions have as empiric argumentation the method in which experts selected inferred concepts from manually assessed conversations: similarity to a concept or idea from the text, but also to the whole text (and the entire text is represented as the document vector and as an extrapolation of the already identified topics).

Figure 2a presents the conceptualization space of a user including initial topics and the automatically inferred concepts. From a computational point of view, we determine all similarities between the previous words, considered nodes within the network, and select as edges solely the ones that are above the threshold selected from the interface (30% in this particular case). The actual size of each node is directly proportional to a Social Network Analysis metric derived from the previous graph (e.g. betweenness). All multidimensional scaling and visualization of the graph as a force-based layout are performed by Gephi (Bastian et al., 2009).

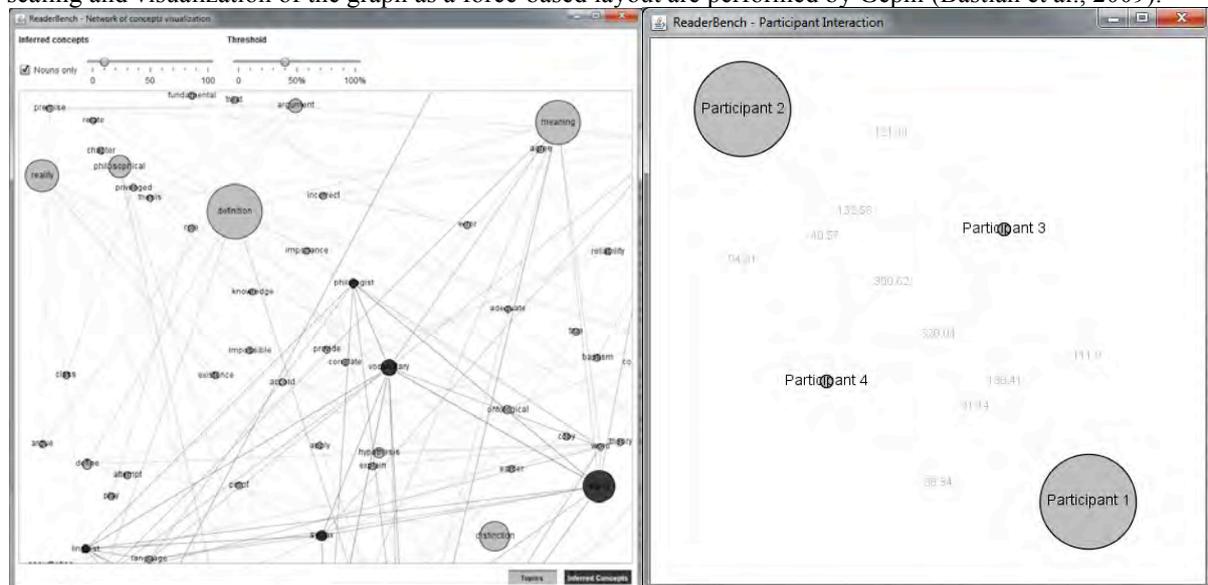


Figure 2. a. Concepts Visualization and b. Individual View of the Interaction Graph.

Besides the identification of topics for each participant in the discussion, *ReaderBench* also supports participant interaction modeling covering two major dimensions: a preliminary *quantitative* one, generated by the actual number of interactions per participant, and a deeper *qualitative* dimension, obtained after summing up intervention scores. Internally, an interaction graph is built with participants as nodes and the weight of links equal to the sum of interventions scores multiplied by the cohesion function with the referred element of analysis, extracted from the cohesion graph. Therefore, by performing social network analysis on the previous participant interaction graph, the scale of analysis is shifted towards an individual perspective (see Figure 2b).

Collaboration Assessment

In order to achieve genuine collaboration, which can be seen as true polyphony in terms of Bakhtin's dialogic theory (1984), the conversation must be generated by a dense intertwining of voices, derived from key concepts of the conversation and covering all participants (Trausan-Matu & Rebedea, 2009). Moreover, by enforcing a personal and social knowledge-building process (Bereiter, 2002; Scardamalia, 2002; Stahl, 2006) at utterance level and by using cohesion as a bridge between utterances seen as an overlap of voices, the actual information transfer through cohesion between units of analysis obtains two valences. Firstly, a *personal dimension* emerges by considering utterances with the same speaker, therefore modeling an inner voice or a continuation of the discourse. Secondly, inter-changed utterances having different speakers define a *social perspective* that models collaboration as a cumulative effect.

Therefore, each intervention now has its previously defined *importance score* and a *knowledge-building (KB) effect*, both personal and social. The personal effect is initialized as the intervention's score, whereas the social effect is zero. Later on, by considering all the links from the cohesion graph, each dimension is correspondingly augmented: if the link is between utterances with the same speaker, the previously built knowledge (both personal and social) from the referred utterance is transferred through the cohesion function to the personal dimension of the current utterance; otherwise, if the pair of utterances is between different participants, the social knowledge-building dimension of the currently analyzed utterance is increased with the same amount of information (previous knowledge multiplied by the cohesion measure). In other words, continuation of ideas or explicitly referencing utterances of the same speaker builds an inner dialogue or personal knowledge, whereas the social perspective measures the interaction with other participants, encourages sharing of ideas, fostering creativity for working in groups (Trausan-Matu, 2010b) and influencing the other participants' points of view during the discussion, thus enabling a truly collaborative discussion.

In this manner we can actually measure collaboration through the sum of social knowledge-building effects, starting from each intervention's score corroborated with the cohesion function. Moreover, personal knowledge-building addresses individual voices (participant voices or implicit/alien voices covering the same

speaker), while social knowledge-building, derived from explicit dialog (that by definition is between two entities), sustains collaboration and highlights external voices. By referring to the dialogic model of discourse analysis, besides voices that are derived from the topics of the conversation in correlation to each participant's point of view, echoes are reflected by cohesion in terms of the information transferred between utterances, whereas the attenuation effect diminishes the strength of the cohesion link with the increase in distance between the analysis elements.

Nevertheless, we must also consider the limitations of our implemented model in terms of personal knowledge-building. Collaboration clearly emerges from social knowledge transfer through cohesion as the influence of one's intervention over other participants' discourse. In contrast, the approximation of personal knowledge-building rather represents an upper bound of the explicitly expressed information transfer between one's personal interventions. Similarly to the gain-based approach (Dascalu et al., 2010b; Trausan-Matu et al., 2012), we use a quantifiable approximation of inner dialogue, although limited in terms of underlying cognitive processes. Personal knowledge-building is seen as a reflection of one's thoughts expressed explicitly within the ongoing conversation as cohesive links between interventions of the same chat participant. But this reflection does not necessarily induce personal knowledge-building, only a cohesive discourse. Therefore, we can consider that the computed value of personal knowledge-building is a *maximum value* of the explicit personal knowledge-building effect, modeled during the discourse through cohesive links.

In addition to the estimation of personal and social knowledge-building effects for each utterance and the modeling of their corresponding evolution throughout the conversation (Figure 3), *ReaderBench* automatically identifies *intense collaboration zones* that are intervals of utterances in which participants are actively involved, collaborate and generate new ideas related to the ongoing context of the discussion. The first step within our greedy algorithm for detecting intense collaboration zones consists of identifying social knowledge-building peaks as maximum local values. Afterwards, each peak is expanded sideways within a predefined slack (experimentally set at 2.5% of the number of utterances). This slack was important due to our focus on the macro-level analysis of collaboration and due to the possible intertwining of multiple discussion threads. In the end, only zones above a minimum spread of 5 utterances are selected as intense collaboration zones. In other words, after identifying the utterances with the greatest social knowledge-building effect, the algorithm expands each zone to the left and to the right, in a non-overlapping manner to previously identified zones, by considering solely utterances above the mean social knowledge-building value and that are in the previously defined slack. If in the end, the zone covers more than the specified minimum spread, it is considered an intense collaboration zone. From a different point of view and highly related to social knowledge-building, cohesion binds utterances within an intense collaboration zone in terms of on-topic relatedness.

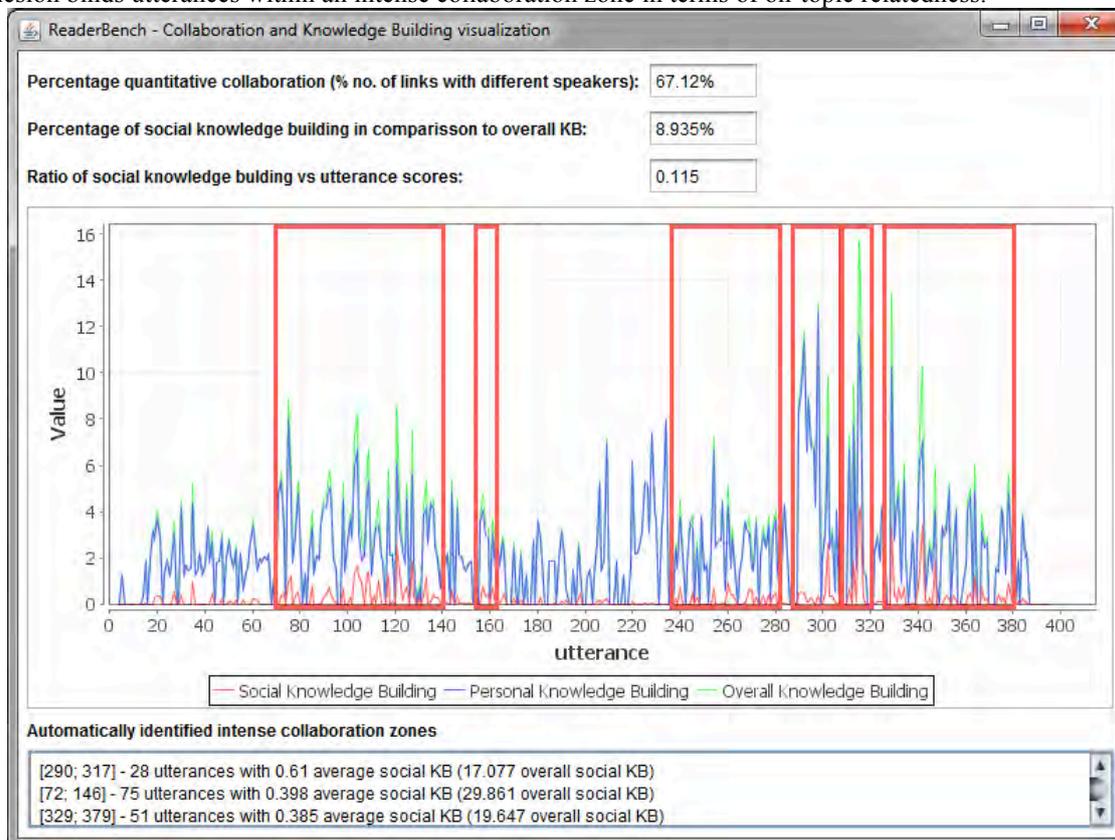


Figure 3. Collaboration Assessment and its Evolution in Time.

From a holistic perspective, three factors were implemented in order to best characterize the overall collaboration within the discussion. Firstly, *quantitative collaboration* is determined as the percentage of links from the cohesion graph having different speakers in comparison to the number of links automatically identified. Although rough as estimation, this measurement provides good insight with regards to the actual information exchange between participants. Secondly, the *overall social knowledge-building score* is compared to the overall knowledge-building effect. Thirdly, the *ratio* between the *overall social knowledge-building score* and the *overall intervention importance scores* is computed for highlighting the amount of information that is transferred through collaboration in comparison to what was withheld initially within each utterance. Low values for the latter factors are normal due to the fact that cohesion, after normalization in the [0, 1] interval, usually integrates low similarity values between utterances of a chat conversation in terms of LSA and LDA.

First Validation Study

Preliminary experiments were conducted in order to validate the dialogic model used for evaluating chat conversations, with emphasis on participant involvement and collaboration assessment. Three chat conversations conducted in an academic environment, with students from the 4th year undergoing the Human-Computer Interaction course and debating on CSCL technologies, were manually assessed by 4 tutors. More specifically, each student had to focus on a CSCL technology (chat, wiki, blog or forum), to present and debate on its benefits in specific use case scenarios generated throughout the conversation. These three conversations (Team 4, Team 34 and Team 36) were selected for detailed analysis after an overview of approximately 50 discussions engaging more than 200 students. Although high discrepancies were noticed in terms of the quality of the content, the involvement and the collaboration of its participants, these conversations were considered representative for the entire sample and the preliminary evaluations were conducted only on these conversations due to the high amount of time it takes to manually assess a single chat conversation (2 to 4 hours for a deep understanding of involvement and of collaboration).

Additionally, the *time evolution modeling* interface depicted in Figure 4 was developed in order to facilitate the manual evaluation of chats in terms of intense collaboration zones. In this context, the presentation of the conversation follows the timeline and models the inter-twining of utterances, based on the cohesion graph. This component is useful for manually identifying: 1/ *breaks within the conversation, zones with limited or no collaboration*, due to the fact that within a specific time-frame we have a monologue of a participant, without any interventions from other users, and 2/ *zones with high collaboration* due to the dense inter-animation of utterances between different participants. In the particular case presented in Figure 4, all utterances with identifiers between 30 and 50 belong a single user, within a limited time-frame, therefore making the social knowledge-building effect zero. Afterwards, as multiple participants get involved in the ongoing discussion, collaboration increases.

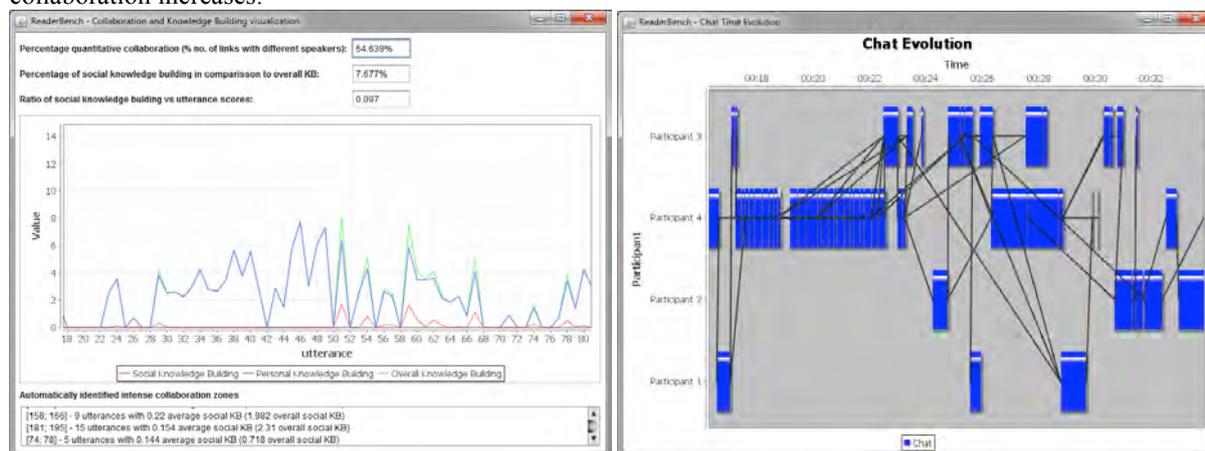


Figure 4. Time slice of a conversation highlighting cohesion links and a monologue within the conversation.

Table 1 presents the correlation between different evaluation factors extracted from *ReaderBench* and the final grades assigned by the experts. Although the participant's identifiers coincide, each conversation had different students attending it. Moreover, in order to ensure the equitability of our analysis, the correlations between the factors automatically determined by the system and the average values of the grades manually assigned by the experts were computed after combining the participants' scores from all conversations.

As an interpretation of the results presented in Table 1, we can observe in Team 4 conversation a discrepancy, as the involvement of the participants from a personal point of view was good, while the actual collaboration throughout the conversation is highly unbalanced. Team 34 conversation has the lowest scores in

all the factors, whereas Team 36 conversation, that was considered the best by the tutors in terms of both quality and involvement, has the highest scores assigned by the system.

By analyzing each factor's correlation, it becomes quite clear that the tutors emphasized on the quality of the interventions, not on the mere number of utterances. As the Intraclass Correlation Coefficient (ICC) was .61 on single measures and .86 on average, results in terms of intervention importance scores and social knowledge-building correlate extremely well with the average expert grades. Additionally, the personal knowledge-building dimension is better correlated with the expert's grades; this sustains the empiric observation that the personal effect reflected in each participant's involvement has a higher impact in the expert's evaluation than the social dimension, through which collaboration is attained. In other words, it is easier to grasp one's active participation than his social interaction from the perspective of collaboratively building knowledge.

Table 1: Correlation between manual and automatic participants' evaluations.

Participant Name	Number of Utterances	Overall Interventions Score	Overall Personal KB	Overall Social KB	Expert Grade					
					1	2	3	4	Avg.	
Team 4										
Participant 1	90	215.34	262.19	27.63	9.0	10.0	8.0	9.5	9.13	
Participant 2	61	99.78	114.91	15.39	7.0	7.0	8.0	6.5	7.13	
Participant 3	120	220.92	248.99	12.76	8.0	9.0	8.5	10.0	8.88	
Participant 4	118	205.79	239.96	29.19	8.0	6.5	9.0	8.0	7.88	
Team 34										
Participant 1	23	60.13	77.91	11.42	6.0	6.0	7.0	6.0	6.25	
Participant 2	34	64.17	72.58	6.66	5.0	7.0	7.0	6.0	6.25	
Participant 3	73	120.03	139.51	13.22	8.0	7.0	8.0	7.0	7.50	
Participant 4	60	124.89	141.12	4.55	7.0	4.0	7.0	5.5	5.88	
Team 36										
Participant 1	54	160.89	200.18	37.95	9.0		9.5	8.0	8.83	
Participant 2	67	201.57	251.98	49.46	10.0		8.0	10.0	9.33	
Participant 3	119	327.45	412.78	58.83	9.0		8.0	8.5	8.50	
Participant 4	57	161.39	200.60	40.33	9.0		9.5	8.0	8.83	
Average correlation	.64	.83	.84	.77						

In terms of intense collaboration zones, manual annotations and automatically identified zones are presented in Table 2.

Table 2: Overlap between manual and automatic identification of intense collaboration zones.

Conversation	Number of utterances	Manually annotated collaboration zones	Automatically identified collaboration zones
Team 4	389	[90; 160], [320; 360]	[19; 43], [72; 146], [157; 166], [240; 280], [290; 317], [329; 379]
Team 34	190	[90; 120], [170; 178]	[51; 67], [74; 78], [94; 121], [127; 131], [158; 166], [181; 195]
Team 36	297	Relatively uniform distribution	[38; 58], [69; 122], [136; 157], [202; 221], [286; 290]

By analyzing all previous results, besides the global evaluation factors that correlate well with the human evaluations, the following *indicators of bad collaboration* (mostly in Team 34 conversation) were observed: the high number of automatically identified zones containing 1 to 3 utterances which were not considered intense collaboration zones in the end, the low average value of social knowledge-building effects and no automatically identified collaboration zone with a wide spread (over 50 utterances, although it was the shortest conversation of the three). In contrast, *conversations with good collaboration* (namely Team 36 which had the best overall collaboration) have more balanced collaboration zones, higher average values of social knowledge-building and a better distribution and coverage over the entire conversation.

Further investigation and analysis of results is required on an extended sample of discussions, but by taking into consideration the subjectivity of the tasks at hand, preliminary results are quite promising and prove that our implemented dialogic model provides measures relevant in terms of the evaluation.

Conclusions

ReaderBench has been designed to support tutors and learners in various educational scenarios that involve the use of chats or forum discussions for solving different tasks, with emphasis on collaborative problem solving or debating. Based on the results of our first validation study, we strongly believe that *ReaderBench* can be used to support tutors in the time consuming process of manually assessing chat conversations.

Overall, starting from a dialogic model of discourse, our integrated analysis system can be used to identify cohesion gaps between utterances, to analyze participants' involvement and to evaluate collaboration individually and holistically, through the process of social knowledge-building. The collaboration assessment model can be considered a cornerstone as it induces dialogistic approach and emphasizes the social knowledge-building perspective of collaboration in CSSL environments. In extent, by combining two different perceptions of cohesion, CSSL participants can use this system to assess to what extent utterance cohesion reflects group cohesion, as an outcome of collaboration depicted from the interaction graph.

References

- Bakhtin, M. (1984). *Problems of Dostoevsky's poetics*. Minneapolis, MN: University of Minnesota Press.
- Bakhtin, M. (1986). *Speech genres and other late essays*. Austin: University of Texas.
- Bastian M., Heymann S., & Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. *Proc. Int. AAAI Conference on Weblogs and Social Media*.
- Bereiter, C. (2002). *Education and Mind in the Knowledge Age*. Mahwah: Lawrence Erlbaum Associates.
- Blei, D., Ng, A., & Jordan, M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3(4-5), 993–1022.
- Dascalu, M., Trausan-Matu, S., & Dessus, P. (2010a). *Utterances assessment in chat conversations*. Research in Computing Science, 46, 323-334.
- Dascalu, M., Rebedea, T., & Trausan-Matu, S. (2010b). *A Deep Insight in Chat Analysis: Collaboration, Evolution and Evaluation, Summarization and Search*. In D. Dicheva and D. Dochev (Eds.), Proceedings of AIMSA 2010, LNCS 6304, Springer, 2010, 191-200.
- Dong, A. (2005). The latent semantic approach to studying design team communication, *Design Studies*, 26(5), 445-461.
- Fujita, N., & Teplovs, C. (2009). Automating the analysis of collaborative discourse: Identifying idea clusters. In C. O'Malley, D. Suthers, P. Reimann & A. Dimitracopoulou (Eds.), *Proceedings of the Int. Conf. CSSL 2009* (pp. 162-164). Rhodes: ISLS.
- Galley, M., & McKeown, K. (2003). Improving Word Sense Disambiguation in Lexical Chaining. *Proc. 18th International Joint Conference on Artificial Intelligence (IJCAI'03)*, Acapulco, Mexico.
- Graesser, A. C., McNamara, D. S., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers*, 36, 193–202.
- Landauer, T.K., & Dumais, S.T. (1997). A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychol. Rev.*, 104(2), 211–240.
- McNamara, D., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14(1), 1–43.
- Miller, G. (1995). WordNet: A Lexical Database for English. *Communications of the ACM*, 38(11), 39–41.
- Manning, C., Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, Mass: MIT Press.
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith & C. Bereiter (Eds.), *Liberal Education in a Knowledge Society* (pp. 67-98). Chicago: Open Court.
- Stahl, G. (2006). *Group cognition. Computer support for building collaborative knowledge*. Cambridge: MIT Press.
- Tapiero, I. (2007). *Situation models and levels of coherence*. Mahwah: Erlbaum.
- Trausan-Matu, S., Rebedea, T., Dragan, A., & Alexandru, C. (2007). *Visualisation of learners' contributions in chat conversations*. In J.Fong & F.L. Wang(Eds.), *Blended learning* (pp. 217-226). Singapour: Pearson.
- Trausan-Matu, S., & Rebedea, T.(2009) *Polyphonic Inter-Animation of Voices in VMT*. In Stahl, G. (Ed.), *Studying Virtual Math Teams* (pp. 451-473), Boston, MA: Springer US.
- Trausan-Matu, S. (2010). Computer Support for Creativity in Small Groups using chats. *Annals of the Academy of Romanian Scientists, Series on Science and Technology of Information*, 3(2), 81-90.
- Trausan-Matu, S., Dascalu, M., & Rebedea, T. (2012). *A System for the Automatic Analysis of Computer-Supported Collaborative Learning Chats*. In proceedings of ICALT 2012, IEEE 2012, 95-99.

Acknowledgments

The research presented in this paper was partially supported by the project No. 264207, ERRIC-Empowering Romanian Research on Intelligent Information Technologies/FP7-REGPOT-2010-1.

Going Deep: Supporting Collaborative Exploration of Evolution in Natural History Museums

Pryce Davis, Michael Horn, Laurel Schrementi, Northwestern University, Evanston, IL 60208

Florian Block, Brenda Phillips, Harvard University, Cambridge, MA 02138

E. Margaret Evans, University of Michigan, Ann Arbor, MI 48109

Judy Diamond, University of Nebraska State Museum, Lincoln, NE 68588

Chia Shen, Harvard University, Cambridge, MA 02138

Email: pryce@u.northwestern.edu, michael-horn@northwestern.edu, laurels@u.northwestern.edu

Abstract: We provide an analysis of pairs of children interacting with a multi-touch tabletop exhibit designed to help museum visitors learn about evolution and the tree of life. The exhibit's aim is to inspire visitors with a sense of wonder at life's diversity while providing insight into key evolutionary concepts such as common descent. We find that children negotiate their interaction with the exhibit in a variety of ways including *reactive*, *articulated*, and *contemplated* exploration. These strategies in turn influence the ways in which children make meaning through their experiences. We consider how specific aspects of the exhibit design shape these collaborative exploration and meaning-making activities.

Introduction and Background

Evolution is a central organizing principle of modern biology that accounts for the diversity of life on Earth. Despite its importance, evolution remains poorly understood by the general public, particularly in the United States (Rosengren et al., 2012; Miller, Scott, & Okamoto, 2006). In this paper we present a qualitative analysis of an interactive tabletop exhibit called *DeepTree* that we have designed to help museum visitors explore key evolutionary concepts. The exhibit presents an interactive visualization of the "tree of life" consisting of over 70,000 species that visitors are free to explore through a *deep zoom* interaction technique (Figure 1). We emphasize the idea that life on Earth is not only astonishingly diverse but also related through common ancestry. A key design challenge is to provide visitors with the means to explore a vast information space, instilling in them a sense of wonder at life's diversity while providing insight into evolutionary landmarks.

After briefly describing the exhibit, we present a study involving pairs of 9- to 15-year-old children interacting with the exhibit at two natural history museums. Our analysis focuses on three questions related to the use of multi-touch tabletops to support collaborative learning in museums: First, how do dyads negotiate their moment-to-moment exploration of the exhibit? Second, how do dyads negotiate meaning through their interaction? And, finally, how do specific aspects of the exhibit design shape these collaborative activities? Our contribution in this paper is a framework that describes dyadic interaction along with an account of the role of design in allowing visitors to make sense of large *information visualization* exhibits.

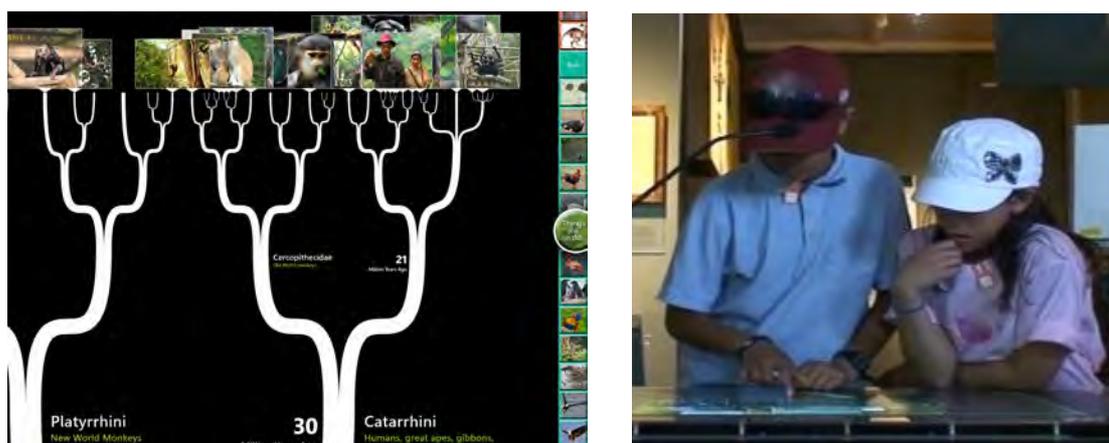


Figure 1: Screenshot from the *DeepTree* exhibit (left). A dyad (Gabrielle and Max) interacting with the exhibit at a natural history museum (right).

Learning Evolution

Studies have demonstrated a variety of challenges that learners face in attempting to grasp core concepts of evolution (see Rosengren, Brem, Evans, & Sinatra, 2012 for a review). These challenges are amplified in museums where engagement times tend to be short and visitors have freedom to move from one exhibit element to the next (Humphrey & Gutwill, 2005). Even depicting the evolutionary relationships of a small number of

species can be confusing for learners (Novick & Catley, 2012; MacDonald & Wiley, 2010). While we embrace the usefulness of simplified representations of scientific concepts (Davis, Horn, & Sherin, 2013), it can be difficult to convey the vast scale and dynamic processes of evolution using simplified static representations alone.

Information Visualization and Large Data Sets

Scientific organizations are actively compiling databases intended to describe all known species inhabiting the Earth today. Current estimates put the number of eukaryotic species around 8.7 million with additional millions of prokaryotic species (Mora, et al., 2011). These organizations face unprecedented challenges related to the processing and visualization of information on such a massive scale. To meet these demands, researchers across scientific disciplines are developing advanced computational methods to visualize information in order to find unexpected patterns and anomalies (Fayyad et al., 1996; Frankel & Reid, 2008). These methods will increase in importance as the capacity for data collection, storage, and communication expands. Due to this importance and utility, educators are beginning to find ways to leverage these tools to visualize large scientific data sets for public consumption in museums.

Interactive Tabletops

As information visualizations become increasingly important for scientific practices, they are slowly beginning to appear in museums in the form of interactive exhibits (e.g. Hinrichs et al., 2008). In particular, we argue for the utility of *interactive tabletops*—surfaces that allow direct touch interaction with a computational environment for multiple users. In recent years, interactive tabletops have moved out of research labs and into classrooms, museums, and public spaces. Preliminary research on the use of interactive tabletops to support collaborative learning has found that tabletop environments can promote physical engagement, reflection, and collaboration (e.g. Harris et al., 2009; Piper & Hollan, 2009; Rick et al., 2011; Shaer et al., 2011; Schneider et al., 2012). Furthermore, researchers have documented some of the interactional arrangements (Hinrichs & Carpendale, 2011) and group dynamics (Rick, Marshall, & Yuill, 2011) that shape interactive tabletops as collaborative forums. Because tabletops support multi-user interaction, they seem remarkably well suited for use in museums. However, while many tabletop museum exhibits now exist (e.g. Geller, 2006; Hornecker, 2008; Antle et al., 2011), few have been rigorously evaluated. In previous work, we have attempted to define measures for successful interaction with multi-touch tabletops and use these measures to evaluate our own design for a table-base evolution exhibit (Horn et al., 2012). In the current study we expand on this previous work to develop a more in-depth qualitative analysis of dyadic interaction around an interactive, information visualization exhibit.

DeepTree Design

The *DeepTree* exhibit is an interactive visualization of the tree of life showing the ancestral relationships of 70,000 species starting from the origins of life some 3.5 billion years ago (see Block et al., 2012, for more detail). *DeepTree* currently runs on a multi-touch Microsoft PixelSense™ surface. The exhibit was designed around five related learning goals: (1) All life on Earth is related; (2) biodiversity is vast; (3) relatedness is derived from common descent; (4) species inherit shared traits from common ancestors; and (5) evolution is ongoing and happens over very long periods of time.

The design has three major components (see Figure 1). The main display area allows visitors to zoom and pan through the entire tree of life using standard multi-touch gestures. Pulling the tree down from the top of the screen allows visitors to zoom in to reveal more information, starting from the root of the tree to its canopy, displaying individual species. Touching and holding an image of an organism causes the display to automatically “fly” through the tree to the selected species. The tree uses a fractal-based layout algorithm so that branches emerge as the user zooms in or out. Unlike static depictions of trees that simplify information by limiting the number of species, the fractal design allows for the depiction of every species in the tree of life while still reducing visual complexity.

The second component is a scrolling image wheel along the right side of the screen containing a subset of 200 species representing important evolutionary groups. Visitors scroll through the images to select and pull out any species onto the main display. When an image is held, a transparent chord points to the species' location in the tree and the system automatically flies toward it. Holding two images points toward both species' location, allowing visitors a glimpse at both species' relative positions on the tree of life.

The final component is an action button centrally located on the image wheel. When pressed the action button reveals a *relate* function that allows visitors to select any two species from the image wheel and the tree automatically highlights their shared lineage and flies to their most recent common ancestor. Once there, the tree prompts the learner to press an icon to initiate an embedded learning activity. This activity presents a simplified tree depicting the two species' shared lineage and highlighting major evolutionary speciation points. These points can be activated to reveal further information about common ancestors and major traits.

We developed *DeepTree* through an iterative process of design and evaluation with a team of computer scientists, learning scientists, biologists, and museum curators. Over the course of a year, we implemented and evaluated twelve prototype designs with 250 visitors in a large natural history museum.

Research Design

In order to evaluate the design, we placed a tabletop surface running *DeepTree* and another interactive design (called *FloTree*, which we discuss in Chua et al. 2012) in two prominent natural history museums in the United States (one in the Northeast and one in the Midwest). We recruited 250, 9- to 15-year-olds ($M=11.56$, $SD=1.68$; 126 females, 124 males) in dyads and randomly assigned them to one of four conditions. In condition A, participants interacted with both *DeepTree* and *FloTree*. In condition B, participants interacted with *DeepTree* only. In condition C, participants watched a 10-minute video about the tree of life (http://archive.peabody.yale.edu/exhibits/treeoflife/film_discovering.html). This comparison group was meant to reflect common, non-static museum exhibit design. Finally, condition D was a control group that received no intervention. For 10 minutes each dyad (except those in group D) freely interacted with one of the tabletop exhibits or watched the video. We video recorded children's physical and verbal interactions in order to capture discourse, behavior, and collaboration.

Following the interaction, all participants (group D included) were interviewed individually. Each interview lasted roughly fifteen minutes and involved open-ended and closed-ended questions about participants' ideas and understanding of evolution. To assess children's breadth of knowledge, we asked about common descent, common ancestry, natural selection, biodiversity, and the on-going nature of evolution. Parents completed a demographic form that included questions on their children's interaction patterns and a survey on their understanding of evolution. There were no significant differences between conditions in parent completion/non-completion of college, parents' or children's self-reported knowledge of evolution, religiosity, or compatibility of evolution with their religious beliefs.

The focus of this paper is on a qualitative analysis of dyadic interaction with the exhibit. Briefly, however, an analysis of close-ended responses revealed that dyads in both tabletop conditions were more likely than those in the control group to agree that humans, other animals, plants, and fungi had ancestors in common, a long time ago. Furthermore, dyads in condition B (*DeepTree* only) were most likely to interpret a tree of life graphic accurately and agree that all living things share DNA ($ps < 0.05$). All were multi-question measures (Evans et al., 2013). A full description of these results is forthcoming.

Descriptions of Interaction

As stated earlier, three research questions drive our analysis:

1. How do dyads negotiate their moment-to-moment exploration of the tabletop exhibits?
2. How do dyads negotiate meaning making through their interaction?
3. How do specific aspects of the exhibit design shape these meaning-making and exploration activities?

In order to begin the process of answering these questions, we adopted the frame of *interaction analysis* (Jordan & Henderson, 1995) that uses video as a primary data source and involves repeatedly viewing data in order to provide a deep analysis of the interactions that shape thought and behavior through talk, nonverbal cues, and artifacts. Based on this approach, we first created *content logs*—rough descriptions of the action with annotations of particularly compelling sections—of the videos. These logs guided analysis, in which we co-viewed the videos and discussed the *micro-level* interactions in order to isolate more general patterns of interaction. This analysis is ongoing and a fully representative account of interactions is beyond the scope of this paper. Nevertheless, in the following sections we describe preliminary interaction analysis using examples from three dyads representing differing levels of successful interaction with the table.

Negotiating Exhibit Exploration

Our first question concerns the ways in which dyads negotiate their exploration of the exhibit from moment-to-moment. Large tabletop displays support collaborative interaction that is potentially much different from other electronic devices. Without the constraint of a single input device (like a mouse or a keyboard), individuals are free to interact at any time, and, as is the case of the *DeepTree* exhibit, individual actions often affect the state of the entire system. For example, if one child decides to zoom or pan the display, the picture that the other child happens to be looking at can disappear. So, individual actions can work at cross-purposes, forcing dyads to frequently negotiate their exploration of the exhibit. This negotiation could be as simple as saying “wait” or physically grabbing the other's hand. We observed the formulation and execution of goals at different levels of granularity lasting from a few seconds to over a minute. An important dimension seemed to be whether or not goals were articulated and agreed upon. Based on our preliminary analysis, most dyads' collaborative exploration seemed to take one of three forms—*reactive*, *articulated*, or *contemplated* (Table 1). These forms do

not necessarily describe the overarching pattern of interaction, but rather a moment-by-moment analysis—a single dyad may employ one or more forms during their interactions.

Types of Goal Negotiation	
Reactive	Independent, moment-to-moment actions in response to the tabletop and other people
Articulated	Short-term goals are expressed (physically or verbally) and agreed upon
Contemplated	Longer-term goals that are articulated and verbally elaborated
Patterns of Meaning Making	
Serendipitous	Chance discovery through exploration of exhibit content
Making Connections	Drawing parallels between outside sources and exhibit content
Goal-oriented	Meaning making directed by the pursuit of overarching goals

Table 1. Forms of dyadic interaction in goal-negotiation and meaning making.

We call the first type of interaction *reactive* negotiation, because it seems driven by reciprocal reaction to immediate individual actions on the tabletop. This is especially evident in the dyad of Diego and Anna's (both 12) interaction with *DeepTree* (Figure 2). Diego and Anna seldom spoke while interacting with the exhibit and most of their moment-to-moment goals seemed independently construed. They appeared to learn how to use various aspects exhibit by watching each other, but their actions also frequently came into conflict, forcing momentary episodes of spontaneous negotiation. For example, at one point Diego begins resizing images using a two-finger spread and pinch. Anna is observing this, but not touching the screen. Anna notices the action button pulsating, points to it over Diego's arms and says, "Oh look." Diego pulls his arms back, looks at the action button, moves his hand towards the button as if to touch it, then looks back toward the main display while moving his arms back over Anna's, and continuing to resize the images. In this instance, Diego's goal— independent of Anna—is the manipulation of the images. When Anna notices and draws Diego's attention to the action button, through both a gesture and an utterance, she introduces a new goal. Diego momentarily considers this goal, employing a pointing gesture, before he wordlessly rejects it via a gesture that actively suppresses Anna's previous gesture. This is an example of two divergent goals clashing and requiring a negotiation between the actors, which takes the form of a brief consideration followed by cursory dismissal or acceptance. Throughout Diego and Anna's interaction we see a cycle of parallel goals conflicting when both require simultaneous use of the table and fleeting negotiations wherein one goal overrides the other, only to start the cycle anew. This arrangement of goals and negotiation is apparent in many of the dyads.



Figure 2: Dyads interacting with the DeepTree exhibit: Anna and Diego (right) and Chloe and Braden (left).

At other times the dyads actively *articulate* their goals through speech or gesture. *Articulated* negotiations generally involve less independent interaction than the *reactive* and sometimes result in mutual agreement on the goal. At one point, Chloe (9) and Braden (11) both begin tapping on images. Their taps result in the image enlarging and Chloe says, "Yeah, let's try that." They both then zoom in and out of the images together. Likewise, Leo (13) and Hope (9) begin their interaction with Leo explicitly asking, "Where do you want to start?" And Hope moving her hands over the table while saying, "Let's start... uhhhhh... here," and pointing at an image. Leo and Hope then work together to discover the deep zoom function of the tree. These kinds of articulations result in more joint action, even when both parties do not adopt the articulated goals. For example, Chloe and Braden are engaged together with a joint action of panning and zooming on individual images. After zooming in on a particular organism, a block of text appears with a description of the species. When this happens, Chloe begins to read the text, thus introducing a new goal. Braden does not adopt her goal and instead continues to pan and zoom. Chloe says she can't read the text and Braden moves his hand away so

the words are no longer obscured, however he continues to pan and zoom the image. In this case, their goals are independent, but they are able to simultaneously achieve their articulated goals, without the strong conflicts or dismissals seen in reactive negotiations. In these interactions, previously articulated goals seem to prevent some level of conflict and support mutual table use.

Occasionally dyads will vocalize explicit overarching goals for their explorations and then negotiate or refine these goals through relatively smooth verbal exchanges. We call this *contemplated* negotiation. *Contemplated* negotiation is similar to *articulated* negotiation, but instead of just being a moment-to-moment goal setting, it involves the setting of larger goals that result in more directed interaction. The dyad of Gabrielle (12) and Max (14) (Figure 1) frequently demonstrate this type of negotiation. Less than two minutes into their interaction with the table Gabrielle says, “Let’s try...” then glances at the pulsating action button, points at it and finishes, “let’s go to things you can do.” Max then presses the button and chooses the *relate* function (Figure 3). Gabrielle then says, “Ok, relating to...? What could we relate to?” In this exchange they have quickly negotiated an overarching goal of “relating to” for their activity, and for the rest of their interaction they only use the *relate* function. Because they have established this higher order goal, they only need to negotiate the specifics of its enactment. After trying several different relations, the following exchange takes place:

Gabrielle: Let’s try... Maybe something that you would think would be the total opposite. See, if, some—somewhere that you think would be the total opposite that you think would never relate.

Max: So something with four legs, or no legs.

Gabrielle: Yeah.

Max: So, let’s try a fish.

Gabrielle: Against a four-legged animal, ok.



Figure 3: The *relate* function compares two species.

In this exchange, we can see that they are still working under the “relate to” goal, but Gabrielle suggests a refinement on their goal, and one that she seems to think will have surprising results. Max agrees and proposes a more specific comparison to work from. Gabrielle then agrees, Max suggests an animal, and Gabrielle offers a comparison. This dialogic agreement and back-and-forth building of a goal allows them to demonstrate that they both understand and can engage the new task. Furthermore, it shows that even though one actor suggested the new goal there is no “leader” in the task and they must work together. We argue that their setting of the higher order goal in the beginning guided the moment-by-moment exploration and allowed it to run smoothly. In other words, having an overarching activity in place puts them both on the same page, so the possible space of sub-goals is constrained and easier to articulate when small disagreements arise. While *reactive* and *articulated* goals generally seem to correlate with an undirected exploration of the table, *contemplated* goals appear to lend themselves to experimentation. In the above exchange, Gabrielle’s sub-goal is presenting an implicit hypothesis—opposites are not related. All of their sub-goals seem like mini-tests of their hypotheses about relationships. Do these collaborative goals actually help the dyads make sense of the content of the exhibit? In the next section, we discuss the patterns of meaning making that we see in the interactions.

Negotiating Meaning

Our second question focuses on how the dyads collaborate to make meaning from the content presented by the exhibit. It has been argued that an important aspect of collaboration is *convergence*—how people construct shared meaning through their interactions (Roschelle, 1992). For surface level understandings convergence may be quite easy to achieve. One person reading a label out loud and another person overhearing and applying the label in order to name an animal is a relatively simple convergence of meaning (and one that is common in museums). However, working towards deeper conceptual change involves progressively more complex systems of convergence. Convergence is also pragmatic—meaning that individuals develop specific strategies moment to moment as they negotiate the meaning with one another and the exhibit. Across our dyadic interaction data we identified three broad patterns in the way children construct an understanding of the content: through *serendipitous discovery*, by *making connections* with prior experiences, and through cooperative, *goal-oriented discovery* (Table 1).

By *serendipitous discovery* we mean that children gained insights about evolution by chance exploration. For example, as part of the exhibit’s design when a player holds her finger down on a picture of a species the screen zooms in for a closer and closer view of that species’ location in the tree of life. This “fly-through” motion zooms past more distantly related species before homing in on the closest relations. This whole

process can take between fifteen and thirty seconds—a seemingly long period of time in the course of interaction. However, during that time players catch a glimpse of hundreds of related species and gain an appreciation for the diversity of species in the tree of life. It is common for players to express their surprise and excitement at seeing so many species fly past. For example, at the start of Chloe and Braden’s interaction, Braden holds his finger down on the “modern human” picture. After a few seconds of zooming, he says, “wow,” followed shortly by “hey, how far is this?” After twenty-five seconds of zooming in, Braden and Chloe finally land on “modern human”—having seen hundreds of other species along the way. Braden expresses his appreciation for that huge number by saying, “that was a lot!” Chloe agrees with this, saying, “Yeah!” We consider this type of discovery to be serendipitous because it is not intentional. Braden’s goal was to find out what would happen if he zoomed all the way in on the species “modern human.” Braden and Chloe converge on an understanding of the great biodiversity of the tree of life directly related to the design of the exhibit.

Chloe and Braden also offer an excellent example of the second pattern of meaning making: *making connections* with prior experiences. In this example Chloe verbalizes a connection she makes between a concept she encounters on the table and at another exhibit in the museum. The other museum exhibit is a working laboratory where museum scientists study DNA. Visitors can look through a glass panel and watch the scientists at work. During one interaction with the exhibit, Chloe and Braden come across a picture of DNA when they are using the *relate* feature. Chloe says, “I guess that’s like the stuff we saw in the glass,” and Braden replies, “It’s DNA.” Chloe then points to the picture of DNA and says, “Cool, that’s the molecules I think.” Braden restates, “Yeah, that’s the DNA.” Chloe ends the conversation by saying, “So DNA was what they were studying in the thing.” In this example Chloe seems to recognize the picture of DNA as something she also saw in the other exhibit. Chloe and Braden converge on a (simplified) understanding of DNA molecules through increasingly more specific talk that draws on the greater museum context to make meaning of the information on the table.

Serendipitous discovery and making connections are patterns seen on a moment-to-moment basis and at any time during interaction with the exhibit. The third pattern we identified, *goal-oriented discovery*, is a result of a dyad’s broader goal negotiation. As discussed earlier, we found that Max and Gabrielle frequently negotiate contemplated goals for the exhibit. This in turn allowed them to jointly develop a big picture goal that drove their interactions with the visualization. While in pursuit of their higher-level goal, Max and Gabrielle took advantage of specific opportunities to make meaning about smaller components of the big picture. For example, during their second trial using the *relate* function, Max and Gabrielle relate modern humans to clown fish. This is part of their contemplated goal to compare opposite kinds of species. A simplified tree appears on the screen showing the traits that humans and clown fish have in common. Max reads aloud the text on one of the branches of the tree, “Jaws perfected to chew food.” Gabrielle says, “Okay. Yeah we can chew.” Then she asks, “So apparently they can too, right?” In this example, Max and Gabrielle are reasoning about how to read this new diagram. They know that the diagram mentions jaws for chewing and they also know that humans can chew. In order to converge on the conclusion that fish can chew—as Gabrielle does when she says, “Apparently they can too”—the pair needs to understand that the graphic shows shared traits between the two species. In fact, later in the session they use this understanding about the graphic again when they are comparing humans with bacteria. Max expresses surprise when the modified tree appears and shows that humans and bacteria have only one trait in common—that they are both made of cells. Max says, “So basically they’re the exact opposites.” Looking down at the graphic, Gabrielle adds, “Yeah, but they’re living cells. That’s pretty much it.” Here again their interpretation of the graphic allows them to make meaning about the relatedness of different species. Furthermore, this example of meaning making is nested into the pursuit of their larger goal. In this case, Max and Gabrielle’s convergence on more surface level meaning (Fish have jaws), allows them to also converge on an understanding of a higher order evolutionary concepts (common descent).

Design Supports for Exploration and Meaning Making

Our third research question relates to the role of design in shaping children’s collaborative interaction around the tabletop. Specifically, given the diverse types of exploration and meaning-making activities that we observed, how does the *DeepTree* design function to make visitor experiences more worthwhile?

Suchman (2007) uses an analogy of a person confronting river rapids in a canoe to help illustrate the concepts of *planning* and *situated action*. We extend this analogy to consider dyads interacting with the *DeepTree* exhibit. Imagine two inexperienced paddlers in a tandem kayak floating in the middle of a large body of water. Each person has a paddle that can be used with immediate effect—move the paddle in the water and the boat moves in response, if not necessarily in a predictable way. Because both kayakers are inexperienced, they are still learning how to most effectively steer the boat in a desired and consistent direction. And, since both paddlers are interacting at the same time, coordination is required. This is complicated by the fact that it can be difficult to figure out how each person is causing the boat to move if both partners are paddling at the same time. So, the kayakers must simultaneously figure out how to use the paddles (the interface), decide on a mutually agreeable direction (a goal), and figure out how to coordinate actions (negotiation and reciprocal learning). Inevitably, novice paddlers spend a period of time splashing around and not making much progress in

any observable direction. We hope that the relationship between the tandem kayak and dyadic interaction with tabletop exhibits is clear. The body of water corresponds to the information space that visitors can explore with the *DeepTree* exhibit. The paddlers are the youth themselves, and the paddles are their fingers, hands, and arms (the input devices).

With this analogy in mind, we see two critical roles for design. The first relates to the body of water. Effective design shapes this open expanse into a river with a gentle but persistent current. Along this river are landmarks or points of interest. This relates to the second goal of design, which is to include a collection of appealing and strategically placed features that invite attention. Earlier we mentioned that Diego and Anna seem to be reactive in their goal negotiation—they are like two rowers each paddling in their own direction, at their own speed, and with their own intentions. This could result in a great deal of effort with no discernible outcome. However, because opposing movements on the table cancel each other out, the table forces their goals into conflict, requiring them to negotiate and coordinate their efforts. In fact, Diego and Anna's independent movements result in the table zooming. While this was not either of their intentions, the result causes them to both hold the zoom to fly through the tree and have the same “wow” experience we saw earlier with Chloe and Braden. In this instance, the exhibit design guided their exploration and in so doing allowed them to spontaneously find and make meaning out of a “landmark”—the fly-through that portrays massive biodiversity. So, for learners in a reactive form of goal setting, the table guides the exploration in a persistent direction toward interesting information—just as the river's current pulls rowers past interesting viewpoints downstream. In other words, even if reactive dyads, like Diego and Anna, still tend to explore surface features of the exhibit; nonetheless, the exhibit design elicits apparently spontaneous meaning making and leads them to some level of understanding of evolutionary concepts.

What about dyads who already articulate or contemplate goals that are also supported by the exhibit design? As previously discussed, Gabrielle and Max explicitly articulate higher-level goals that drive their moment-by-moment interaction with the exhibit. This dyad can be viewed as tandem rowers who are in harmony in terms of the direction they wish to follow (even if they are still learning how to paddle more effectively). They work together to explore and experiment with the exhibit and to discover meaning, directed by their articulated goals. But, just as with the discordant rowers, the exhibit is not merely an inert tool for synchronized rowers. Though Gabrielle and Max control the direction of their kayak, the current of the river brings them to their goal more rapidly than they could have achieved on their own. Contemplated dyads, such as Gabrielle and Max, quickly move past the surface level, and the exhibit guides them to a feature, such as the *relate* function, which allows them to surge more deeply into the content and construct richer understandings.

The design of the *DeepTree* exhibit affords many strategies for goal negotiation, and both spontaneous and contemplated meaning making. Some meaning making, such as making connections to outside knowledge, is not directly supported, but by driving collaboration in the service of convergence, *DeepTree* encourages learners to find meaning through whatever interactive strategy they happen upon.

Conclusion and Future Work

Despite its centrality to modern biology, evolution remains a challenging subject for learners and its understanding persists to be an elusive goal of science education—particularly in informal educational environments. In this paper, we have examined a design for a novel museum exhibit that conveys the rich complexities of dynamic evolutionary processes through an interactive visualization. In a study, we find that children negotiate their exploration of the exhibit in a variety of ways including *reactive*, *articulated*, and *contemplated* exploration, and that these negotiations impact the ways children make meaning from the exhibit content and their interactions with one another. We argue that particular aspects of the design guide visitors in their interaction and collaboration. For example, the “fly-through” motion supports the *serendipitous discovery* of biodiversity, while the *relate* function encourages experimentation and the *goal-oriented discovery* of common descent. By encouraging a flow through the exhibit and providing specific landmarks for discovery, the *DeepTree* exhibit allows learners to make sense of evolution through their own free choice interactive techniques. In future work we plan to operationalize the framework proposed in this paper and systematically apply it to all of the video data collected in our study in order to help build and strengthen theories on collaborative learning in science museums.

References

- Allen, S. (2004). Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education*, 88(S1), S17-S33.
- Antle, A.N., Bevans, A., Tanenbaum, J., Seaborn, K., & Wang, S. (2011). Futura: design for collaborative learning and game play on a multi-touch digital tabletop. In *Proc. Tangible, Embedded, and Embodied Interaction (TEI'11)*, ACM, 93-100.
- Block, F., Horn, M.S., Phillips, B.C., Diamond, J., Evans, E.M., Shen, C. (2012). The DeepTree Exhibit: Visualizing the tree of life to facilitate informal learning. *IEEE Transactions on Visualization and*

- Computer Graphics*, 18(12), 2789-2798
- Chua, K. C., Qin, Y., Block, F., Phillips, B., Diamond, J., Evans, E.M., Horn, M., Shen, C. (2012). FloTree: A multi-touch interactive Simulation of Evolutionary Processes. In *Proc Interactive Tabletops and Surfaces (ITS'12)*, ACM, 299-302.
- Davis, P.R., Horn, M.S., & Sherin, B.L. (2013). The right kind of wrong: A “Knowledge in Pieces” approach to science learning in museums. *Curator: The Museum Journal*, 56(1), 34-41.
- Evans, M.E., Phillips, B., Horn, M., Block, F., Diamond, J., & Shen, C. (2013). Active prolonged engagement: When does it become active prolonged "learning"? 2013 Biennial Meeting of the Society for Research on Child Development. April 10-13, 2013. Seattle, Washington.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37.
- Frankel, F., & Reid, R. (2008). Big data: distilling meaning from data. *Nature*, 455(7209), 30-30.
- Geller, T. (2006). Interactive tabletop exhibits in museums and galleries. *Computer Graphics and Applications, IEEE*, 26(5), 6-11.
- Harris, A. Rick, J., Bonnett, V., Yuill, N., Fleck, R., Marshall, P., & Rogers, Y. (2009). Around the table: Are multiple-touch surfaces better than single-touch for children’s collaborative interactions? In *Proc. CSCL '09*. ACM.
- Hinrichs, U., Schmidt, H. & Carpendale, S. (2008). EMDialog: Bringing information visualization into the museum. *IEEE Transactions on Visualization and Computer Graphics*, 14(6), 1181-1188.
- Hinrichs, U., & Carpendale, S. (2011). Gestures in the wild: studying multi-touch gesture sequences on interactive tabletop exhibits. In *Proc. Human Factors in Computing Systems (CHI'11)*, ACM, 3023-3032).
- Horn, M., Atrash Leong, Z., Block, F., Diamond, J., Evans, E. M., Phillips, B., & Shen, C. (2012). Of BATs and APEs: an interactive tabletop game for natural history museums. In *Proc. Human Factors in Computing Systems (CHI'11)*, ACM, 2059-2068.
- Hornecker, E. (2008). “I don’t understand it either, but it is cool.” Visitor interactions with a multi-touch table in a museum. In *Proc. IEEE TABLETOP 2008* (pp. 113-120). IEEE.
- Humphrey, T., & Gutwill, J. P. (2005). *Fostering active prolonged engagement: The art of creating APE exhibits*. San Francisco: Left Coast Press.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4(1), 39-103.
- MacDonald, T., & Wiley, E.O. (2010). Communicating phylogeny: Evolutionary tree diagrams in museums. *Evolution: Education and Outreach*, 1-15.
- Miller, J.D., Scott, E.C., & Okamoto, S. (2006). Public acceptance of evolution. *Science*, 313, 765–766.
- Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G., & Worm, B. (2011). How many species are there on Earth and in the ocean? *PLoS biology*, 9(8).
- Novick, L. R., & Catley, K. M. (2012). Reasoning About Evolution’s Grand Patterns: College Students' Understanding of the Tree of Life. *American Educational Research Journal*.
- Piper, A. M., & Hollan, J. D. (2009). Tabletop displays for small group study: affordances of paper and digital materials. In *Proceedings of the 27th international conference on Human factors in computing systems* (pp. 1227-1236). ACM
- Rick, J., Marshall, P., & Yuill, N. (2011). Beyond one-size-fits-all: How interactive tabletops support collaborative learning. In *Proceedings of the 10th International Conference on Interaction Design and Children* (pp. 109-117). ACM.
- Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences*, 2(3), 235-276.
- Rosengren, K. S., Brem, S. K., Evans, E. M., & Sinatra, G. M. (Eds.). (2012). *Evolution challenges: Integrating research and practice in teaching and learning about evolution*. New York: Oxford University Press.
- Shaer, O., Strait, M., Valdes, C., Feng, T., Lintz, M., & Wang, H. (2011). Enhancing genomic learning through tabletop interaction. In *Proc. Human factors in Computing Systems (CHI'11)*, ACM, 2817-2826.
- Schneider, Bertrand, Strait, M., Muller, L., Elfenbein, S., Shaer, O., & Shen, C. (2012). Phylo-Genie: engaging students in collaborative “tree-thinking” through tabletop techniques. In *Proc. Human factors in Computing Systems (CHI'12)*, ACM, 3071–3080.
- Suchman, L. (2007). *Human-Machine Reconfigurations: Plans and Situated Action*. Cambridge: Cambridge University Press.

Real-time Collaboration for Web-Based Labs

Luis de la Torre, Ruben Heradio, Sebastian Dormido, Systems Engineering and Signal Theory Department,
Spanish Open University, Spain

Email: ldelatorre@dia.uned.es, rheradio@issi.uned.es, sdormido@dia.uned.es

Carlos Jara, Computer Science and Automatic Department, University of Alicante, Spain, carlos.jara@ua.es

Abstract: Web-based labs are key tools for distance education that help to illustrate scientific phenomena which require costly or difficult-to-assemble equipment. We propose the extension of two open source tools: (i) the learning management system Moodle, and (ii) the application to create web-based labs Easy Java Simulations (EJS). Our extension provides: (i) synchronous collaborative support to any lab developed with EJS (i.e., existing labs written in EJS can be automatically converted into collaborative with no cost), and (ii) support to deploy synchronous collaborative labs into Moodle. Thanks to our approach, students and/or teachers can invite other users enrolled in a Moodle course to a real-time collaborative experimental session, sharing and/or supervising experiences at the same time they practice and explore experiments using labs. The experimental evaluation of our work shows statistical significant (i) increase in student engagement and (ii) higher exam grades for students trained with collaborative labs.

Introduction

It is commonly accepted that digital media (such as simulations, videos, interactive screen experiments or web labs) can positively impact student knowledge, skills and attitudes (Kozma, 1994). Consequently, tools such as *Learning Management Systems* (LMSs) and *web-based labs* have become widespread in distance education in the last decade. LMSs support the administration, documentation, tracking, and reporting of training programs, classroom and online events (Ellis, 2009). Web-based labs make possible to illustrate scientific phenomena that require costly or difficult-to-assemble equipment (Chang et al., 2005). There are two complementary approaches for web-based labs:

1. *Virtual Labs* provide computer based simulations which offer similar views and ways of work to their traditional counterparts (Guimaraes et al., 2011). Nowadays, simulations have evolved into interactive graphical user interfaces where students can manipulate the experiment parameters and explore its evolution.
2. *Remote Labs* use real plants and physical devices which are teleoperated in real time (Wannous, 2010).

Even though constructivist web learning environments and Virtual/Remote Labs (VRLs) already exist, there is still a lack of: (i) convergence and interoperability between both tools (Gravier, 2008), and (ii) real-time interaction between users when they work with VRLs (Gravier et al., 2008; Ma & Nickerson, 2006) and/or within a LMS. We consider that several advantages could be achieved covering these two drawbacks, especially for distance education of practical experiences in technical or scientific subjects. This paper presents a new approach that solves this scientific gap. Currently, there are two different types of collaborative environments according to the moment when the student-student (or student-teacher) interaction takes place: asynchronous and synchronous (Bafoustou & Mentzas, 2002). The first ones allow data exchange in flexible timetables and remote access to web-based course materials to carry out activities in an asynchronous way. They use collaborative tools such as e-mail or forums for on-line communication. This is the typical approach and tools offered by most classic LMS. However, this type of communication can cause feelings of isolation in the student and hence reduces his/her motivation (Boulos et al., 2005). Furthermore, students do not receive instant feedback from their questions and cannot talk in real-time about results obtained in the learning activities. These limitations have been solved by applying synchronous technologies (Marjanovic, 1999), as we have performed in the approach presented in the paper.

It is from the intersection of these previous ideas that the concept of synchronous collaborative VRLs deployed into LMSs is born. The approach presented is based on in this concept by means of the use of two valuable software applications for e-learning and VRL development: Moodle and Easy Java Simulations (EJS). Moodle is a widespread used LMS (more than 50 million registered users, according to its official webpage) that supports constructivist learning, offering its users communication and interaction facilities. EJS (Christian & Esquembre, 2007; Christian et al., 2011) is a tool designed for the creation of discrete computer simulations. During the last few years, EJS has grown for helping to create web-accessible labs in control engineering education. With this objective in mind, recent releases of EJS support connections with external applications, such as LabView and Matlab/Simulink. Hence, EJS not only is useful to create virtual labs, but also the GUIs of their remote counterparts (Heradio et al., 2011).

This paper describes an extension for Moodle and EJS we have developed to provide synchronous collaborative support to any VRL developed with EJS, i.e., thanks to our extension, any existing VRL written in

EJS can be automatically converted into a collaborative lab with no cost. Our approach not only supports the teacher's presentation or explanation of course material by emulating a traditional classroom on the Internet. More interestingly, it also supports collaborative learning activities centered on students' exploration or application of the course material through VRLs. That is, students working in groups of two or more, mutually searching for understanding, solutions, or meanings.

We have evaluated our approach on a course of Experimental Techniques in Physics at the Spanish Open University (UNED), where students voluntarily performed lab assignments using VRLs. The results show (i) a correlation between the student exam grades and the number of completed lab assignments, (ii) that the collaborative feature we offer encourages student engagement (i.e., students that use the collaborative feature tend to complete more lab assignments), and (iii) that our synchronous collaboration approach helps to make the most of the lab assignments (i.e., students trained with collaborative labs get better exam results than those trained with non-collaborative labs).

Synchronous Collaborative VRLs

Moodle includes a good number of tools that provide asynchronous collaborative support (e.g., forums, the messaging system...). Our proposal takes advantage of such features by deploying VRLs into Moodle. In addition, we enrich Moodle collaborative support by providing a new feature: the synchronous collaboration among the VRLs that are included into a Moodle course. Our approach satisfies the following requirements:

1. **Supporting Deep Collaboration.** To the extent of our knowledge, existing proposals on synchronous collaborative VRLs limit collaboration to multimedia streams coming from the equipment server and from the users (Bochicchio & Longo, 2009). Thus, the only shared elements are audio, video and/or images. Under our approach, VRLs are deployed into Moodle, which has several plugins to provide synchronous sharing of audio, video and images (e.g., the Skype module available on http://docs.moodle.org/22/en/Skype_module). Therefore, our proposal supports such type of synchronous collaboration. In addition, our approach provides a higher collaboration level. For each participant in a collaborative session, there is a running instance of the shared VRL. The state of all the instances is synchronized, i.e., whenever a participant acts over its VRL instance, the changes produced on the VRL state are propagated to the remainder of the participants' VRL instances. For instance, Figure 1 shows a collaborative version of the "Three Tank" VRL (Dormido et al., 2008), which helps control engineering students to learn in a practical way many fundamental aspects of control processes. In the figure, two students work together to parametrize a Proportional-Integral-Derivative (PID) controller to get an overshoot and a settling time smaller than 20% and 1000s in Tank 1 and 15% and 500s in Tank 2, respectively. The areas in the Figure labeled "Virtual Lab" and "Remote Lab" visualize the lab state (i.e., the level of liquid in the tanks). Note that such state is the same for both students. Thus, although there are running two instances of the VRL, the students have the feeling of being working on the same VRL.
2. **Maximizing Software Reuse.** Building a VRL from scratch is too expensive, so it should be avoided "reinventing the wheel" every time a VRL requires collaborative support. Thanks to our approach, any existing VRL created with EJS is automatically converted into a collaborative lab by just clicking a single button. Thus, VRL developers can be focused on creative activities, avoiding the routine ones.
3. **Usability.** Our approach provides a high level of usability (i.e., the ease of use and learnability of a human-made object) for all the existing roles in the development, management and use of VRLs:
 - a. The *VRL developer* creates VRLs by using EJS. Thanks to the EJS extension we have built, any VRL can be automatically converted into a collaborative lab by just clicking a single button.
 - b. The *LMS administrator* deploys VRLs into Moodle, controls user access to the deployed labs, and performs maintenance activities related to the labs (e.g., VRL backup and restore). Such functionalities are graphically supported by our Moodle extension.
 - c. The *teacher* and the *students* participate in collaborative sessions by using an adaptive visual interface. That is, to simplify the user interface and prevent errors, the interface dynamically changes to only make available the correct options for a given state of the collaborative session. For instance, a student visualizes the "participate as an invited student" button (Figure 3.a) only when s/he has been previously invited to a collaborative session.
4. **Scalability.** Our approach is highly scalable, i.e., many collaborative sessions may be running at the same time. We have adopted a peer-to-peer (P2P) approach which avoids that multiple collaborative sessions overload the server that host the Moodle portal and the VRLs. When a collaborative session begins, users just interact with the server by downloading the applet that implements the VRL they are going to use in the session. Then, an instance of the applet is locally run in each participant's computer. The instances communicate each other through a server-less collaboration over TCP and UDP protocols. Thus, the communication between the server and the participants' computers is limited to simple messages of session creation, session pause, session close, etc.

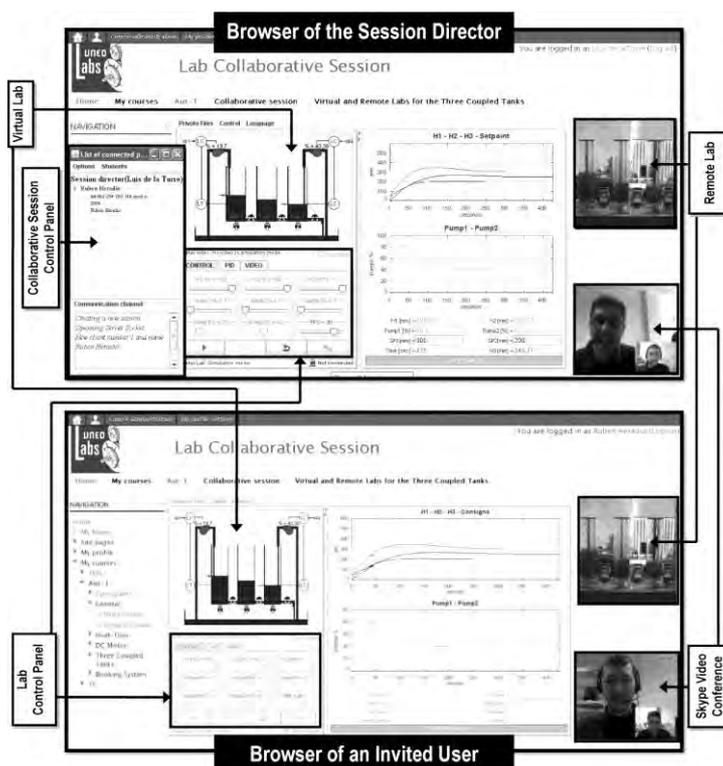


Figure 1. Example of Two Users participating in a Collaborative Experimental Session.

A fundamental issue in a synchronous collaborative system is the *Floor Control* (Dommel & Garcia-Luna, 1997). This term points out how the system components share the computational resources. Our proposal tries to offer shared VRLs that can be controlled in real-time by the different members of a virtual class. In our case, the shared VRL is composed of a Java applet generated with EJS. There are two main kinds of components to coordinate: one *session director's* applet and some *invited user's* applets. The session director is responsible for starting, monitoring and closing a collaborative session. Thanks to the Moodle and EJS extensions we have developed, the session director's applet manages in real-time the virtual class and synchronizes all the invited user's applets. S/he has a list of invited users connected to the virtual session and can disconnect any invited user's at any moment. In order to have a suitable floor control, connected invited user's applets are locked and they cannot interact with the shared VRL in a first moment. They are only allowed to see in real-time what the session director is doing in the shared application. This way, the collaborative session avoids collisions of events which can cause unwanted and incoherent results. One example of this problem could be that the real equipment which controls the VRL becomes uncontrollable because of unsuitable user interactions.

Extending Moodle

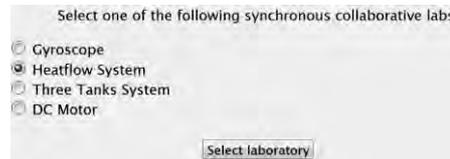
In the following lines, the behavior of the *EJSApp Collab Session* block, which extends Moodle to support synchronous collaborative sessions of VRLs, is illustrated:

1. **From the session director point of view**, a collaborative session is composed of the following steps:
 - a. A session is created by clicking the button "Create collaborative session" (Figure 2.a).
 - b. The session director selects then the potential participants to the session he is creating (Figure 2.c). Selected participants are *potential* in the sense that they may or may not decide to participate into the session. When the "Invite participants" button is clicked, they will be notified with an e-mail and a Moodle internal message.
 - c. The VRL is accessed in collaborative mode, i.e., the session director's applet manages the virtual class and synchronizes all the invited user's applets.
 - d. The collaborative session is finished by clicking the "Close collaborative session" (Figure 2.d).
2. **From an invited user point of view**, a collaborative session is composed of the following steps:
 - a. Once invited, the user clicks on the button "Participate as an invited student" (Figure 3.a). To prevent misuses of EJSApp Collab Session, its graphical interface changes to show just the valid choices available to a given situation (see Figures 2.a, 2.d and 3.a). So, the "Participate as an invited student" button is only visible because the user has been invited to, at least, one collaborative session.

- b. From all the received invitations, the user selects in which session s/he wants to participate (Figure 3.b). Note that a course member can be invited to several collaborative sessions, but s/he can only participate in one of them at the same time.
- c. The VRL is accessed in collaborative mode.
- d. The user stops participating in the session when (i) s/he decides to leave it or (ii) when the session director closes it. In the former case, the user is free to enroll either to that session again or to any of the other current available invitations.



(a) Starting a Synchronous Collaborative Session.



(b) Selecting the Collaborative VRL to be used within the Session.



(c) Selecting the Session Participants.

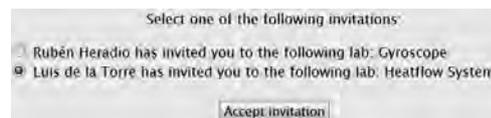


(d) Closing a Collaborative Session.

Figure 2. Synchronous Collaborative Session in Moodle from the Session Director Point of View.



(a) Accepting an Invitation to a Collaborative Session.



(b) Selecting one of the Available Invitations.

Figure 3. Synchronous Collaborative Session in Moodle from an Invited User Point of View.

Extending EJS

We have extended EJS to add synchronous collaborative support to any VRL developed with this tool. The last EJS release, its version 4.3.7, includes the collaborative approach described in this paper. This is done by *TCP* and *UDP* connections that periodically share and synchronize the VRL state of the session director with the VRLs of the invited users. The extension provides the session director, as an additional feature related to the synchronous collaboration, with the “Collaborative Session Control Panel” shown in Figure 1. This panel includes a list of the invited users connected to the collaborative session (e.g., control panel in Figure 1 shows that “Luis de la Torre” is the session director and “Ruben Heradio” is an invited user). Using such list, the session director can perform the following tasks:

1. Supervising which users have already connected to the collaborative session in order to call the roll before starting the experimentation.
2. Disconnecting any invited user at any moment.
3. Assigning the *chalk* to an invited user. With this feature, the session director gives permission to control the shared VRL to a specific invited user, by selecting him from the list. The chalk enables a student to manage the VRL, but not the collaborative session (i.e. the control panel is always commanded by the session director).

Figure 4 depicts the communication framework that underlays the collaborative sessions. When a session begins, users just interact with the Moodle server by downloading the applet that implements the VRL they are going to use in that session (see dashed lines in Figure 4). On the other hand, users participating in a session interact each other through a server-less collaboration over TCP and UDP protocols (see solid lines in Figure 4). Thus, the communication framework we propose supports multiple simultaneous sessions without overloading the Moodle server.

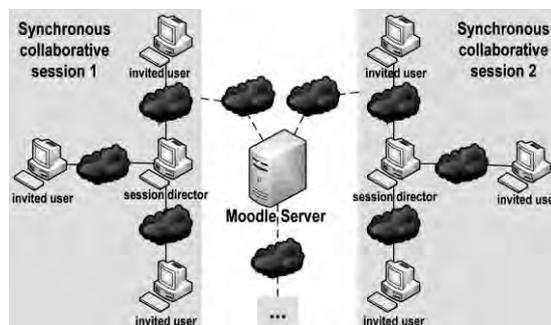


Figure 4. A Network of Collaborative Sessions

Invited users' applets are connected directly to the session director's applet in a P2P centralized overlay network. In contrast with server-based approaches, our e-learning system is focused in a server-less architecture. This communication method avoids delays caused by the server processing in the data flow because the communication engine is embedded in the Java applets downloaded by the users. In addition, the number of network connections can be substantially decreased because the session director's applet can manage the session, the floor control, and the data exchange having higher control over the invited user applets. As stated, the P2P network is centralized around the session director's applet. This last application is the central node of the collaborative class and contains a multithread communication module which manages the synchronization of all the applets that compose the shared VRL. Invited users' applets are connected to the central node over TCP and UDP sockets performing a centralized network.

To synchronize in real-time all the applets connected to the virtual class, a method based in Java object tokens (Dommel & Garcia-Luna, 1997) is used. Java object tokens are small update messages which contain a String object that defines the action to be performed by other applets of the same session. The small amount of sent information optimizes the network use and reduces the connection delay.

Since all the applets must be in the same state at any time, it is necessary to synchronize them. The developed communication framework provides a transport service suitable for all update data: a TCP-based channel for reliable messages and a UDP-based channel for fast messages. The TCP channel is used to update all the variables of the application because the transmission of the values needs the reliability of an ACK-based protocol. The UDP channel is used to transmit the small changes in the user-interface and this requires to be quickly updated in the rest of the applets.

Experimental Evaluation

In terms of number of students, the Spanish Open University (UNED), with more than 260,000 scholars, is the biggest university of Spain and the second one of Europe, next to the English Open University. To support their students, UNED is composed of a network of associated learning centers scattered around the world (more than 60 centers distributed across Spain, Europe, America and Africa). Unfortunately, the geographical dispersion of the students makes impossible to provide the scientific courses of UNED with traditional labs at a reasonable cost. Since the nineties, the Department of Computer Science and Automatic Control of UNED has been very concerned about this problem and has been working in new ways to illustrate scientific phenomena that require costly or difficult-to-assemble equipment. The UNEDLabs web portal (<http://unedlabs.dia.uned.es>) is the fruit borne by such work. It hosts a rich network of VRLs for students of UNED and other Spanish Universities, such as the Leon University and the Almeria University. All VRLs in UNEDLabs have been developed using the approach described in this paper. This section reports the experimental evaluation of our work on a course of *Experimental Techniques in Physics* supported by UNEDLabs.

Participants

The experimental evaluation of our approach was performed on two consecutive academic courses of *Experimental Techniques in Physics* at UNED: 2010-11 and 2011-12. In both years, students were encouraged to carry out five voluntary lab assignments supported by the following VRLs:

1. A motorized rotatory laser to illustrate the Snell's law (de la Torre et al., 2011).
2. A motorized optical bench to estimate the focal of thin lenses.

3. A Hooke’s law simulator (de la Torre et al., 2011).
4. A Geiger counter based VRL to experiment with radioactive disintegration laws.
5. An RC Circuit.

Whereas the 2010-11 course had 53 students and the lab assignments were individual (i.e., no collaborative support was available), the 2011-12 course had 62 students and the assignments were performed in groups of two/three students by using the collaborative features described in this paper. Table 1 and Figure 5 describe the dataset of our experimental evaluation, which is composed of the number of lab assignments completed by the students and their grades on the course final exam (note that exam grades are rated on a 10-point scale).

Table 1: Dataset Descriptive Statistics.

		Mean	Standard Deviation	Median	Skew	Kurtosis
Course 2010-11	Exam Grades	3.91	2.50	3.00	0.56	-0.52
	Number of Completed Lab Assignments	1.53	1.75	1.00	0.92	-0.53
Course 2011-12	Exam Grades	5.40	2.98	6.00	-0.04	-1.49
	Number of Completed Lab Assignments	2.79	2.10	3.00	-0.19	-1.65

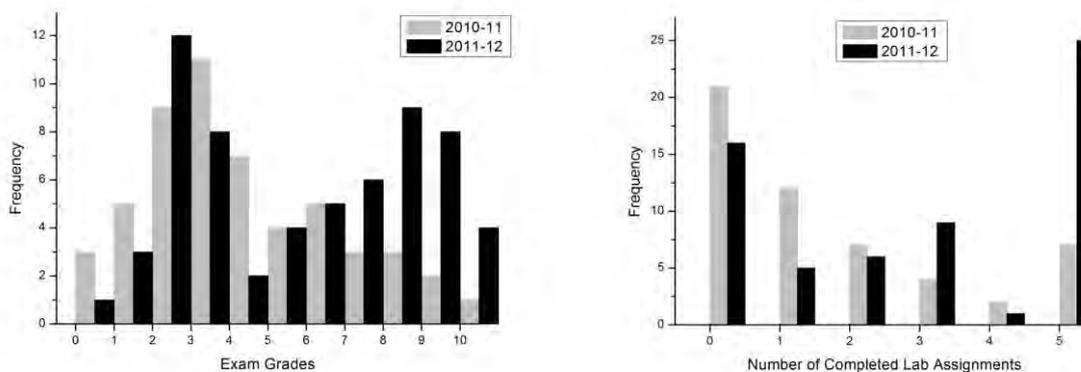


Figure 5. Dataset Histograms.

Results

The Exam Grades and the Number of Completed Lab Assignments are Correlated

The scatter plot in Figure 6 depicts the relationship between the number of completed lab assignments and the exam grades for both courses. Since there are many data points (53+62=115) and significant overlap among them, points have been grouped into colored hexagonal cells. The color range goes from light grey (one single point) to black (when a cell groups 16 points). In addition, Figure 6 includes the linear regression lines of (i) the courses 2010-11 and 2011-12 separately, which just take into consideration their corresponding 53 and 62 points, respectively; and (ii) both courses jointly. Table 2 summarizes the correlation tests of the relation between assignments and exam grades. Since the *p*-values are minor than 0.01, the tests show that the correlation is statistically highly significant.

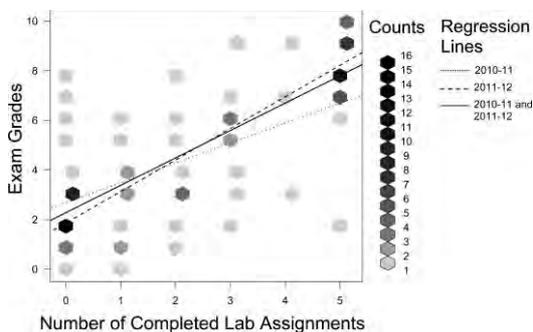


Figure 6. Scatter Plot and Regression Lines for the Dataset.

Table 2: Correlation and Regression Lines between Exam Grades and Completed Lab Assignments.

Courses	Pearson’s product-moment correlation			Regression Line Grade = B ₀ + B ₁ *NumberOfAssignments	
	Correlation factor <i>r</i>	<i>t</i>	<i>p</i> -value	B ₀	B ₁
2010-11	0.561544	4.8465	1.22e-05	2.6804	0.8017
2011-12	0.8941395	15.467	< 2.2e-16	1.854	1.272
2010-11 and 2011-12	0.7877397	13.593	< 2.2e-16	2.271	1.105

Collaborative Labs encourage Student Engagement

Table 1 shows that students who performed the lab practices in a collaborative fashion completed on average more assignments than the ones who made it individually (i.e., whereas the mean and the median for 2010-11 are 1.53 and 1 respectively, for 2011-12 are 2.79 and 3). *Student’s t-test* of the number of completed assignments for 2010-11 and 2010-12 has *t* = 3.4684 and *p*-value = 0.0007417. So, the difference between using

our collaborative approach and not using it is statistically highly significant. In addition, the *Cohen's d* is 0.6465427. Therefore, the difference effect size is moderate (>0.5).

Synchronous Collaboration increases Lab Assignment Performance

As Table 2 shows, the correlation factor for course 2011-12 is higher than for 2010-11 ($0.89 > 0.56$), and the slope of the 2011-12 regression line is steeper than the 2010-11 one ($1.28 > 2.69$). So it looks like students get better exam results when practicing with collaborative labs or, in statistical terms, it seems that the collaborative support *moderates* the effect that the number of lab assignments has over the exam grades (Jaccard & Turrisi, 2003). To check such moderation effect, the two multiple regression models summarized in Table 3 has been used. Whereas, Model 1 just includes variables *NumberOfAssignments* and *HasTheCollaborativeFeature* to explain the exam grades, Model 2 includes the moderation effect encoded by the product *NumberOfAssignments* * *HasTheCollaborativeFeature* as well. To facilitate the interpretation of both models:

1. *NumberOfAssignments* is put in *deviation form*, i.e., every value x is centered to the mean: $x_{centered} = x - \text{mean}_{\text{NumberOfAssignments}}$. Thus, the regression coefficient B_1 is 0 when *NumberOfAssignments* is equal to its mean.
2. *HasTheCollaborativeFeature* is encoded as (i) 1 for collaborative assignments, and (ii) 0 for non-collaborative ones.

Table 3: Moderation Effect Evaluation by using Multiple Regression Models.

Moderation effect?		Coefficient values	Coefficient <i>p</i> -values	R ²
No: Model 1 Grade = $B_0 + B_1 * \text{NumberOfAssignments} + B_2 * \text{HasTheCollaborativeFeature}$	B_0	3.90566	$< 2e-16$	0.6209
	B_1	1.09685	$< 2e-16$	
	B_2	1.49757	1.62e-05	
Yes: Model 2 Grade = $B_0 + B_1 * \text{NumberOfAssignments} + B_2 * \text{HasTheCollaborativeFeature} + B_3 * \text{NumberOfAssignments} * \text{HasTheCollaborativeFeature}$	B_0	3.9057	$< 2e-16$	0.6446
	B_1	0.8017	4.87e-08	
	B_2	1.4976	9.78e-06	
	B_3	0.4703	0.00754	

Hence, the interpretation of the regression coefficients for Model 2 is:

- The estimated grade for a student that has completed the average number of lab assignments without using the collaborative feature is $B_0 = 3.9057$.
- The average return per lab assignment completed without using the collaborative feature is $B_1 = 0.8017$.
- The difference in grade between completing the average number of lab assignments using the collaborative feature and not using it is $B_2 = 1.4976$.
- The difference in the grade by completed assignments slope between non-collaborative and collaborative labs is $B_3 = 0.4703$.

The following points support the existence of a statistically significant moderation effect:

1. Comparing both models, the *NumberOfAssignments* coefficient B_1 decreases, i.e., it becomes less important when the interaction *NumberOfAssignments* * *HasTheCollaborativeFeature* is considered. Besides, in Model 2 the moderation effect coefficient B_3 has *p*-value 0.00754, i.e., the interaction term is statistically highly significant.
2. Whereas Model 1 explains 62% of the variance in the exam grades, Model 2 explains 64% of the variance (i.e., R^2 is 0.6209 and 0.6446 for Models 1 and 2, respectively).
3. ANOVA model comparison for both models has $F = 7.4083$ and $\text{Pr}(> F) = 0.00754$, i.e., it is statistically highly significant that Model 2 estimates the exam results better than Model 1.

Discussion and Concluding Remarks

To the extent of our knowledge, previous works on synchronous collaborative support for VRLs are limited to the usage of communication tools such as chat or video-conference applications (Tsovaltzi et al, 2010; Bochicchio & Longo, 2009; van Joolingen et al., 2005). Our approach not only provides that kind of collaboration but also a new way of communication, based on the VRL itself. For each participant in a collaborative session, there is an instance of the shared VRL running. The states of all the VRL instances are synchronized, i.e., whenever a participant acts over its VRL instance, the changes produced on the VRL state are reflected in the other participants' VRL instances. This way, participants have the feeling of working together on the same VRL.

Gravier et al. (Gravier et al., 2008) have surveyed forty-two different remote labs finding that every project implements its own software architecture with no reuse. Both building a VRL from scratch and creating its collaborative support requires a huge effort. Our work alleviates such effort since EJS is a code generator that speeds up the VRL development, and our approach automatizes the addition of collaborative support to existing EJS VRLs. Thus, we avoid "reinventing the wheel" every time a VRL requires collaborative features.

Finally, there is experimental evidence of the usefulness of our work. In particular, the statistical analysis reported in this paper shows (i) a correlation between the student exam grades and the number of completed lab assignments, (ii) an increase in student engagement thanks to the collaborative feature we propose, and (iii) a moderation effect of our synchronous collaboration approach and the number of completed lab assignments. Given the success of this pilot project, we plan to extend the use of our collaborative approach to other UNED courses with a major number of students.

References

- Bafoustou G., & Mentzas, G. (2002). Review and functional classification of collaborative systems. *International Journal on Information Management*, 22, 281-305.
- Bochicchio, M., & Longo, A. (2009). Hands-on remote labs: Collaborative web labs as a case study for it engineering classes. *IEEE Trans. on Learning Tech.*, 2(4), 320-330.
- Boulos, M., Taylor, A. & Breton, A. (2005). A synchronous communication experiment within an online distance learning program: A case study. *Telemedicine journal and e-health*, 11(5). 583-593.
- Chang, G.W., Yeh, Z.M., Chang, H.M., & Pan S.Y. (2005). Teaching photonics lab using remote-control web technologies. *IEEE Transactions on Education*, 48(4), 642 - 651.
- Christian, W., & Esquembre, F. (2007). Modeling physics with easy java simulations. *The Physics Teacher*, 45. 475-480.
- Christian, W., Esquembre, F. & Barbato, L. (2001). Open source physics. *Science*, 334(6059). 1077-1078.
- Dommel, H.P., & Garcia-Luna, J.J. (1997). Floor control for multimedia conferencing and collaboration. *Multimedia Systems*, 5, 23-38.
- Dormido, R., Vargas, H., Duro, N., Sanchez, J., Dormido-Canto, S., Farias, G., Esquembre, F., & Dormido, S. (2008). Development of a web-based control lab for automation technicians: The three-tank system. *IEEE Trans. on Education*, 51(1), 35-44.
- Ellis, K. (2009). Field guide to learning management systems. Technical report, ASTD Learning Circuits.
- Gravier, C., Fayolle, J., Bayard, B. Ates, M. & Lardon, J. (2008) State of the art about remote labs paradigms - Foundations of ongoing mutations. *International Journal of Online Engineering*, 4(1).
- Guimaraes, E., Cardozo, E., Moraes, D., & Coelho, P. (2011). Design and implementation issues for modern remote labs. *IEEE Trans. on Learning Tech.*, 4(2), 149-161.
- Heradio, R., de la Torre, L., Sanchez, J., Dormido, S. & Vargas, H. (2011) An architecture for virtual and remote labs to support distance learning. *Research in Engineering Education Symposium, Madrid, Spain*.
- Jaccard, J.J., Turrisi, R. (2003). Interaction Effects in Multiple Regression. *SAGE Publications, Series: Quantitative Applications in the Social Sciences*.
- van Joolingen, W.R., de Jong, T., Lazonder, A.W., Savelsbergh, E.R., & Manlove, S. (2005). Co-lab: research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior*, 21. 671-688.
- Kozma, R. (1994). Will media influence learning? Reframing the debate. *Education Technology Research and Development*, 42(2), 7-19.
- Ma, J., & Nickerson, J.V. (2006) Hands-on, simulated, and remote labs: A comparative literature review. *ACM Computing Surveys*, 38(3).
- Marjanovic, O. (1999). Learning and teaching in a synchronous collaborative environment. *Journal of Computer Assisted Learning*. 15. 129-138.
- de la Torre, L. Sanchez, J., Dormido, S., Sanchez, J.P., Yuste, M., and Carreras, C. (2001). Two web-based labs of the fisl@bs network: Hooke's and Snell's laws. *European Journal of Physics*, 32(2), 571-584.
- Tsovaltzi, D., Rummel, N., McLaren, B.M., Pinkwart, N., Scheuer, O., Harrer, A., & Braun, I. (2010). Extending a virtual chemistry lab with a collaboration script to promote conceptual learning. *International Journal of Technology Enhanced Learning*, 2. 91-110.
- Wannous, M., & H. Nakano, H. (2010). Nvlab, a networking virtual web-based lab that implements virtualization and virtual network computing tech. *IEEE Trans. on Learning Tech.*, 3(2). 129-138.

Fostering learning and collaboration in a scientific community – evidence from an experiment using RFID devices to measure collaborative processes

Julia Eberle¹, Karsten Stegmann¹, Kris Lund², Alain Barrat³, Michael Sailer¹, Frank Fischer¹

¹Chair of Education and Educational Psychology, University of Munich, Leopoldstr. 13, 80802 München, Germany

²CNRS, ICAR UMR 5191, Université de Lyon, Université Lumière Lyon 2, ENS de Lyon, 69007 Lyon, France

³CNRS, CPT UMR 7332, Aix Marseille Université, 13288 Marseille, France and Data Science Laboratory, ISI Foundation, 10123 Torino, Italy

Email : julia.eberle@psy.lmu.de, karsten.stegmann@psy.lmu.de, kristine.lund@ens-lyon.fr, michael.sailer@psy.lmu.de, alain.barrat@cpt.univ-mrs.fr, frank.fischer@psy.lmu.de

Abstract: In this study, the integration of new members into a scientific community that comprised to a large extent members from the CSCL community was investigated. New members usually lack the necessary knowledge to interact successfully with more experienced members of a scientific community and to find collaboration partners. We investigated how the level of community participation and support for community knowledge were related to the building of new collaborative relationships during a scientific conference. Participants' interaction behavior was tracked using RFID devices; social network questionnaires and a bibliographic analysis provided additional data. We found that newcomers do not interact less with other participants than experienced members, but develop fewer collaborative relationships. The chances that newcomers' interactions lead to the building of new collaborative relationships were increased by access to explicit relevant community knowledge. Making such knowledge explicit seems to be a useful means for supporting newcomers in scientific communities.

Integrating new members into the CSCL community

The CSCL research community has already been the object of research in the past. We have learned that it is a broad interdisciplinary community comprising researchers from more than 11 disciplines who are distributed all over the world but with a majority in North America and Europe and a growing number of members in Asia (Hoadley, 2005; Kienle & Wessner, 2006; Stahl, Spada, Miyake, & Law, 2011). We know that there is an ongoing discussion about what CSCL is, what it comprises, and what a common and shared theory could be; also perspectives on the community differ between members of different geographical locations. Nevertheless, Hoadley (2005) found that the CSCL community consists of a stable core of leading persons. Also a stabilizing trend of CSCL conference participants was observed in 2006 (Kienle & Wessner) which might be seen as a sign for a maturing community.

A healthy scientific community must constantly integrate new members to secure its existence and to bring new ideas in. However, it was observed at previous CSCL conferences that a large percentage of newcomers participated only once and did not participate in later conferences (Kienle & Wessner, 2006). This might be considered an alarming sign and the community might wish to take actions to change this situation and focus more on the integration of new members. So far, research on scientific communities is mostly based on bibliometric analyses focusing on co-authorship or citation analysis of conference proceedings or journal papers (see for example Hoadley, 2005; Newmann, 2004; Lee, Ye, & Recker, 2012). However, papers are artifacts that become only available to the scientific community with a large timely distance to their creation and many factors mediate between the publication of a paper and the beginning of a collaboration between members of the scientific community. To identify those factors that are directly related to the integration of newcomers, such bibliographic analyses are therefore not the optimal choice. For this reason, we conducted a study using a new approach to measure collaboration: RFID devices that all participants of a conference wore and which tracked their interaction with other participants. The aim of this study was to investigate what happens during scientific meetings and how the integration of newcomers can be fostered at such occasions.

Scientific Communities and New Members

The CSCL community can be seen as a scientific community, which is a special form of a community of practice (Kienle & Wessner, 2006). Kienle and Wessner (2005) collected essential characteristics of scientific communities: They consist of a heterogeneous group of members who are usually involved in several scientific communities and are therefore used to switching roles, from expert in one scientific community to less experienced in another scientific community. Members of scientific communities are often geographically

distributed and belong to different organizations; in the CSCL community, members are even located across the whole world, although most members can be found in North America and Western Europe (Hoadley, 2005; Kienle & Wessner, 2006). In many scientific communities, members have backgrounds in different disciplines and scientific cultures, resulting in the use of different methods and theories. What brings them together is a joint field of research interests. Communication and interaction mostly takes place via written artifacts like journal publications and using computer-mediated channels, but there are also regular opportunities for face-to-face meetings, usually organized in the form of conferences and workshops.

Wenger, McDermott, and Snyder (2002) distinguish different levels of community participation in communities of practice located in and across organizations: outsiders, peripheral members, active members, and core members; in addition, some communities of practice have a coordinator. For scientific communities, these participation levels were adapted (Kienle & Wessner, 2005). Outsiders are person who do not intend to contribute to the scientific community, but benefit from its work; for example by reading single papers and maybe exporting ideas and results from them to another scientific community. Peripheral members are, according to the original model, persons who contribute only sporadically to the community and often lack the abilities and knowledge to contribute to more complex tasks. Therefore, newcomers to a community usually start as peripheral members. In a scientific community, peripheral members could be graduate students preparing their first papers but also more experienced researchers who explore a research field that is new to them. While focussing only on observable contributions to conferences, Kienle and Wessner (2005) suggested viewing only passive conference participants as peripheral members. Following that rule, they defined all paper authors as active members. This is, however, only one possible way to conceptualize this participation level as the transition between peripheral and active membership is smooth and graduate students who wrote and presented their paper could still be seen as peripheral members in the first learning stage. Active members usually are defined as those persons who regularly contribute to a community and have the necessary knowledge and skills to do so. Core members are those active members who additionally take over substantial responsibility for the whole community and make efforts to influence its directions. In a scientific community, this can include conference program organizers, journal editors, or scientific board members.

Communication and Collaboration in Scientific Communities

In scientific communities, especially in interdisciplinary ones, successful collaboration can be understood as one of the most desired results of researchers' learning and of the scientific community's cohesion. Successful interdisciplinary collaboration in a scientific community requires an integration of the contributing disciplines on some level, for example the mutual integration of concepts, theories, methodologies, and epistemological principles (van den Besselaar & Heimeriks, 2001). The development of mutual understanding and the building of shared representations are important for fruitful communication between experts of multiple domains (Fischer, 2000).

To enable an individual researcher to benefit from and collaborate in a scientific community, this person must to some extent be integrated and has to acquire several types of knowledge that are shared in the community. Successful collaboration requires shared knowledge, including several different types of relevant knowledge, for example about contents or methods, but also about attitudes in the community as well as about the individual members (Cannon-Bowers & Salas, 2001). However, peripheral members usually have only little knowledge about the new community and need to acquire it first to become able to contribute more and in a proper way (Levine & Moreland, 1999). Compared to mono-disciplinary communities, this might be even more complicated in multi- and interdisciplinary communities because of the variety of research lines. The CSCL community still is divided on several questions, no underlying theory or methodology can be found on which all members would agree as being the basis of CSCL. Attitudes and beliefs of a community are even harder to grasp as they are usually not made explicit. Knowledge about individual community members seems comparably easy to acquire as most CSCL researchers present their bios and publications on their website. But for a newcomer, it is hard to identify the 'important' persons in a community or those who could be relevant for their own work. This makes it hard for newcomers to gain relevant knowledge about content, methods, community attitudes, and members within CSCL on their own.

A usual way to learn about a scientific community is reading papers. But to read only some of them can give a peripheral community member a very biased idea of the community. In this respect, face-to-face meetings are of high relevance to scientific communities; among other things, they provide possibilities for peripheral members to gain community knowledge and interact with other members. Such meetings make communication easier, especially in scientific communities like CSCL which consist of members with different native languages complicating the distribution of results and effective communication (Kienle & Wessner, 2005). Workshops and conferences are used to foster researchers' communication and learning about the findings and approaches of others, but also to integrate newcomers. Such events bring participants together and allow them to focus on learning activities and on community building, and can be called encapsulation. Although encapsulation is a widely used strategy in different contexts (Levine & Moreland, 1991), it can be organized in different ways:

workshops usually allow for more one-on-one interaction, while (larger) conferences usually focus on other types of communication. However, it is unclear how one-on-one interaction is related to researchers' learning in the scientific community. Access to community knowledge, especially to knowledge about other members, seems in particular to be also very relevant and it might be helpful to foster it during encapsulation events.

In this study we investigate factors which influence the integration and learning of newcomers in a scientific community. We adopt a social network perspective on learning and integration by focusing on the building of collaborative relationships between community participants as the visible and desired consequence of integration and learning. The social network approach offers two different ways to look at the building of collaborative relationships. First, we can look at individual persons and how successful they were in building new relationships; second, we can also look at all individual relations between two community members and what factors influence the probability of a random relation to become a collaborative relationship:

1. Persons-related RQ: To what extent can support for community knowledge and participants' level of participation predict a participant's number of newly built collaborative relationships?
2. Relation-related RQ: To what extent can time spent interacting, support for community knowledge, and a persons' level of community participation predict the development of a new collaborative relation?

Method

Study Design, Context and Participants

The study was planned with a quasi-experimental design in which the factor *support for community knowledge* was varied across different workshops and the factor *community participation level* varied naturally among participants. This design was implemented at a small conference organized by a European research community which is closely related to the CSCL community. The aim of the conference was to bring together researchers from multiple disciplines who worked in the field of technology-enhanced learning. The conference was organized in a non-standard way and consisted of 8 workshops and a doctoral consortium. The workshops were organized in two series of 4 workshops taking place in parallel. The number of participants for each workshop varied between 14 and 22 persons. Each workshop lasted one and a half days while the doctoral consortium lasted for the whole 4 days. The conference took place at a hotel in a remote place and all workshop organizers and participants also lived in this hotel during the time. Participants were selected based on a review process of papers they had submitted.

All together, 152 persons participated in the conference. The majority of them came from European countries, but there were also participants from many other countries. For this study, only persons who had participated in one of the workshops were taken into account. Persons who were only involved in the doctoral consortium were left out because of their special conditions during the conference. Also the data of participants who had missing values or whose answer patterns made their credibility questionable was left out. For the analysis of the two research questions, further constraints (explained below) resulted in different sample sizes. For RQ 1, the sample consisted of 89 participants. For RQ 2, the sample consisted of 742 dyadic relations in which 125 persons were involved.

Data Collection and Instruments

Data about the participants' interaction during the conference, their collaborations with other participants beforehand, and their intended collaborations after participation to a workshop were measured using two different tools: RFID devices and social network questionnaires. Further information about the participants was taken from the application form for the conference. Additionally, a bibliographic analysis of co-authorships listed in Google Scholar was performed.

Tracking face-to-face proximity with RFID devices

During the conference, each participant wore an RFID device, developed by the SocioPatterns collaboration (<http://www.sociopatterns.org>) that was integrated into the name badge. The devices engage in bidirectional low-power radio communication. As the human body acts as a shield for the used radio frequency, and as the badges are worn on the chest, badges can exchange radio packets only when the individuals wearing them face each other at close range (about 1 to 1.5 m). The measuring infrastructure can capture that there was a close face-to-face proximity between two individuals with a temporal resolution of 20 seconds, and gives therefore access to the amount of time that two participants spent together (see Cattuto et al. (2010) for a detailed description of the infrastructure). Only two participants of the conference declined to wear the devices.

Social network questionnaires

After each workshop, participants were asked to fill in a social network questionnaire about their relations within the workshop: they were given a list of all workshop participants' names and were asked to indicate with whom they had collaborated already before the conference and with whom they had found potential for future

collaboration. As some conference attendees participated in two workshops, 160 questionnaires were handed out, from which 150 were returned.

Bibliographic analysis on publications listed in Google Scholar

About 1.5 years after the conference had taken place, we performed a Google Scholar search to have an indicator if the subjective indications in the social network questionnaire had lead to an objective measurable collaborative outcome.

To identify if two participants had collaborated with each other before but forgot to indicate that in the social network questionnaire, we performed a search for joint publications before the conference. We restricted this search to papers published in 2010 or earlier (query term: “as_yhi=2010”). For each author’s name, at least two variations were included (query term: “(author: Doe J OR author: Doe John)”). For each possible pair of workshop participants, a separate search was conducted by combining them, for example the query term “(author: Doe J OR author: Doe John) AND (author: Smith S OR author: Smith Samantha)”.

To identify joint papers after the workshop, we performed a second Google Scholar search similar to the first one, but restricted to papers published in 2011 or later (query term: as_ylo=2011).

Dependent and Independent Variables

Level of prior community participation

Each participant was allocated to a community participation category, either as a peripheral community member or as an active community member. The allocation was based on the participant’s previous participation in the scientific community (similar to the allocation criteria used by Kienle and Wessner, 2006). Those participants who fulfilled at least one of the following criteria were assigned as active members: they attended the previous conference 1.5 years ago, they were organizers of one of the workshops at the present conference, or they were members in one of the boards of the scientific community. Peripheral members fulfilled none of these criteria.

Access to community knowledge

In 3 of the 8 workshops, support for community knowledge was given with the aim to foster collaboration between participants. Support for community knowledge was implemented as knowledge about the individual community members. A brochure with the following information about all participants was compiled: their name, picture, contact information, affiliations, background, research interests, and exemplary publications. This brochure was handed out to the participants at the beginning of the workshop without further instructions.

Number of newly built relationships

To answer RQ1, we computed for each person the number of new collaboration partners. Participants with whom the person had already collaborated before the conference were not included as new collaborative relationships; this was either indicated by the person, the partner, or both of them in the questionnaire or by previous collaborative publication between the person and the partner found in the bibliographic analysis. Three different types of newly built relationships were computed based on different measures: the number of new interaction relationships, the number of new subjective collaborative relationships, and the number of new objective relationships.

Number of new interactive relationships. The number of new interaction partners was computed for all persons from their face-to-face time with other workshop participants (recorded by the RFID infrastructure). All participants of the workshop with whom a person had interacted during the workshop (but not collaborated before the workshop) were counted.

Number of new subjective collaborative relationships. This outcome variable was computed from the social network questionnaire. We counted the number of participants with whom the particular person had indicated to have identified potential for future collaboration, but only if the respective participant had also indicated to have identified potential for future collaboration with this person.

The sample for RQ 1, in which the number of newly built relationships was the outcome variable, included only a sub-sample of the conference participants. Only those participants were taken into account who had indicated in the questionnaire to have not collaborated before with at least 10 of the other workshop participants. This constraint was made because of statistical reasons: A person who had only the chance to build a new relationship with 2 other participants of the workshop (because he or she had collaborated with all others before) would bias the results because this person might have built more new relationships if possible.

Interaction time

Interaction time was computed for each dyad of participants in a workshop. The time was taken from the RFID-based measurements. For the relations within one of the 4 workshops in the first part of the conference, only the interaction time from the beginning of the conference to the end of the workshop was taken into account. For the

relations within one of those 4 workshops that took place in the second part of the conference, only the interaction time from the beginning of the workshop until the end of the conference was taken into account. Observed interaction times between pairs of participants ranged from 0 seconds to 75 minutes.

Development of a new collaborative relation

To answer RQ 2, we analyzed all possible new relation between two participants of the same workshop who had not collaborated with each other before. Therefore, relations were excluded if one or both persons in the relation had indicated in the social network questionnaire to have collaborated with each other already before the workshop or if a previous joint publication was identified in the bibliographical analysis. Three different types of possible new relations between two participants were computed based on different measures; all three types were dichotomous variables: the development of an interactive relationship, the development of a new subjective collaborative relationship, and the development of a new objective relationship.

Development of an interactive relationship. If a face-to-face interaction had been recorded with the RFID devices measurement between two persons who formed a possible new relation, this was taken as a newly developed interactive relationship (= 1). If no face-to-face interaction was recorded, the relation was treated as one without a newly developed interactive relationship (= 0).

Development of a new subjective collaborative relationship. If both persons in a relation had indicated in the social network questionnaire to have identified potential for future collaboration after the workshop, this was taken as a newly developed relation (=1). If none or only one of the two persons in the relationship had indicated to have identified potential for future collaboration, this was seen as no relationship (= 0).

Development of a new objective relationship. Taking the results of the Google Scholar search of co-authorships, we looked for each possible pair of two workshop participants, if they had published a joint paper after the conference. If a joint paper was found this was taken as newly developed objective relationship (= 1). If no jointly published paper was found, this was taken as no newly developed objective relationship (= 0).

Data Analysis

To answer RQ 1, the data was analyzed using hierarchical linear modeling and applying a HLM model (using the software HLM 6.08 by Raudenbush, Bryk, and Congdon, 2004) as each person was nested within a workshop. HLM allowed us to control for random effects caused by differences in the workshop which could not be controlled otherwise. As the dependent variable was Poisson-distributed, a logarithmic link function was used.

For RQ 2, instead of looking at the outcome of individual persons, we looked at the *relations* between two persons. In this way, data of the same person appeared several times in the dataset and the relation-data was cross-classified within two persons. To control for person-specific effects we used a HCM2 model (using the software HLM 6.08). Each relation appeared twice in the dataset, so each of both persons in a relation was identified as cross-classification variable twice. After carrying out the analysis, we divided the degrees of freedom in half again to deliver appropriate results for the real sample. Additionally, it was necessary to apply a logarithmic link function to the model to account for the binomial (dichotomous) distribution of the outcome variable. Additionally, a χ^2 -test was applied to investigate the development of new objective collaborative relationships because the percentage of identified new objective collaborative relationships was too small to apply HCM here as well. For all analyses, the significance level was set to .05.

Results

Person-related RQ1: Influences on a person's number of new interactive relationships. No significant effects of community participation level, support for community knowledge or of an interaction of those variables on the participants' number of interaction partners during the workshops could be identified.

Person-related RQ1: Influences on a person's number of new subjective collaborative relationships. The population-average model to predict a participant's number of new collaboration partners after the workshop revealed the following results: An average participant who was an active member of the scientific community and had not received support for community knowledge acquired on average 3.4 new collaboration partners ($\gamma = 1.222$, $SE = 0.142$, $t(7) = 8.622$, $p < .01$). Peripheral community members, in contrast, acquired on average only 2.0 new collaboration partners ($\gamma = -0.319$, $SE = 0.153$, $t(86) = -2.087$, $p = .04$). However, peripheral community members who received support for community knowledge were able to find on average 3.6 new collaboration partners ($\gamma = 0.437$, $SE = 0.213$, $t(86) = 2.051$, $p = .04$). No significant effect of support for community knowledge was found on active community member's number of new subjective collaborative relationships.

Relation-related RQ 2: Influences on the development of a new interactive relationship. The HCM model showed that for two random participants of a workshop who did not receive support for community knowledge, the average probability to develop a new interactive relationship was 30.3 % ($\theta = -0.832$, $SE = 0.150$, $t(698) = -5.541$, $p < .01$). This probability varied significantly across individual workshop participants. If they

received support for community knowledge in their workshop, their probability to develop a new interactive relationship was significantly reduced by 10.3 % ($\theta = -0.551$, $SE = 0.235$, $t(698) = -2.340$, $p < .01$). No significant effect of community participation level was found on the probability to develop a new interactive relationship.

Relation-related RQ 2: Influences on the development of a new subjective collaborative relationship.

We identified the following HCM model to predict the probability of a random relation between two workshop participants to develop a new subjective collaborative relationship after they had participated in the same workshop: If both persons were active community members and they had not interacted face-to-face with each other, their probability to mutually identify potential for future collaboration was on average 27.9 % ($\theta = -0.948$, $SE = 0.162$, $t(697.5) = -5.861$, $p < .01$). However, this varied significantly across individual persons. For every minute two persons spent interacting, the probability for them to develop a new subjective collaborative relationship increased significantly (θ (in seconds) = 0.001, $SE < 0.001$, $t(697.5) = 2.057$, $p = .03$). For example, a relation in which the two persons had spent 1 minute interacting with each other, had an increased probability to develop into a new subjective collaborative relationship by 1.2 % compared to a relation in which no direct interaction was measured. If one of the two persons was a peripheral member, the probability for the relation to develop into a new subjective collaborative relationship was significantly reduced by 6.7 % ($\theta = -0.361$, $SE = 0.197$, $t(697.5) = -1.836$, $p < .04$). No significant effect was found for support for community knowledge.

Relation-related RQ 2: Influences on the development of a new objective collaborative relationship.

There was a significant difference between the relations of participants who had received support for community knowledge and those relations in which no support was available regarding the development of new objective collaborative relationships ($\chi^2(1,699) = 21.11$, $p < .01$) favoring those who had received the means of support. No significant differences was identified between peripheral and active members. Also no significant difference was found between relations in which persons had directly interacted with each other and those without direct interaction during the workshop. No difference was found as well between those relations in which a subjective collaborative relationship was reported and those in which no such relationship was reported.

Conclusion

Encapsulation events like conferences and workshops are an important means of scientific communities to bring their members together and foster integration of new members. Therefore, peripheral members of scientific communities are an important group of participants. Although peripheral community members seem not to differ from active community participants regarding the number of interactions they have during a conference or the probability for interacting with a random other participant, they seem disadvantaged regarding the outcome of these interactions. Peripheral members' chances to develop subjective collaborative relationships are lower resulting in a lower number of new subjective relationships after the conference. However, we do not find these results in the more objective measurement on real collaborative outcomes, but this might be due to the fact that measuring joint papers 1.5 years after the conference is still a bit early. The process to plan a joint research project, collect and analyze data, write a paper and successfully publish it takes usually a long time and it seems advisable to rerun a Google Scholar search at a later point in time.

In this study we assumed that the disadvantages of peripheral community members could be based on their lack of community knowledge. Therefore, we supported the participants of some of the workshops with explicit community knowledge. Providing participants with this support reduced the chances that a participant interacted with a random person, but did not reduce the number of their interaction partners during the conference. This can be seen as a hint that participants who had the community knowledge support were able to identify more precisely who would be a relevant interaction partner and enabled participants to use their few time more efficiently by focusing on interactions with those participants. This is in line with the result that the longer two persons spent interacting with each other, the higher their chances to build a new collaborative subjective relationship. However, our community knowledge support was not directly related to the chances for building a new subjective collaborative relationship or the number of newly built subjective collaborative relationships. Looking at the few results of joint papers we identified so far, we see a clear relation between receiving community knowledge support and successful objective collaborative relationships. Taking all these results together, we can assume that community knowledge support enables participants of scientific meetings, especially those who still possess little community knowledge, to identify promising partners for collaboration more efficiently and to focus on longer and more effective interactions with those, which increases the chances of plans for future collaboration to become real and visible collaborations.

Additionally to the limitations of the objective collaborative relations measurement through our Google Scholar search, also some technical problems with the RFID devices have to be reported: we can not claim to really have measured all face-to-face interactions between workshop participants because some participants lost or forgot their RFID devices for some time or the devices run out of battery. Also, the name badges in which the devices were integrated flipped quite often, so the body of the participant wearing the badge shielded the radio signals. Although we surely missed some interactions between participants, the results show that RFID devices

can work as a promising new method to measure collaboration. But researchers who want to use this technique in the future can surely improve its use by taking care of the reported problems.

Aside from technical questions, the results of this study confirm the importance of shared community knowledge for collaboration (Cannon-Bowers & Salas, 2001). The relation between community knowledge has so far been studied on the cognitive level, but our findings confirm them also from a social network perspective with visible outcomes. Acquiring shared community knowledge seems important for the building of new collaborations and the integration of peripheral members into a scientific community. Additionally, our results confirm finding from previous studies about tactics which can be used by communities of practice to integrate and support their newcomers and peripheral members (Eberle, Stegmann, & Fischer, 2012). This study had identified explicit access to community knowledge as an important means to foster the learning of peripheral members and foster their collaboration with more active members. This finding seems to be transferable to scientific communities as a special type of community of practice.

We can also draw some practical implications from this study for future CSSL conference and workshop organizers as well as for their participants: Organizers can support the integration of peripheral participants by providing explicit access to information about the community and its members and by planning for enough time for their participants to interact on a one-on-one basis with each other. Participants, on the other hand, can positively influence their workshop and conference experiences by informing themselves beforehand about the other participants and their backgrounds and by focusing on longer one-on-one interactions with other participants.

References

- van den Besselaar, P. & Heimeriks, G. (2001). Disciplinary, Multidisciplinary, Interdisciplinary: Concepts and Indicators. In M. Davis and C.S. Wilson (Eds.), *ISSI 2001* (pp. 705-716), Sydney: UNSW 2001.
- Cannon-Bowers, J. A. & Salas, E. (2001). Reflections on Shared Cognition. *Journal of Organizational Behavior*, 22, 195-202.
- Cattuto, C., Van den Broek, W., Barrat, A., Colizza, V., Pinton, J.-F., Vespignani, A. (2010). Dynamics of person to person interactions from distributed RFID sensor networks. *PLoS ONE*, 5(7), e11596.
- Eberle, J., Stegmann, K., & Fischer, F. (2012). Legitimate peripheral participation in academic communities of practice – How newcomers' learning is supported in student councils. 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. 386-390). Sydney, NSW, Australia: ISLS.
- Fischer, G. (2000). Social Creativity, Symmetry of Ignorance and Meta-Design. *Knowledge-Based Systems [Special Issue on Creativity & Cognition]*, 13(7-8), 527-537.
- Hoadley, C. (2005). *The shape of the elephant: Scope and membership of the CSSL community*. Paper presented at CSSL 2005, Taipei, Taiwan.
- Kienle, A., & Wessner, M. (2005). Principles for cultivating scientific communities of practice. In P. v.d. Besselaar, G.D. Michelis, J. Preece, & C. Simone (Eds.), *Communities and Technologies 2005, Proceedings of the 2nd International Conference on Communities and Technologies*, Berlin: Springer.
- Kienle, A., & Wessner, M. (2006). The CSSL community in its first decade: Development, continuity, connectivity. *International Journal of Computer-Supported Collaborative Learning*, 1(1), 9-33.
- Levine, J. M., & Moreland, R. L. (1991). Culture and Socialization in Work Groups. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (2nd ed., pp. 257-279). Washington, DC: American Psychological Ass.
- Levine, J. M. & Moreland, R. L. (1999). Knowledge Transmission in Work Groups: Helping Newcomers to Succeed. In L. L. Thompson, J. M. Levine & D. M. Messick (Eds.), *Shared Cognition in Organizations: The Management of Knowledge* (pp. 267-296). Mahwah/London: Erlbaum.
- Lee, V. R., Ye, L., & Recker, M. (2012). What a long strange trip it's been: A comparison of authors, abstracts, and references in the 1991 and 2010 ICLS Proceedings. 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. 172-176). Sydney, NSW, Australia: ISLS.
- Newman, M.E.J. (2004). Co-authorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences (PNAS)*, 101 (Suppl. 1), 5200-5205.
- Raudenbush, S.W., Bryk, A.S, & Congdon, R. (2004). HLM 6 for Windows [Computer software]. Skokie, IL: Scientific Software International, Inc.
- Stahl, G., Spada, H., Miyake, N. & Law, N. (2011). Introduction to the Proceedings of CSSL 2011. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSSL2011 Conference Proceedings. Volume I – Long Papers*. ISLS.
- Wenger, E., McDermott, R., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Boston, Mass.: Harvard Business School Press.

Supporting Active Wiki-based Collaboration

Adam Eck, Leen-Kiat Soh, and Chad Brassil, University of Nebraska-Lincoln, Lincoln, NE
Email: aeck@cse.unl.edu, lksoh@cse.unl.edu, cbrassil@unl.edu

Abstract: Prior research has established that active participation and collaboration by students results in multiple benefits during wiki-based CSCL activities. However, achieving such behavior can be a challenge without external motivation. To increase active participation and collaboration by users, we developed an enhanced wiki called the Written Agora. Using popular Web 2.0 features, our wiki provides additional means of participation and aims to encourage direct communication and collaboration between users. Additionally, using intelligent features, we enable the wiki system itself to also participate during collaboration. In this paper, we analyze the results of a study using the Written Agora in a classroom for two semesters. We discover that simply including such features is not necessarily enough to cause their use or improve collaboration. However, encouraging the use of these features resulted in not only greater than expected use, but more diverse and higher quality collaboration by users.

Introduction

Within recent years, one popular tool for computer supported, collaborative learning (CSCL) has been wiki software. Wikis have been used for a wide-variety of CSCL activities, including hosting supplementary material to classroom lectures (Cole, 2009), building glossaries of important terms (Peterson, 2009), group essay writing (Khandaker & Soh, 2009), and contributing to a publically shared knowledge-base (Lampe et. al, 2012).

The use of wikis for CSCL has resulted in several positive, documented results. For example, Wheeler et. al (2008) found that students wanted to create high quality content given the possibility and excitement of broad information dissemination, as well as a self-reported increase in student writing and critical thinking skills. Cress and Kimmerle (2008) developed a model of student interactions with a wiki, focusing on assimilation (addition of new knowledge) and accommodation (reconstruction of existing knowledge) both *internally* within students and *externally* in the wiki between a group of collaborating users. Based on this model, Moskaliuk, Kimmerle, and Cress (2008; 2012) verified that features of wiki content (e.g., levels of incongruity, redundancy and polarity with student prior knowledge) can encourage student learning. Lampe et. al (2012) observed that some students were motivated to continue contributing to wiki systems such as Wikipedia (<http://www.wikipedia.org>) after producing content as part of classroom activities. Thus, using a wiki for CSCL can benefit students both during classroom activities, as well as beyond the classroom.

However, one common concern has been revealed from the use of wikis in CSCL. Specifically, several studies (e.g., Ebner et. al, 2008; Cole, 2009) report that students tend not to participate in wiki activities without proper external motivation (e.g., requiring participation by assigning points towards grades). This lack of participation is troublesome because without contributing during wiki-based activities, students will fail to achieve the aforementioned benefits of using a wiki for CSCL. Furthermore, research outside of CSCL has also documented the benefits of and concerns over active participation during wiki activities. For example, users who are more active while first exploring the wiki system are much more likely to continue participating in the community in the future (e.g., Panciera et. al, 2009; Antin et. al, 2012). Thus, initial buy-in is very important. Furthermore, Kittur and Kraut (2008) and Arazy and Nov (2010) found that articles on Wikipedia that contained more active collaboration by users (e.g., high levels of activity on the corresponding “discussion” page where users can leave comments for other users) achieved a higher subjective quality (e.g., content accuracy and completeness). Therefore, active participation by users is paramount to achieving the benefits of wiki-based activities both for individual users (e.g., student learning) and the system (e.g., better quality content).

In order to improve the use of wikis as a tool for CSCL, we propose an advanced, intelligent wiki system called the **Written Agora**. Within the Written Agora, we augment the traditional wiki framework with additional features designed to offer more modes of collaboration to encourage greater participation by users. To achieve this goal, we leverage popular features common to other collaborative Web 2.0 applications (e.g., Amazon, <http://www.amazon.com>; Facebook, <http://www.facebook.com>; Reddit, <http://www.reddit.com>) with which users are likely already familiar, such as page ratings, keyword tagging, and threaded discussions. These additional features enable users to participate in different ways than they would in a traditional wiki (e.g., just by viewing and editing pages), hopefully endearing or empowering users and subsequently increasing participation. That is, users who might not have been comfortable or well equipped to participate in traditional wiki-based activities now have additional ways to contribute to the collaboration process, such as providing feedback or summarizing the key content of pages. These software features also provide additional means for external assimilation and accommodation (Cress & Kimmerle, 2008) through adding additional information to wiki pages (e.g., ratings) and coordinating transformations of knowledge (e.g., threaded discussions), which could boost student learning similar to features in wiki content (Moskaliuk, Kimmerle & Cress, 2008; 2012).

Furthermore, we also add *intelligent* features to enable the wiki itself to be a *proactive participant* in collaboration while supporting users with their tasks. For instance, as a user browses pages in the Written Agora, we provide automated recommendations of similar pages the user might be interested in viewing based on the content of those pages. This improves the end-user experience by both helping the user navigate through the broad expanse of topics present in a collaborative wiki to target topics of interest in greater depth, as well as improving user knowledge by encouraging increased usage of the wiki. Moreover, it potentially decreases the burden of initial system usage, which could result in more active and sustained participation by users. Similar intelligent features have been demonstrated to be beneficial to wiki participation both (1) within a CSCL setting, such as intelligently forming groups which increases participation (Khandaker & Soh, 2009), and (2) outside of CSCL, such as recommending pages to edit or content to include in order to increase page quality and coverage (Cosley et. al, 2007; Kong et. al, 2010).

In this paper, we evaluate an experimental study conducted to investigate the impact of including additional Web 2.0 and intelligent features within the Written Agora to increase active participation by students. We consider the results of deploying our wiki system over the course of two semesters in an undergraduate classroom setting. Consistent with prior results of general wiki usage (e.g., Ebner et. al, 2008; Cole, 2009), we find that simply including such features is not necessarily enough to result in their use or improve collaboration. However, encouraging the use of these features resulted in greater than expected use and *more diverse* and *higher quality* collaboration by users. Therefore, their inclusion does result in a net benefit for students, but does not necessarily address the active participation concern. Based on our results, we hypothesize several possible avenues to tackle this important problem without requiring external motivation (e.g., graded participation).

The Written Agora

The features of the Written Agora can be categorized into three primary categories: (1) **traditional** features commonly found in other existing wiki systems, (2) additional **Web 2.0** features, used to offer further modes of participation and to enhance communication between students, and (3) **intelligent** features, enabling the system itself to actively participate in collaboration and support user activities. Table 1 summarizes these features. In the following, we elaborate on the additional Web 2.0 and intelligent features studied in this paper.

Table 1: Features of the Written Agora

Category	Purpose	Features
Traditional	Provide support for common wiki-based activities.	<ul style="list-style-type: none"> • Create, view, edit, and delete pages with rich text • View and compare page revision history • Control viewing and editing access to pages • Add, view, and delete multimedia attachments • Browse and search for pages
Web 2.0	Enable more modes of participation and advanced user collaboration behavior.	<ul style="list-style-type: none"> • Rate pages • Tag pages using keywords • Converse in threaded discussions
Intelligent	Enable the system to proactively participate during collaboration and support users' activities.	<ul style="list-style-type: none"> • Track user behavior during activities • Extract important keywords from pages • Recommend pages based on keyword similarity

Web 2.0 Features

Beyond traditional wikis, the Written Agora includes additional Web 2.0 features common to many other types of contemporary collaboration systems on the web (e.g., social networking sites). These features provide additional means of participation for users, and also enable advanced collaboration and enhanced communication between users to promote improved collaborative work and higher quality pages. Additionally, the features enable the wiki system to store not only the *end product* of collaboration (e.g., shared knowledge), but also *by-products* (e.g., ratings, consensus) and serve as a self-contained *process* for collaboration without the need for coupling with external tools (e.g., email, instant messaging), which could be inconvenient for users.

Page Ratings: Users can rate pages based on their quality using a scale from 1 (bad) to 5 (great) stars. Ratings from multiple users are scored using a cumulative average to define an overall user-supplied quality metric on pages. The overall and current user's ratings are displayed on each page, as well as when browsing for pages, to assist users in quickly evaluating whether a page might be worth reading to increase their own knowledge, or whether or not a page is a good candidate for their editing to improve the overall knowledge within the system. Prior work has shown that visualizing information about pages can improve trust in the information stored in wiki systems (Kittur et. al, 2008), which we believe might also be achieved through community ratings.

Keyword Tagging: Users can also tag pages with important keywords, useful for both (1) summarizing the important concepts within a page, as well as (2) organizing pages around similar topics. Moreover, tag clouds

displaying the most popular keywords based on their frequency of use allows users to observe a current snapshot of the current knowledge within the wiki system. Keywords also assist with page navigation towards topics of interest—clicking on a keyword either on a page or in the tag cloud searches for all pages either tagged with that word or containing the word in its content.

Threaded Discussions: Each page supports collaborative discussions between users through threaded comments, allowing users to start new topics of discussion and respond to one another's comments. For example, users might organize their edits for the page, mediate conflicts, or propose new ideas and suggestions to improve the quality of the page. They can also ask questions of one another and receive direct responses to promote enhanced understanding of the page's content. These discussions provide organized, topic-based communication, in contrast with a more free-form and less organized style of discussion, as in discussion pages on Wikipedia.

Intelligent Features

Another novel aspect of the Written Agora is the inclusion of intelligent features, common to other applications with intelligent user interfaces. These features enable the system to provide its own active support to individual and group user activities. For example, the system can assist users when searching for desired information by leveraging its own knowledge of the system's contents, and it can organize its content and link related pages based on shared topics. Our intelligent features are powered by a multiagent system adhering to the Adaptive Knowledge Assistants framework (Eck & Soh, 2012), where individual and system agents are used to provide tailored support to user activities within the Written Agora.

Keyword Extraction: Similar to keyword tagging by users, the Written Agora also automatically analyzes every page and extracts the most important keywords, which assists users in (1) finding relevant topics within pages, as well as (2) organizing pages containing similar concepts and ideas.

Page Recommendations: Moreover, using the keyword content found by automated extraction and tagged by users, the system also recommends similar pages to the one the user is currently viewing. These recommendations are presented in an unobtrusive side panel to help the user navigate through the wide expanse of the system's content without distracting the user from her current activities. Such recommendations are useful for assisting the user improve her knowledge through both the breadth of related concepts to the current page in the wiki, as well as the depth of understanding of the current topic by targeting pages describing the topic in more detail. Of note, this feature differs from prior recommendations in wikis (e.g., Cosley et. al, 2007; Kong et. al, 2010) in that our recommendations are intended to grow each *individual's* internal knowledge through exploring existing pages, rather than intending to grow the *community's* external knowledge through expanding the shared information in the system. In the future, we plan to explore both types of recommendations.

User Tracking: The Written Agora monitors and records all user activity within the wiki, including which page revisions are viewed by users, which keywords are tagged or removed from pages, and what recommended pages are viewed. For each activity, the system tracks: (1) who performed the activity, (2) what activity was performed, (3) when the activity occurred, (4) what page the user was viewing, and (5) any object corresponding to the activity (e.g., rating, comment). Using this tracked information allows us to evaluate the collaborative behavior of users and provides information to know how best to support users during their collaboration.

Study and Methods

As described previously, the primary purpose of the design of the Written Agora was to create an advanced wiki system that (1) encourages opportunities for participation between users, and (2) actively participates with users in the collaboration process. We conducted a user-based study evaluating the impact of support provided by the Written Agora on user activity and collaboration through both additional Web 2.0 and intelligent features. In the following, we outline (1) the research questions guiding our study, (2) our proposed hypotheses answering these questions, (3) the data set used for our analysis, and (4) the methods used to evaluate our hypotheses.

Research Questions

Guiding our research in CSCL through the use of the Written Agora are two primary research questions, each corresponding to different types of support for active participation and collaboration:

Q1: How does the inclusion of Web 2.0 features intended to encourage more opportunities for participation, as well as active communication and collaboration between users, affect the activities and performance of users?

Q2: How does the inclusion of intelligent features intended to enable the system to become an active participant during collaboration through interactions with users affect the activities and performance of users?

With respect to Q1, we aim in this study to assess the impact of including (1) page ratings, (2) keyword tagging, and (3) threaded discussions in the wiki system in order to potentially increase active participation and

collaboration amongst users. With respect to Q2, we aim to assess the impact of including (1) automated keyword parsing, and (2) related page recommendations on the overall collaboration activities of users.

Hypotheses

Based on these two research questions, we propose several hypotheses stating our expectations about the impact of Web 2.0 and intelligent features on user participation and collaboration:

H1: The inclusion of additional Web 2.0 features will increase the amount of user activity and collaboration.

H2: The inclusion of additional Web 2.0 features will provide more opportunities for participation and collaboration, spreading out activity from only a few users (e.g., Panciera et. al, 2009; Antin et. al, 2012) to most users.

H3: The inclusion of intelligent features will result in more page views (through searches for related pages with similar extracted keywords and followed recommendations).

H4: The inclusion of Web 2.0 and intelligent features will boost the quality of collaboration through more active collaboration amongst users and with the system.

While these hypotheses are intuitive responses to the research questions posed above, they are not guaranteed to hold true in practice. For example, we might observe that collaboration fundamentally follows a power law distribution (Antin et. al, 2012) where only a few users perform nearly all activities, regardless of the type of activity (e.g., editing, rating, commenting) while other users perform few if any activities. If so, the additional features might not result in any increase in total participation as the few active users are already near a maximal amount of activity without these features. Furthermore, the inclusion of intelligent features could have no impact on user behavior as users might not trust or simply ignore the system's active participation.

Data Sets

For this study, we consider two semesters of deployment (Fall 2011 and Spring 2012) of the Written Agora branded as the Duckweed Paper Exchange (DPE, <http://duckweed.unl.edu>), a component of the Duckweed Project (<http://www.unl.edu/cbrasil/duckweed-project-0>) within the School of Biological Sciences at the University of Nebraska-Lincoln. In this project, students conduct group-based lab experiments studying the growth of *Spirodela polyrhiza* in different treatment conditions. Based on these experiments, students write their own individual reports about their group's activities, and then the entire group collaboratively forms a final group report detailing their experiment. Within the Duckweed Project, the DPE serves as a tool and repository for the creation and archival of these group reports. Moreover, the reports within the DPE constitute a student-produced journal, where students consider past reports in the design of their own experiments and reference prior work by other students in their own reports. Using the DPE, students practice scientific writing.

Prior to our study, the DPE was prepopulated with one previous semester's group reports (written without the DPE) as an initial seeding of content, as well as one semester's initial usage of the DPE as a pilot study to evaluate the feasibility of the DPE. Thus, between two and three semesters worth of prior reports were available during our study for students to view, rate, and discuss in order to guide their experiments and writing. In both semesters, students were only required to create their final group report using the DPE. Alternatively, they were also allowed to create their own individual drafts within the system before forming a group report, although this step could also be performed outside of the DPE. In the Fall 2011 semester, Web 2.0 features were offered for use but students were not required to use these features. Later, in the Spring 2012 semester, students were required to perform a minimum level of collaborative activities to encourage further collaboration amongst students. These requirements included rating 3 pages, tagging 2 keywords, and offering 5 comments. Moreover, for this later semester, we added the intelligent features considered in our study (automated keyword parsing and related page recommendations). Overall, 47 and 41 users from the Fall 2011 and Spring 2012 semesters agreed to participate in our study, respectively. Of these users, 36 and 28, respectively, agreed also to allow us to consider their earned grades in our study. Thus, our data sets consist of all of the activities performed by these 88 users, including the use of the Web 2.0 and intelligent features in the DPE, as well as the grades earned by 64 users. We would like to note that the grading of student reports was performed by impartial graduate teaching assistants assigned to the course who were not part of our research project. They were made aware that a study was ongoing, but were not given information about what the study measured or our analysis approach. Thus, there was no bias in the grading to impact our results.

Evaluation Methods

To evaluate our hypotheses, we propose the following methods. First, we consider the level of participation and collaboration by users *cumulatively* in each semester, measured by counting the number of times users performed each type of action: ratings, keywords added, comments, recommendations followed, edits, and views. These values are compared against one another both (1) *within* each semester to assess how users collaborated as a collective whole and how the use of one type of activity affected the other types of activities performed by

users, and (2) *across* semesters to determine how the requirement of using Web 2.0 features affected user collaboration behaviors. Second, we also look at the relationships between feature usage (measured using correlations) to learn more about how individual users who exploited or ignored the Web 2.0 and intelligent features behaved in general. Our goal is to better understand the relationship between these activities and user behavior, including whether or not users adopted specific collaborative roles through the use of these features. Finally, we also evaluate the *quality* of collaboration by comparing the grades received for the groups' reports. Here, we aim to understand what relationship exists between the level and type of participation and collaboration performed by users and the quality of the end product of collaboration.

Results

Use of Additional Web 2.0 Features and their Effect on Collaboration

First, we analyze the use of the additional Web 2.0 features during the users' wiki-based activities while writing their reports in wiki pages. To evaluate how often these features were exploited during collaboration, we present the number of actions performed per user, ranked in decreasing order, in terms of (a) ratings made, (b) keywords tagged, and (c) comments posted from both the Fall 2011 and Spring 2012 semesters in Figs. 1a-c.

From these figures, we observe that in the Fall 2011 semester when usage was not required, very few students used these features during their wiki-based activities. In fact, of the 47 users participating in our study, only 3, 7, and 9 users rated pages, tagged keywords, or posted comments, respectively. Moreover, the few users who did exploit these features used them very infrequently. Thus, simply including additional types and modes of collaboration did not necessarily increase the level of participation by students. This matches the results from previous studies of wiki-based activities for CSCL where few users participated without external motivations (e.g., graded requirements) (Ebner et. al, 2008; Cole, 2009). However, we observe a dramatic change of behavior in the Spring 2012 semester *when minimum levels of activity were required*. Here, a larger percentage of users not only used these features to meet the minimum requirements (represented by horizontal black lines in Figs. 1a-c), but most users went *above and beyond* what was necessary. This implies that once users were encouraged to use the features, they perceived a greater value in their use through experience and made greater use of these features than necessary. We hypothesize that *the use of Web 2.0 features in wiki-based collaboration requires reaching a (albeit small) "critical mass" where enough users make use of the features for their benefits to be perceived and their use sustained by the community*. In the future, we plan to further investigate how to achieve and sustain such a crowd-based effect without requiring external motivations.

Second, we analyze the effect of the use of Web 2.0 features encouraging more active collaboration amongst users on both the (1) behavior and (2) quality of collaboration. We begin by presenting the number of pages edited per user in decreasing order in Fig. 1e. We observe that the use of Web 2.0 features appear to have had a significant effect on the editing behavior of users. On the one hand, in the Fall 2011 semester when users made little use of the Web 2.0 features, users generally performed similar numbers of edits. We believe this was due to users focusing on only one type of collaborative action (editing), and each user tried to contribute equally to the group project, so users each had to perform similar numbers of edits. On the other hand, in the Spring 2012 semester when users exploited the Web 2.0 features, we observe that editing behavior shifted to where only a few users made the majority of the edits to the groups' reports, whereas other users contributed instead through the additional Web 2.0 features. Thus, it appears that the use of Web 2.0 features caused role diversification within the groups, where users contributed in different ways. For example, some users tagged keywords to summarize the report and organize it within the context of the other pages in the wiki, some users actively discussed page content through comments, and others carried out designated edits. Moreover, other users contributed more to the class in general by rating many pages within the wiki, rather than contributing to their group's page. Therefore, adding additional Web 2.0 features achieved our goal of increasing opportunities for participation by different users, which led to greater overall participation by users.

Furthermore, the increased diversification of collaboration in the Spring 2012 semester also resulted in higher quality collaboration than the less diverse Fall 2011 semester. Table 2 shows that the grades earned for the group reports were much higher in Spring 2012 than in Fall 2011. Additionally, the standard deviation in student grades was also much smaller in Spring 2012, indicating that the quality of reports was *consistently* better in Spring 2012. Therefore, *the increased use of the Web 2.0 features resulted in not only more active collaborative activity amongst users, but also higher quality collaboration, thus benefiting the wiki system and users*.

Use and Effect of Intelligent Features

Next, we analyze the use of intelligent features by users only during the Spring 2012 semester (since these features were not available during Fall 2011). These features encourage the user to explore the other pages within the wiki by either directly recommending such pages or indirectly helping the user find related pages by organizing pages with similar keywords. To evaluate the impact of these effects, we present per user the number of such recommendations followed and the number of pages viewed in Figs. 1d and 1f.

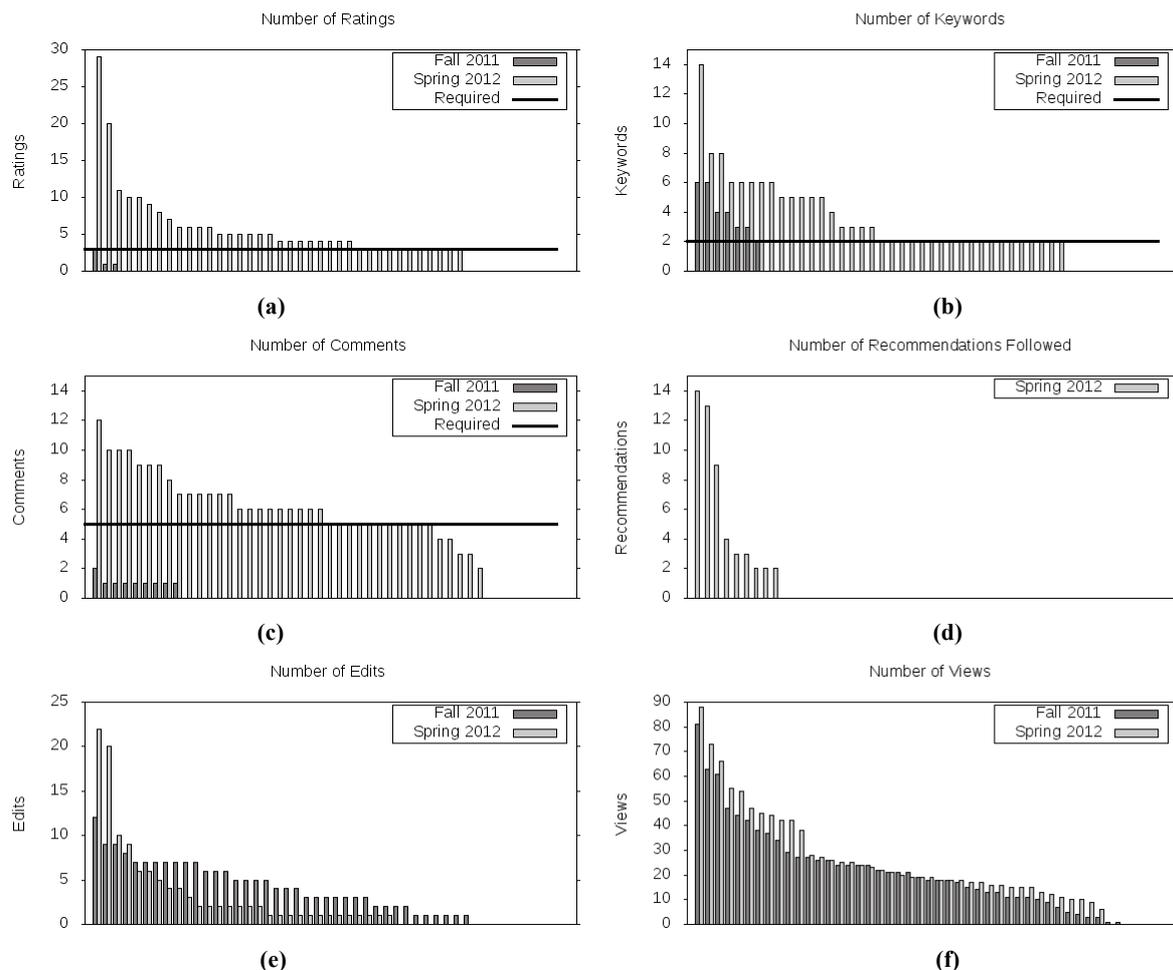


Figure 1. Ranked Ordering of Activity Counts for (a) Ratings Made (b) Keywords Tagged (c) Comments Posted (d) Related Page Recommendations Followed (e) Pages Edited (f) Pages Viewed
 Note: the horizontal black lines in (a)-(c) represent the number of activities required for Spring 2012

Table 2: Grades Earned for Group Reports

	Fall 2011	Spring 2012
Mean	87.6944%	96%
Max	100%	100%
Min	72%	92%
SD	7.7527%	2.7756%

First, usage of this intelligent feature did *not* follow quite the same trend as the Web 2.0 features in Fall 2011. Specifically, despite the fact that in both cases the use of these features was not required, we observe that a *larger number of users* exploited the recommendations made by the system in Spring 2012 than used the Web 2.0 features in Fall 2011, and users did so *more frequently* as well. This result indicates that impactful, sustained use might be easier to achieve for intelligent rather than Web 2.0 features. Furthermore, such use could be enhanced through increased awareness of these features. Particularly, recommendations were made in an unobtrusive side panel located near the bottom of a page; with better visibility, this feature could become more useful to more users, similar to the effect we observed for required Web 2.0 feature usage.

Additionally, we observe that the inclusion of intelligent features positively affected the total viewing behavior of users. Most importantly, we note that the least active users (in the tails of the viewing distributions) viewed a higher number of page views in the Spring 2012 semester that included the intelligent features. This implies that including intelligent features encouraged users to more actively participate in the wiki-based activities by viewing more pages. This result could be due to the system making it easier for users to explore the collaborative knowledge stored within the wiki system, thereby lowering the costs of entry by the least active users.

Relationships Between Features

Finally, we analyze the relationships between the use of the different types of features and with the quality of collaboration. We consider the correlations between activity counts from each feature type and the correlation

between activity counts and grades. These results are presented in Tables 3 (Fall 2011) and 4 (Spring 2012). We highlight the results found to be statistically significant at the 0.01 and 0.05 significance levels.

In Table 3 describing the Fall 2011 semester where few students made use of the Web 2.0 features, we observe that a significant, positive correlation existed between the keyword tagging and page view and edit actions. Thus, the few students who used the advanced keyword tagging feature were active users within the wiki, indicating that including this feature didn't boost participation by inactive users, but was instead an additional way for active participants to collaborate. On the other hand, students who rated pages were also significantly likely to leave comments, indicating that a *second type of participants emerged*: students who offered feedback on pages, but didn't contribute additional content to wiki pages. Unexpectedly, the students' grades were not significantly correlated to any of the particular actions, indicating that even students who performed many actions (especially viewing and editing) did not necessarily achieve high grades for their reports, so the quantity and quality of wiki-based activity were *unrelated*. All other activities were not significantly correlated.

In contrast, in Table 4 we observe additional significant, positive correlations for the Spring 2012 semester. First, usage of all of the Web 2.0 features was highly correlated. Thus, using some features might have helped influence the use of others, which could be beneficial with assisting the system to reach "critical mass" of their usage and boost overall participation. Second, usage of the Web 2.0 and intelligent features was also highly correlated. That is, *users were more likely to use any of the advanced features once they used one of them*. Finally, we also observe several significant, positive correlations between user actions and their grades. This indicates that, unlike in the Fall 2011 semester, users who performed larger *quantities* of collaboration were also likely to achieve higher *quality* collaboration. Therefore, we have more evidence that increased participation through the inclusion of Web 2.0 and intelligent features led to better collaboration between users.

Table 3: Correlations between Activity Counts and Grades for Fall 2011 Note: ** = p < 0.01, * = p < 0.05

	View	Edit	Comment	Keyword	Rating
Edit	0.7250**				
Comment	0.1745	0.1144			
Keyword	0.3474*	0.3413*	0.2753		
Rating	0.0964	0.2505	0.2888*	0.0893	
Grade	0.0964	0.0141	-0.1411	0.3268	-0.3224

Table 4: Correlations between Activity Counts and Grades for Spring 2012 Note: ** = p < 0.01, * = p < 0.05

	View	Edit	Comment	Keyword	Rating	Rec.
Edit	0.7531**					
Comment	0.5514**	0.3870*				
Keyword	0.6061**	0.4764**	0.5726**			
Rating	0.1798	-0.0344	0.3829*	0.3336*		
Rec.	0.3714*	0.1391	0.4867**	0.5322**	0.3819*	
Grade	0.4724*	0.4268*	0.1650	0.1906	-0.1137	0.2050

Discussion

Based on these results, we finally evaluate our proposed hypotheses for our research study and begin to answer our research questions. First, we found support for hypothesis H3 because we observed that the inclusion of intelligent features led to an increase in the number of pages viewed by both the most and least active users in the system. Thus, recommending related pages and organizing pages by extracted keywords from the page content led users to explore the shared collaborative knowledge stored within the wiki, (1) boosting participation by less active users, and (2) potentially improving the individual knowledge of users after reading those pages.

Second, we found evidence both in favor of and opposing hypotheses H1, H2, and H4. We observed that simply including Web 2.0 features did not result in their usage (in Fall 2011), and thereby did not affect collaboration. However, once usage was encouraged through minimum requirements (in Spring 2012), not only were the features used more frequently, but their usage *often exceeded the requirements*. This led to (1) increased amounts of collaboration between users, supporting H1, (2) a diversification of roles in the collaboration process, supporting H2, and (3) increased quality of collaboration through higher grades earned, supporting H4.

Overall, with respect to studying wiki-based CSCL, we draw the following primary conclusion:

*Including advanced (e.g., Web 2.0 and intelligent) features to support and promote active participation and collaboration amongst student users of wiki-based systems is valuable and can lead to higher quality collaboration, but must be **appropriately** encouraged.*

Specifically, such encouragement does not simply mean imposing minimum requirements for collaboration as in our study and considered elsewhere in the literature (e.g., Ebner et. al, 2008). Instead, such encouragement could possibly arise through *improved interface design*, such as better highlighting the existence and benefit of

such features. Moreover, *increased education* of the use of the interface could also encourage more advanced feature usage. Our users were simply assigned their projects with little to no education in the usage of the wiki tool or more than a brief introductory text explaining its features. Finally, periodic pop-ups or other *encouragement from the system itself* could also lead to increased usage of advanced features and the resulting benefits of more active collaboration. This last approach has been effective in other collaborative systems to encourage users to participate (Wash & Lampe, 2012). In the future, we plan to study these potential methods of encouraging users to participate in order to promote more active and effective collaboration amongst users and reach a “critical mass” where sustained active collaboration benefits both the system and its users. We also intend to add surveys and possibly interviews to our data collection (1) to better understand the interplay between student knowledge, technology experience, and motivation with our advanced software features and (2) more precisely measure increases in student learning (e.g., assimilation and accommodation, Cress and Kimmerle, 2008).

References

- Antin, J., Cheshire, C., & Nov, O. (2012). Technology-mediated contributions: editing behaviors among new Wikipedians. *Proceedings of the 2012 ACM conference on Computer-supported cooperative work*, 373-382.
- Arazy, O., & Nov, O. (2010). Determinants of Wikipedia quality: the roles of global and local contribution inequality. *Proceedings of the 2010 ACM conference on Computer-supported cooperative work*, 233-236.
- Cole, M. (2009). Using wiki technology to support student engagement: Lessons from the trenches. *Computers & Education*, 52, 141-156.
- Cosley, D., Frankowski, D., Terveen, L., & Riedl, J. (2007). SuggestBot: using intelligent task routing to help people find work in Wikipedia. *Proceedings of the 2007 international conference on Intelligent user interfaces*, 32-41.
- Cress, U., & Kimmerle, J. (2008). A systematic and cognitive view on collaborative knowledge building with wikis. *Computer-Supported Collaborative Learning*, 3, 105-122.
- Ebner, M., Kickmeier-Rust, M., & Holzinger, A. (2008). Utilizing wiki-systems in higher education classes: a chance for universal access? *Universal Access in the Information Society*, 7(4), 199-207.
- Eck, A., & Soh, L.-K. (2012). Adaptive knowledge assistants. *Proceedings of the 2012 international workshop on Human-agent interaction design and models*, 107-114.
- Khandaker, N., & Soh, L.-K. (2009). ClassroomWiki: a collaborative wiki for instructional use with multiagent group formation. *IEEE Transactions on Learning Technologies*, 3(3), 190-202.
- Kittur, A., & Kraut, R.E. (2008). Harnessing the wisdom of crowds in Wikipedia: quality through coordination. *Proceedings of the 2008 ACM conference on Computer-supported cooperative work*, 37-46.
- Kittur, A., Suh, B., & Chi, E.H. (2008). Can you ever trust a wiki? Impacting perceived trustworthiness in Wikipedia. *Proceedings of the 2008 ACM conference on Computer-supported cooperative work*, 477-480.
- Kong, N., Hanrahan, B., Weksteen, T., Convertino, G., & Chi, E.H. (2010). VisualWikiCurator: human and machine intelligence for organizing wiki content. *Proceedings of the 2010 international conference on Intelligent user interfaces*, 367-370.
- Lampe, C., Obar, J., Ozkaya, E., Zube, P., & Velasquez, A. (2012). Classroom Wikipedia participation effects on future intentions to contribute. *Proceedings of the 2012 ACM conference on Computer-supported cooperative work*, 403-406.
- Moskaliuk, J., & Kimmerle, J., & Cress, U. (2008). Learning and knowledge building with wikis: The impact of incongruity between people’s knowledge and a wiki’s information. *Proceedings of the 2008 international conference for the learning sciences*, 2, 99-106.
- Moskaliuk, J., Kimmerle, J., & Cress, U. (2012). Collaborative knowledge building with wikis: The impact of redundancy and polarity. *Computers & Education*, 58, 1049-1057.
- Pancierera, K., Halfaker, A., & Terveen, L. (2009). Wikipedians are born, not made: a study of power editors on Wikipedia. *Proceedings of the 2009 GROUP conference*, 51-60.
- Peterson, E. (2009). Using a wiki to enhance cooperative learning in a real analysis course. *PRIMUS*, 19(1), 18-28.
- Wash, R., & Lampe, C. (2012). The power of ask in social media. *Proceedings of the 2012 ACM conference on Computer-supported cooperative work*, 1187-1190.
- Wheeler, S., Yeomans, P., and Wheeler, D. (2008). The good, the bad, and the wiki: Evaluating student-generated content for collaborative learning. *British Journal of Educational Technology*, 39(6), 987-995.

Acknowledgments

This research was supported by grants from the Dept. of Education (P200A040150), the NSF (DBI-0743783, DEB-0953766), and an NSF Graduate Research Fellowship (DGE-054850). We thank the developers of the Biofinity Project (from which the Written Agora was created), and the TAs assisting with the DPE classes.

Inhibiting Undesirable Effects of Mutual Trust in Net-Based Collaborative Groups

Tanja Engelmann, Richard Kolodziej, & Michail Kozlov,
 Knowledge Media Research Center, Schleichstrasse 6, D-72076 Tuebingen, Germany,
 Email: t.engelmann@iwm-kmrc.de, r.kolodziej@iwm-kmrc.de, m.kozlov@iwm-kmrc.de

Abstract: Experimental studies have demonstrated the effectiveness of the knowledge and information awareness approach by Engelmann and colleagues for improving computer-supported collaborative problem-solving. This approach informs group members about their collaborators' knowledge structures and underlying information visualized by digital concept maps. In our study, we investigated whether this approach may reduce undesirable effects of mutual trust. Trust is an important influencing factor with regard to behavior and performance of groups. High mutual trust can have a negative impact on group effectiveness because it reduces mutual control and, as a result, the detection of the others' mistakes. In an empirical study, 20 triads collaborating with the knowledge and information awareness approach were compared with 20 triads collaborating without this approach. The members of a triad were spatially-distributed and collaborated computer-supported. The results demonstrated that the availability of the knowledge and information awareness approach overrides the negative impact of too much mutual trust.

1. Challenges of our Information Age

The need for collaboration, especially for experts in different fields of knowledge, is ever rising in our information age, and certainly the distribution of experts around the world is not a real problem these days: These experts do not have to share the same physical space to work together on common tasks; instead, for this purpose, they can use current computer technologies such as video chatting or collaborative writing tools.

There are many advantages of computer-supported collaboration, for example, the mentioned spatial flexibility (Engelmann, Dehler, Bodemer, & Buder, 2009). However, this is not easy to achieve in an effective way. There are many reasons why interaction problems, especially regarding communication and coordination, often occur in a computer-mediated environment (Janssen, Erkens, Kanselaar, & Jaspers, 2007): One problem is the reduced contextual information, that is, missing information (e.g. non-verbal signals) which would be available in a face-to-face setting (Kiesler, Siegel, & McGuire, 1984). Another difficulty for virtual groups is that the members often do not know each other before they have to collaborate on a common task, and therefore, they do not know what their collaborators know. However, different lines of research have demonstrated the importance of knowing what the collaborators know (cf. Engelmann & Hesse, 2010): Research on *Audience Design* (e.g., Dehler-Zufferey, Bodemer, Buder, & Hesse, 2011) gives evidence that individuals adapt their texts depending upon the addressee. According to the *Knowledge Imputing* approach (Nickerson, 1999) effective communication requires a sufficient amount of correct knowledge about the communication partner's knowledge. If one overestimates the partner's knowledge, the partner might not even be able to understand statements (Nickerson, 1999). Studies on the *Theory of Transactive Memory System* (Wegner, 1986) confirm that the groups whose members know who is an expert on which topics achieve more in group tasks (e.g. Liang, Moreland, & Argote, 1995).

However, prior research has shown that it is not easy to acquire correct knowledge about the collaboration partner's knowledge (cf. Engelmann & Hesse, 2010): This knowledge was derived from both behavioral and categorical information as well as from the interaction with the collaboration partners. During this process, a lot of perception or evaluation mistakes can slip in (cf. Krauss & Fussell, 1991; Nickerson, 1999). In addition, according to the theory of transactive memory system (Wegner, 1986), sufficient common time is required to acquire this knowledge. Furthermore, there are situations in which the possibilities of acquiring knowledge about the partners' knowledge are strongly restricted (Engelmann & Hesse, 2010), for example, in a computer-supported collaborative setting with spatially distributed group members, who must deal with reduced contextual information (Kiesler et al., 1984).

2. The Approach for Fostering Knowledge and Information Awareness

In order to find a solution for the need for and the problem of acquiring knowledge about the collaboration partners' knowledge in computer-supported collaborative settings, Engelmann (née Keller) and colleagues developed an implicit approach, called the *Approach for Fostering Knowledge and Information Awareness* (Keller, Tergan, & Coffey, 2006). They defined knowledge and information awareness as being informed about the collaborators' knowledge structures and information underlying these structures (e.g. Engelmann & Hesse, 2011). The approach for fostering knowledge and information awareness provides the spatially distributed

group members with their collaborators' knowledge structures and their collaborators' information underlying these structures, both visualized by means of digital concept maps.

Concept maps are a well-established kind of knowledge visualization representing conceptual knowledge by means of hierarchically ordered concepts (in form of labeled nodes) and relations between the concepts (in form of labeled links) (Novak & Gowin, 1984). Digital concept mapping technologies moreover allow for adding nodes with links for accessing further information (e.g. Alpert, 2005). Concept maps are, therefore, well suited for fostering knowledge and information awareness.

Several experimental studies have confirmed that not only knowledge and information awareness can be quickly acquired by being provided with the collaborators' digital concept maps, but also that this approach improves collaboration and collaborative problem-solving of spatially distributed group members (e.g. Engelmann, Tergan, & Hesse, 2010; Engelmann & Hesse, 2010). Collaborative problem-solving fosters learning (e.g. Hausmann, Chi, & Roy, 2004). In addition, the experimental results have shown that this approach may also assist in overcoming important collaboration barriers: The study by Engelmann and Hesse (2011) provides evidence that the knowledge and information awareness approach fosters sharing, discussing, and cognitive processing of unshared information. In the study by Schreiber and Engelmann (2010), it was shown that this approach furthers the development of a transactive memory system. In the study by Engelmann and Kolodziej (2012), it was demonstrated that also in unstructured group situations self-created concept maps for fostering knowledge and information awareness may reduce the needed collaborative problem solving time and, therefore, the coordination effort within the group. In the current paper, we will focus on investigating a collaboration barrier having to do with the concept of mutual trust.

3. The Impact of Mutual Trust on Behavior and Performance of Groups

Trust is an important influencing factor with regard to behavior and performance of groups (Salas, Sims & Burke, 2005). Changes in the situation can have an impact on the role of trust in groups (e.g. Kramer, 1999). For example, the role of trust is dependent on the degree of structure in the situation (Dirks & Ferrin, 2001; Jarvenpaa, Shaw, & Staples, 2004): In situations with a low degree of structure, trust has a direct effect on group variables. In such situations, it is difficult to interpret the others' behavior. Therefore, their behavior is interpreted depending on the amount of trust that the group members have in each other. In situations with a moderate degree of structure, trust is a moderating factor. Factors for interpreting the others' behavior are given; however, trust influences how these factors are interpreted. In situations with high structure, the others' behavior can be directly evaluated. Trust is not used to interpret the others' behavior and, therefore, does not have any impact on group measurements.

In situations in which trust has an effect on group variables, the following relations are to be expected: In numerous publications (e.g. Jarvenpaa, Knoll, & Leidner, 1998), it is argued that mutual trust is an important influencing factor for group effectiveness. This was also confirmed by several empirical studies (e.g. Kanawattanachai & Yoo, 2002). Further empirical studies, for instance by Aubert and Kelsey (2003) as well as Jarvenpaa et al. (2004), have shown that trust has an effect on group efficiency, but not on group effectiveness.

These contradictory results could possibly be explained by another influencing factor, namely, correctness of individual performances: If group members with high mutual trust work without mistakes, this should result – according to Aubert and Kelsey (2003) as well as Jarvenpaa et al. (2004) – in a faster and, therefore, more efficient collaboration, since it is to be expected that high mutual trust reduces mutual control. When free from errors, high mutual trust should not have an impact on group effectiveness. However, if group members with high mutual trust make mistakes, these mistakes might not be discovered due to the reduced mutual control caused by having high mutual trust. This should lead to reduced group effectiveness (cf., Jarvenpaa et al., 2004; Dirks & Ferrin, 2001). Due to the fact that efficiency is defined as effectiveness per time, the time saved while performing the task has to be very high in order to obtain good efficiency with low effectiveness. Therefore, it is expected that low effectiveness will lead to poor efficiency.

Contrarily, low mutual trust should increase mutual control and, therefore, the needed time; that is, it should reduce group efficiency. However, there is a good chance that the mistakes of the collaboration partners will be discovered. As a consequence, higher group effectiveness can be expected.

Due to the fact that, compared to face-to-face collaboration, computer-supported collaboration is often accompanied by various difficulties (e.g., Janssen, Erkens, Kanselaar & Jaspers, 2007), it is most likely that the group members will make mistakes. For example, while creating their own concept maps visualizing their own knowledge structures, group members could “forget” to include parts of their knowledge or visualize some elements wrongly. As these knowledge representations were used as a starting point for net-based collaboration, these mistakes could decrease group performance. The following argumentation refers only to collaborations in which mistakes appeared.

Aubert and Kelsey (2003) have shown that trust is lower in computer-supported collaborative groups compared to face-to-face groups. Thus, virtual group members have a higher need for mutual control. However, because of the computer-support mutual control ability is limited. In such situations, mutual control is very

effortful. However, as in the study presented in this paper, there are possibilities that allow for mutual control in such virtual groups.

4. Experimental Study

The goal of the study was to investigate the impact of mutual trust depending on the availability of the knowledge and information awareness approach on group effectiveness and group efficiency of solving problems in virtual groups.

4.1 Hypotheses

Without being provided with the knowledge and information awareness approach (control condition), as already mentioned, it is to be expected that trust will affect group effectiveness: A variance regarding the amount of trust also appears in virtual groups despite lower mutual trust in virtual groups compared to face-to-face groups (Aubert & Kelsey, 2003). If mutual trust is high, it is to be expected that there is low mutual control, and therefore, mistakes will not be detected. In reference to our study, the group members would not check whether the contributions of the others are correct. This should decrease group effectiveness and – because of its relation to effectiveness (see above) – efficiency. However, even if mutual control is effortful in computer-supported settings, low trust should lead to mutual control, for example, by asking their collaborators if they are sure that their contributions are correct. This should reduce efficiency, while effectiveness should increase a bit. (Due to the mutual control difficulties, it is to be expected that not all mistakes will be detected.)

In addition, it is to be expected that by direct access to the collaborators' knowledge and information, the availability of the knowledge and information awareness approach (experimental condition) will facilitate mutual control. The ability for easy mutual control can therefore be given also in virtual groups. In prior studies (e.g. Engelmann & Hesse, 2010), it has been confirmed that the knowledge and information awareness approach is used if it is available. Therefore, there should not be an impact of mutual trust on mutual control; that is, there should be mutual control independent of the amount of mutual trust. Consequently, it is expected that the amount of mutual trust will not have an impact on group effectiveness and group efficiency. Due to the fact that first, the collaborators' contributions are checked and, therefore, their mistakes will be detected, and second, because the process costs of mutual control are low, an effective and efficient group performance is expected, compared to groups that collaborate without the knowledge and information awareness approach.

To sum up, we hypothesize the following effects under the assumption of the existence of individual mistakes included in the individual concept maps:

Hypothesis 1: Regarding group effectiveness as criterion, we expect a significant interaction between mutual trust and condition. In more detail, we expect that in the experimental condition, trust will not have an impact on group effectiveness, whereas in the control condition, high trust will reduce effectiveness because of less mutual control and, therefore, less mutual correcting of mistakes.

Hypothesis 2: Regarding group efficiency as the criterion, we only expect a main effect for condition in favor of the experimental condition; that is, the experimental condition will be more efficient compared to the control condition. We expect neither a main effect for trust nor an interaction of trust and condition on group efficiency.

4.2 Method

In the experimental study, an experimental condition consisting of 20 triads being provided with the knowledge and information awareness approach was compared to a control condition consisting of 20 triads collaborating without this approach.

4.2.1 Participants

Participants of the study were 120 students (84 female, 36 male) of a German university from different fields of study with an average age of 23.74 years ($SD = 3.47$). They volunteered to participate for payment. The participants, collaborating in groups of three, were randomly assigned to a control or an experimental condition.

The composition of the groups regarding gender were equal between the conditions; that is, both conditions had the same number of groups with no, one, two, or three women. The members of a group either did not know each other or hardly knew each other. There was no significant difference between the conditions regarding the degree of acquaintance among the members in a group ($F < 1$).

4.2.2 Setting and Materials

The members of a triad were spatially distributed and collaborated computer-supported. They communicated by using Skype (only audio). The experimental environment consisted of several shared and unshared working windows of CmapTools, a digital concept mapping software developed by the Florida Institute for Human and Machine Cognition (USA). The study was held in German. Therefore, for this paper, all contents have been translated into English.

The domain was concerned with rescuing a fictitious type of spruce forest and consisted of 13 concepts, 30 relations between the concepts and 13 pieces of background information (in parts divisible into sub-elements), each linked to a concept. These elements were evenly distributed among the three group members in a way that each member had the same amount of shared and unshared concepts, relations, and background information aspects. The shared elements were shared with either one collaborator or both collaborators.

Online questionnaires and instructions: An online questionnaire for assessing several control measure items (e.g. experience in working with computers and in groups) and for measuring the amount of initial mutual trust was included. For measuring mutual trust, several items taken from Amelang, Gold, and Külbel (1984), from Jarvenpaa, Knoll and Leidner (1998), as well as from Jarvenpaa et al. (2004), were used that were translated into German and partly adapted to our experimental setting. The 15 control measure items and the 13 items for assessing mutual trust were designed as multiple-choice items with five-point rating scales, ranging from complete agreement to no agreement. Examples of these items are: "I can create visualizations by means of a computer" (control measure item) and "In contact with strangers, it is better to be careful until they have provided evidence that one can trust them."

An online knowledge test was used to measure the knowledge of group members regarding their own and their collaborators' knowledge on particular relations and concepts. It consisted of 24 multiple-choice test items. These items were classified with regard to who possessed the requested knowledge, resulting in four types of items: (1) items asking for one's own unshared elements, that is, items measuring knowledge that one had alone in his/her individual map, (2) items asking for the collaborators' unshared elements, that is, items measuring knowledge that only one of the collaborators had, (3) items asking for shared elements that one shared with one of the collaborators, that is, items measuring knowledge that one had together with one of the collaborators, and (4) items asking for shared elements of the collaborators, that is, items measuring knowledge that only the two other collaborators had. For each item the participants stated whether they were certain that they had answered it correctly (rating scale with three possible answers: low, middle, and high certainty). In order to assess the knowledge and information awareness referring only to the collaborators' knowledge and information only the categories 2 and 4 are relevant.

A second online questionnaire was used to evaluate the study, that is, to assess among other things aspects of collaboration and mutual control as well as to subjectively rate the quality of the group performance. In addition, in the experimental condition, only the usefulness of the knowledge and information awareness approach was assessed. Again the items were designed as rating scales with answer categories ranging from one point for no agreement and five points for complete agreement. The questionnaire contained 50 items in the control condition and – due to the additional items – 56 items in the experimental condition.

The group members were provided with paper-based instructions on how to use CmapTools and with paper-based instructions to explain all the phases of the study and the tasks to be completed by the group members.

4.2.3 Procedure

After informing the participants about the set-up of the study and obtaining their signed letter of agreement to take part in the study, the three members of a group were sent to separate rooms each equipped with a desk and a computer. They began the study by individually filling out the online questionnaire for assessing several control measure items and their initial mutual trust. After that each group member practiced using CmapTools. In the subsequent phase, the group members were informed that they should imagine that they were three experts who would have to mutually rescue a spruce forest. They were told that in order to rescue this forest they would have to solve two problems, namely, which pesticide and which fertilizer they would use. The fertilizer problem could only be solved correctly if the pesticide problem was solved correctly. The groups were told that there was only one solution for each problem. They were told further that they should imagine that in the past they had taken some notes regarding these problem domains and that –based on these notes – they had to create their own digital concept map visualizing their own knowledge and information. They had 20 minutes to create their individual concept map. This was enough time for each group member to finish the individual map. The log files of creating the individual maps (by CmapTools) were recorded. After that the members of the experimental groups were additionally provided with their collaborators' individual concept maps for 5 minutes. In order to control the time in the individual phase, the members of the control groups had 5 more minutes for viewing their own individual map.

Then the collaborative problem solving phase started which lasted 35 minutes. In this phase, the groups had to solve the two problems for rescuing the forest. In order to accomplish this, they had to compile their individual conceptual knowledge by creating a digital group concept map together in a shared working window. The background information aspects were irrelevant to the problem, but this was not known to the group members. The group members could speak with each other by using Skype. Besides the shared working window, each member of the control condition had access to their own individual concept map that they had

created in the individual phase (see Figure 1, on the left). The members of the experimental condition were additionally provided with their collaborators' individual concept maps visualizing their collaborators' conceptual knowledge and background information (see Figure 1, on the right).

In this collaborative phase, log files of creating the group maps (by CmapTools), as well as video and audio files (by Camtasia), were recorded.

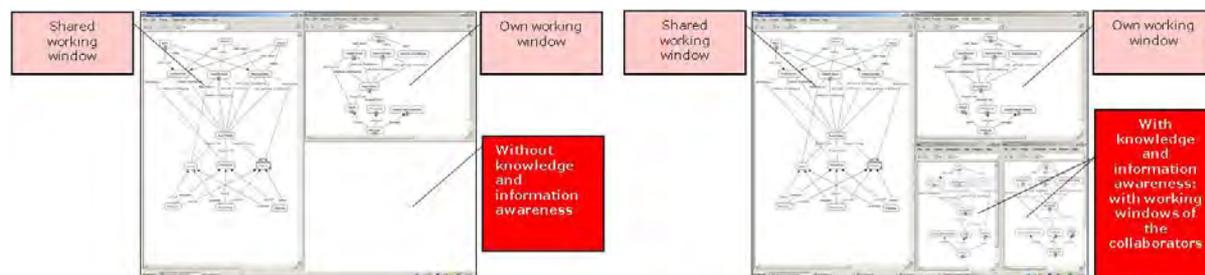


Figure 1. Collaborative phase (left: control condition; right: experimental condition).

4.3 Measures

To answer the hypotheses, besides the two conditions, the following measures were used as predictor measures:

A factor analysis with Varimax rotation with the 13 trust items included in the questionnaire on control measurements resulted in two interpretable factors (cf. Bortz, 1999), namely, “general skepticism regarding others” (in the following this will be called initial skepticism) and “trust in others due to experience” (in the following this will be called initial trust). As expected, there were no significant differences between the conditions regarding these factors (for initial skepticism: $F < 1$; for initial trust: $M_C = 0.16$; $M_E = -0.16$; $F(1, 38) = 1.06$, $MSE = 1.00$, $p = .31$).

Regarding the quality of the problem solutions of the groups as criterion measures for group effectiveness, we differentiated between two dependent measures, namely, solving the pesticide problem correctly and solving the fertilizer problem correctly. If a group solved the pesticide problem correctly, one point was given; if the wrong pesticide was chosen, no points were given. Analogous to this, if a group solved the fertilizer problem correctly, one point was assigned; if the wrong fertilizer was chosen, no points could be attained. The interrater agreement for both measures was Cohens' $\kappa = 1$ indicating a perfect match (Cohen, 1960).

Regarding group efficiency, the following measures were differentiated: Because in this study effectiveness was determined as a dichotomy variable (solved vs. not solved), to determine efficiency measures, only those triads were included that solved the pesticide problem and/or the fertilizer problem correctly. Two measures were differentiated: The variable “efficiency of choosing the correct pesticide solution” refers to the collaboration time needed to decide on the correct pesticide solution. The variable “efficiency of choosing the correct fertilizer solution” refers to the collaboration time needed to decide on the correct fertilizer solution. The interrater agreement was $ICC = .96$ for efficiency of deciding on the correct pesticide solution and $ICC = .96$ for efficiency of deciding on the correct fertilizer solution (two-way mixed single measures, cf. Shrout & Fleiss, 1979).

5. Results

The experimental condition in which the group members were provided with the knowledge and information awareness approach was compared with the control condition in which the group members collaborated without this approach. All analyses presented here are based on the group level, because of both the dependent variables were variables on the group level and the individuals in a group are not independent of each other. This means that also the trust variables were calculated as group means in order to use them on the group level.

The inclusion of a covariate was not necessary because we did not find significant differences between the conditions regarding the control measures. As a manipulation check, it was analyzed whether our knowledge and information awareness approach fostered the acquisition of knowledge and information awareness: Accordant with the results of prior studies (e.g. Engelmann et al., 2010), the knowledge test resulted in a significant higher knowledge and information awareness value for the experimental condition compared to the control condition ($M_C = 18.77$ $M_E = 22.87$; $F(1, 38) = 7.41$; $MSE = 22.66$; $p = .01$; $\eta_p^2 = .16$). This value was calculated as the sum of item categories 2 and 4 each weighted by the correctness certainty (see Section 4.2.2).

Due to the fact that we were interested in interaction effects between condition and variables of trust, moderator analyses were conducted following Aiken and West (1991). The necessary requirements for conducting regression analyses were tested each time: All analyses met the global test statistic (cf. Peña & Slate, 2006).

5.1 Group Effectiveness as Criterion Variable

The regression analyses with effectiveness measures as the criterion variable as well as condition, initial trust, and their interaction as predictor variables led to the following results:

The regression analysis with the solution of the pesticide problem as the criterion variable revealed no significant conditional effect for initial trust ($b = .04$, $SE = .09$; $\beta = .09$, $p = .62$) or for belonging to a particular condition ($b = -.02$, $SE = .07$; $\beta = -.04$, $p = .78$), adjusted $R^2 = -.018$, $F(2, 37) = 0.66$, $p = .52$. However, as predicted, a significant interaction between condition and initial trust appeared ($b = .25$, $SE = .09$; $\beta = .50$, $p < .01$), adjusted R^2 change = .153, F change (3, 36) = 3.35, $p = .03$: Simple slope analyses indicated, as assumed, that higher initial trust significantly impaired the solution of the pesticide problem of the control condition ($b = -.21$, $SE = .08$; $\beta = -.42$, $p = .02$). In the experimental condition trust had, as expected, no significant effect on the solution of the pesticide problem ($b = .29$, $SE = .15$; $\beta = .60$, $p = .06$), but in contrast to our hypothesis, a marginal effect indicating that high trust marginally increased the group effectiveness.

Regarding the measure “solution of the fertilizer problem” as criterion variables, no significant effects resulted. In addition, the regression analyses with effectiveness measures as the criterion variable, as well as condition, initial skepticism, and their interaction as predictor variables, did not result in significant interactions. Therefore, these results are not reported.

5.2 Group Efficiency as Criterion Variable

The regression analyses with efficiency measures as the criterion variable, as well as condition, initial trust, and their interaction as predictor variables, led to the following results:

Accordant with our hypothesis, the regression analysis with efficiency of choosing the correct pesticide solution as the criterion variable revealed a significant main effect for belonging to a particular condition ($b = -.193.78$, $SE = 93.74$; $\beta = -.40$, $p = .05$): The experimental groups needed less time for finding the correct pesticide solution compared to the control groups ($M_C = 19:56$, $SD_C = 7:21$, $M_E = 13:15$, $SD_E = 7:55$). As expected, we did not find a significant main effect for initial trust ($b = -.127.96$, $SE = 117.38$; $\beta = -.23$, $p = .29$), adjusted $R^2 = .16$, $F(3, 21) = 2.50$, $p = .09$, nor did a significant interaction between condition and initial trust appear ($b = -180.10$, $SE = 117.38$; $\beta = -.32$, $p = .14$).

In line with this result, the regression analysis with efficiency of choosing the correct fertilizer solution as criterion variable also revealed, as assumed, a significant main effect for the belonging to a particular condition ($b = -166.27$, $SE = 63.20$; $\beta = -.50$, $p < .05$): The experimental groups needed less time for finding the correct fertilizer solution compared to the control groups ($M_C = 22:40$, $SD_C = 4:42$, $M_E = 17:16$, $SD_E = 5:23$). As predicted, we did not find a significant main effect for initial trust ($b = 33.07$, $SE = 78.99$; $\beta = .09$, $p = .68$), adjusted $R^2 = .15$, $F(3, 23) = 2.52$, $p = .08$, nor did a significant interaction between condition and initial trust appear ($b = -16.43$, $SE = .78.99$; $\beta = -.05$, $p = .84$).

The regression analyses with efficiency measures as the criterion variable, as well as condition, initial skepticism, and their interaction as predictor variables, did not result in significant interactions. Therefore, these results are not reported.

6. Discussion and Implications

Our initial point was the conflicting empirical findings regarding the impact of mutual trust on group performance. While some researchers have demonstrated that trust has an effect on group effectiveness (e.g., Kanawattanachai & Yoo, 2002), others have pointed out that trust has an effect on group efficiency, but not on effectiveness (e.g. Aubert & Kelsey, 2003). In this paper, we argued that these conflicting findings could be explained by including “the amount of individual errors made by group members” as a further factor. We argued that in CSCL-settings, we must expect that group members will make mistakes due to the difficulties caused by the use of collaborative technology (cf. Kiesler et al., 1984; Janssen et al., 2007). In addition, compared to face-to-face situations, in CSCL-settings, the mutual trust is lower, and therefore, the need for mutual control is higher, while mutual control is much more effortful. Hence, we concluded that in CSCL-settings with increasing mutual trust, mutual control will be increasingly reduced, and as a result, it is likely that mistakes will not be detected, decreasing group effectiveness. In contrast, with decreasing trust, we expected increasing mutual control and, therefore, increasing effectiveness.

We further argued that our knowledge and information awareness approach counteracts this effect of trust. This approach provides group members with their collaborators’ externalized knowledge structures and underlying information and, therefore, allows for easy mutual control also in virtual settings. Prior studies have shown that this approach is used if it is available (e.g. Hesse & Engelmann, 2010). Therefore, it was expected that the groups will check each other’s work if provided with this approach, independent of their amount of mutual trust. Accordingly, it was expected that in the experimental condition, trust does not affect effectiveness.

To sum up, we expected a significant interaction between condition and initial trust on group effectiveness in a way that increasing trust will decrease effectiveness in the control condition, while in the experimental condition trust will not have an effect on group effectiveness.

The results of the presented study confirmed our hypothesis: In the control condition with increasing mutual trust, group effectiveness decreased. In the experimental condition, mutual trust did not significantly affect group effectiveness; however, there was a marginal effect indicating that high mutual trust marginally increased group effectiveness. The negative impact of mutual trust in the control condition can be counteracted successfully by the availability of the knowledge and information awareness approach. We explained this result with the fostering of mutual control when the knowledge and information awareness approach is available. The marginal effect may demonstrate that in the experimental condition mutual trust even fostered mutual control. Another explanation could be that the knowledge and information approach leads to a situation with high structure in which trust does not have an effect on group variables anymore (e.g. Dirks & Ferrin, 2001)

However, it is interesting to note that these effects were only found with regard to solving the pesticide problem, but not with regard to solving the fertilizer problem. A reason for this could be that the fertilizer problem could only be solved correctly if the pesticide problem was solved correctly; that is, solving the fertilizer problem depended more on solving the pesticide problem than on other reasons. In addition, it is also interesting that we only found this effect for the factor initial trust, but not for the factor initial skepticism. Initial trust was based on items such as “In most of the groups that I have worked with in the past, the group members trusted each other” or “In the past, I have worked mostly together with trustworthy people”. Therefore, it refers to the amount of general trust in others developed by prior experience. Initial skepticism was based mainly on items such as “One should be very careful if working together with strangers” or “In current times, with so much competition, you should be on the alert or someone will probably take advantage of you” and, therefore, refers mainly to a generalized skepticism about others, based more on a general attitude. This difference seems to be crucial: Initial skepticism seems not to be just the opposite of initial trust. There appears to be quality differences, at least with respect to our factors. Future studies are needed to explain this difference.

With regard to group efficiency, we expected for control groups with high trust also low efficiency because efficiency is dependent on effectiveness. For control groups with low trust, we also expected low efficiency due to much mutual control that takes time. For the experimental groups, we expected, independent of the amount of trust, high group efficiency due to the low process costs for checking the others.

This hypothesis was confirmed: In line with prior study results (e.g. Engelmann & Hesse, 2010), the experimental groups solved both of the problems faster compared to the control groups. As expected, neither a main effect for trust nor an interaction between trust and condition on group efficiency were observed. Together with the findings on group effectiveness, this result demonstrated that mutual trust may have an effect on group effectiveness, but not on group efficiency. This is accordant with Kanawattanachai and Yoo (2002) and Jarvenpaa et al. (2004). Thus, this paper also contributes to solving the conflicting findings in the literature regarding the effects of trust.

Our hypotheses were derived, among others things, from the assumptions regarding mutual control. However, in this study, we did not analyze mutual control. Future analyses could be based on the recorded discussions. However, in order to analyze mutual control in a better way, eye tracking is needed. Eye tracking results could contribute to further clarifying the postulated relations.

To sum up, this study demonstrated that the availability of the knowledge and information awareness approach overrides the negative impact of too high mutual trust that is to be expected especially in CSCL-settings. Additionally, this study further contributes to clarifying the impact of trust on group effectiveness and group efficiency in computer-supported collaborative situations depending on different situational factors such as being provided with a knowledge and information awareness approach or not.

7. References

- Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, California: SAGE Publications, Inc.
- Alpert, S. R. (2005). Comprehensive mapping of knowledge and information resources: The case of Webster. In S.-O. Tergan, & T. Keller (Eds.), *Knowledge and information visualization. Searching for synergies*. (pp. 220-237). LNCS 3426. Heidelberg: Springer-Verlag.
- Amelang, M., Gold, A., & Külbel, E. (1984). Über einige Erfahrung mit einer deutschsprachigen Skala zur Erfassung zwischenmenschlichen Vertrauens (Interpersonal Trust). *Diagnostica*, 30(3), 198-215.
- Aubert, B. A., & Kelsey, B. L. (2003). Further understanding of trust and performance in virtual teams. *Small Group Research*, 34(5), 575-618.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Dehler-Zufferey, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Partner knowledge awareness in knowledge communication: Learning by adapting to the partner. *The Journal of Experimental Education*, 79(1), 102-125.

- Dirks, K. T. & Ferrin, D. L. (2001). The role of trust in organizational settings. *Organization Science*, 12(4), 450-467.
- Engelmann, T., Baumeister, A., Dingel, A., & Hesse, F.W. (2010). The added value of communication in a CSCL-scenario compared to just having access to the partners' knowledge and information. In J. Sánchez, A. Cañas, & J.D. Novak (Eds.), *Concept Maps making learning meaningful: Proceedings of the 4th International Conference on Concept Mapping* (Vol. 1, pp. 377-384). Chile: University of Chile.
- Engelmann, T., Dehler, J., Bodemer, D. & Buder, J. (2009). Knowledge awareness in CSCL: a psychological perspective. *Computers in Human Behavior*, 25(4), 949-960.
- Engelmann, T. & Hesse, F.W. (2010). How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *International Journal of Computer-Supported Collaborative Learning*, 5(3), 299-320.
- Engelmann, T. & Hesse, F. W. (2011). Fostering sharing of unshared knowledge by having access to the collaborators' meta-knowledge structures. *Computers in Human Behavior*, 27, 2078-2087.
- Engelmann, T. & Kolodziej, R. (2012). Do virtual groups recognize situations in which it is advantageous to create digital concept maps? In A. Cañas, J.D. Novak. & J. Vanhear (Eds.), *Concept Maps: theory, methodology, technology. Proceedings of the 5th International Conference on Concept Mapping* (Vol. 1, pp. 172-179). Malta: University of Malta.
- Engelmann, T., Tergan, S.-O. & Hesse, F.W. (2010). Evoking knowledge and information awareness for enhancing computer-supported collaborative problem solving. *The Journal of Experimental Education*, 78, 1-20.
- Hausmann, R. G. M., Chi, M.T. H. & Roy, M. (2004) Learning from collaborative problem-solving: An analysis of three hypothesized mechanisms. In K. D. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Cognitive Science society* (pp. 547-552). Mahwah, NJ: Lawrence Erlbaum.
- Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education*, 49, 1037–1065.
- Jarvenpaa, S. L., Knoll, K., & Leidner, D. E. (1998). Is there anybody out there? Antecedents of trust in global virtual teams. *Journal of Management Information Systems*, 14(4), 29-64.
- Jarvenpaa, S. L., Shaw, T. R., & Staples, D. S. (2004). Toward contextualized theories of trust: the role of trust in global virtual teams. *Information Systems Research*, 15(3), 250-267.
- Kanawattanachai, P. & Yoo, Y. (2002). Dynamic nature of trust in virtual teams. *The Journal of Strategic Information Systems*, 11, 187-213.
- Keller, T., Tergan, S.-O. & Coffey, J. (2006). Concept maps used as a “knowledge and information awareness” tool for supporting collaborative problem solving in distributed groups. In A. J. Cañas & J. D. Novak (Eds.), *Concept Maps: Theories, Methodology, Technology. Proceedings of the Second International Conference on Concept Mapping* (pp. 128-135). San José: Sección de Impresión del SIEDIN.
- Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American-Psychologist*, 39, 1123-1134.
- Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Reviews Psychology*, 50, 569-598.
- Krauss, R. M., & Fussell, S. R. (1991). Perspective-taking in communication: Representations of others' knowledge in reference. *Social Cognition*, 9, 2–24.
- Liang, D. W., Moreland, R., & Argote, L. (1995). Group versus individual training and group performance: The mediating role of transactive memory. *Personality and Social Psychology Bulletin*, 21(4), 384-393.
- Nickerson, R. S. (1999). How we know – and sometimes misjudge – what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, 125(6), 737-759.
- Novak, J. D. & Gowin, D.B. (1984). *Learning how to learn*. New York: Cambridge University Press.
- Peña, E. A., & Slate, E. H. (2006). Global validation of linear model assumptions. *Journal of the American Statistical Association*, 101(473), 341–354.
- Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a “big five” in teamwork? *Small Group Research*, 36(5), 555-599.
- Schreiber, M. & Engelmann, T. (2010). Knowledge and information awareness for initiating transactive memory system processes of computer-supported collaborating ad hoc groups. *Computers in Human Behavior*, 26, 1701-1709.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlation: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.
- Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behaviour* (pp. 185-208). New York: Springer.

Constructing and Deconstructing Materially-Anchored Conceptual Blends in an Augmented Reality Collaborative Learning Environment

Noel Enyedy, UCLA, 2027 Moore Hall, Box 951521, Los Angeles, CA 90095 USA, enyedy@gseis.ucla.edu
Joshua A. Danish, Indiana University, 201 North Rose Ave, Bloomington, IN 47405, jdanish@indiana.edu
David DeLiema, UCLA, 2027 Moore Hall, Box 951521, Los Angeles, CA 90095 USA,
david.deliema@gmail.com

Abstract: Science and math school activities around modeling often involve students stepping into a simulation to play the first-person roles of (often inanimate) components. In this case study, we examine how a student maps her own experience onto a ball to simulate the physics of force and friction. We study this mapping from a conceptual blending perspective, tracking how the narrative structure of a board game, the physical floor materials (e.g. linoleum), the student's first-person embodied experiences, the third-person live camera feed, and the augmented reality symbols become integrated in the modeling activity. The student's concepts of force and friction, in turn, are rooted in the blend between the narrative, the body, and the physical materials.

Introduction

There is a new class of computer-supported tools to aid learning referred to as mixed reality or augmented reality (henceforth AR). In AR environments, the physical world is digitally enhanced by viewing *reality* through a video feed or device that augments the display with a graphical or informational overlay. Studies have shown AR to be successful at promoting learning across the grade levels and across subject domains (Enyedy et al., 2012; Klopfer, 2008). While designing new technologies that effectively promote learning is a laudable goal in and of itself, as learning scientists, our primary goal should be to discover *why* these new technologies work. Further, as learning scientists, we want to turn the question on its head and ask what these new technologies can reveal about the basic processes of learning and instruction. In this paper we suggested that AR is uniquely positioned to support learning through its ability to support students in developing conceptual blends (Fauconnier & Turner, 1998)—cognitive spaces developed through the layering of multiple prior ideas in a way that allows students to draw new inferences.

An example Augmented Reality system

In the Learning Physics through Play (LPP) project, we designed an augmented reality system that uses socio-dramatic play as a form of scientific modeling and helps young students learn the core concepts of force and motion (Enyedy et al., 2012). There are two key components to the LPP system: 1) an augmented reality system that uses computer vision to record and display the students' physical actions and locations, and 2) software that translates this motion into a physics engine and generates a visual display based on the sensing data. We tracked students' physical motion in a 12' x 12' carpet area at the front of the classroom to create a *modeling space*. In this space, young children make predictions by pretending to be objects in motion and they see (simultaneously) their physical motion projected onto a large screen behind them in the form of an animated ball. For example, a student might act out the motion of a ball given a large force by walking quickly over tile and then slowly over an imagined sand pit.

After making predictions by directly modeling motion with their bodies, students in the LPP project seamlessly transition into a physics microworld, comparing their predictions to what happens in the ideal Newtonian simulation. Like other microworlds, LPP allows students to see and manipulate a situation in ways impossible in the real world (e.g., turning off friction). We call students' initial activities in the AR system *play-as-modeling* because students are oriented toward using multiple experiences and resources to model motion as a set of rules. Much like in pretend play, one's activity is governed by and oriented toward articulating the rules of the imaginary situation (Sidnell, 2011). During these play-as-modeling activities, students wear geometric figures mounted on cards or hats. The computer can track the motion of several figures at once, and the scene is displayed on a shared interactive whiteboard. Instead of seeing themselves walk around the rug, students see a ball moving across the floor, propelled by forces and slowed down by friction.

An important part of our pedagogical design was that the students developed all the images of objects, invisible forces, and the background art used in the LPP system during earlier lessons. Inventing these representations increased understanding of the target concepts and helped students create a personally meaningful context for the activities. Moreover, as students refined their symbols collectively, they were also determining which aspects of the phenomenon were important to capture in its representative symbol. In this

way the activities slowly transitioned students from play-as-modeling to reasoning from symbols and concepts in a way that more closely resembles what is commonly recognized as scientific modeling.

In this project, the vast majority of second-grade students significantly improved their understanding of physics (see Enyedy et al., 2012 for full details). To date we have been able to illustrate what learning looked like in this environment focusing on our two design principles of the role of play and the role of progressive symbolization. However, what is needed is a microgenetic account of learning that allows us to pinpoint the details of how the affordances of AR relate to cognition and learning. Our prime candidate is conceptual blends. Conceptual blending describes a type of reasoning where selective inputs from two separate spaces are projected into a new space together to form an emergent structure (Fauconnier, 1994; Fauconnier & Turner, 1998). The blending framework helps us account for how the AR environment laminates virtual semiotic resources onto material structures.

Our goal in this paper extends beyond applying the conceptual blending framework to our data. We wish to extend the blending framework in three ways. First, we wish to argue that there is a special case of conceptual blends where one of the source domains is the student herself. The first-person, subjective perspective as a source domain may create a powerful type of blend that is particularly useful when learning or problem solving. This is consistent with some of the hypotheses of the power of agent based reasoning (Wilensky & Resnick, 1999) and 'liminal worlds' (Ochs, Gonzalez, Jacoby, 1996). Second, conceptual blending has primarily been used to explain stable forms of reasoning. We are particularly interested in learning. We will extend the blending framework by examining learning as the process of construction and partial deconstruction of blends. Third, although the blending framework was developed as a cognitive theory that makes rather large assumptions about the amount of computation that occurs within the head, we follow Hutchins' lead (2005) and examine the material circumstances of blending, attempting to remain agnostic about the degree to which blending happens inside the head. These physical objects in material (Hutchins, 2005) or real (Dudis, 2004) space—including the body, visual symbols, and physical objects—become components in the microworld that structure students' inferences.

In summary, this research project is oriented toward the following research question: How does the conceptual blending framework account for students and teachers' interactions with multiple resources (e.g. bodies, symbols, physical materials, narratives) during a microworld learning environment? What can the blending framework reveal about foundational processes of learning and instruction? Using the conceptual blending framework, we track how the verbal discourse, augmented-reality technology, physical objects, abstract symbols, and students' own bodies selectively fuse together to create a blend through which students reason about physics.

Conceptual Framework

Conceptual blending, an extension of mental spaces theory (Fauconnier, 1985), is a general model for the integration of concepts and the creative construction of meaning. In theory, a conceptual blend is created by coordinating multiple, distinct conceptual spaces, or source domains, and projecting them into a hybrid conceptual space that has emergent properties not found in the source domains (Fauconnier & Turner, 1998). For example, Fauconnier and Turner (1998) offer the hypothetical example of a professor who is lecturing and begins to have an argument with Kant. In this case, one source domain is the modern day professor. A second source domain is Kant himself, dead now for hundreds of years. Projected together into a blend, one can imagine how Kant might comment on the writings of Hegel or Wittgenstein or argue with the modern-day professor he never could have known would exist. That is, the blend has emergent properties that afford the production of new inferences.

The process of conceptual blending is hypothesized to involve three operations. The first operation is composition, where the different source domains are evoked and elements from one source domain are explicitly mapped to another. The second phase is completion, where an inference or a computation is made from the emergent properties of the blend. Often, completion is thought to involve filling in the blend by matching it to memories or frames stored in long-term memory (Coulson & Oakley, 2000). However, we argue that the important aspect of completion is putting the blend in relation to a goal and then using the blend as a tool to achieve that goal. As many have noted about representations and other mental structures, a structure in the absence of activity is meaningless (Greeno & Hall, 1997) and computation assumes that there is a reason for making the computation. Hence, for us completion is fundamentally about putting the blend to use. The third phase is elaboration. Closely related to completion, elaboration involves extending the blend by continuing to bring in new elements, running the blend as a simulation, and extending it to new situations. In our analysis, and for education more generally, this is perhaps the most important part of blending, as it is here where different semiotic resources are put in relation to one another in different combinations to produce new insights.

Materially Anchored Blends

A potential difficulty in using conceptual blends to inform educational research is that, consistent with the norms of cognitive linguistics, CB theory was developed from hypothetical cases rather than empirical cases. As a result, it can be seen as broadly applicable to almost every case of reasoning (Coulson & Oakley, 2000). Further, these hypothetical cases involving purely mental computations and blends can be difficult to verify. In response, many have considered the relationship between (observable) physical materials and conceptual content (Dudis, 2004; Hutchins, 2005; Lidell, 1998). Hutchins (2005), for example, extended this work to a number of empirical cases where one can see the computations in the blend being performed in the material world. These ‘materially anchored blends’ re-envision the composition phase as the construction of material objects that literally superimpose structures on top of one another. For example, in a historical case from nautical navigation, he shows how the 32 points of the compass rose, which represents the cardinal directions, is superimposed with solar time (i.e., a 24 hour clock), dividing 24 hours into 32 45-minute periods. Because these 45 minutes were a good approximation of lunar time and the difference between high tide on consecutive nights, this blended structure was then used to compute at what time high tide would occur at a given port. The blend in this case was external and the computation was done by manipulating the representational state of the material world. However, it is important not to read too far into Hutchins’ examples, as this would preclude the option that some of the structures in the blend are not materially present but are instead made present by the subject through action, talk, or imagination.

Liminal blends

In our case, one of the central resources that is being blended with other semiotic resources is the child’s own body, an example of a real space blend involving gesture or action (Parrill & Sweetser, 2004). In pretending to be an object in motion and physically moving in the AR world, students using their bodies-as-objects are blended with abstract symbols and rules articulated through talk. To understand this special class of materially anchored and embodied blends, we draw on the work of Ochs and colleagues (1996) who coined the term ‘liminal worlds’ to describe cases where “the distinction between the scientist as subject and the physical world as object is blurred” (p. 347). In a study of professional physicists trying to understand emergent theories of the atomic structure of condensed matter, Ochs, Gonzalez, and Jacoby (1996) found that scientists were, “taking on the perspective of (empathizing with) some object being analyzed and by involving themselves in graphic (re)enactments of the physical events” (p. 360). For example, in trying to describe a finding related to atomic spin, a scientist switched into first person language and imagined himself to be the atom as it moved through a series of transitions, saying things such as, “when I come down I’m in the domain state,” (p. 331). Ochs et al. described these linguistic constructs where the participants moved between a normative scientific description of a phenomenon to more personal 1st person description as *liminal worlds*, because they were episodes in which objective facts were blended together with subjective reasoning from a first-person perspective. These liminal worlds created a qualitatively different set of resources from which to reason and were found to be productive in model and theory building. The LPP environment deliberately created this sort of liminal world where one’s subjective understanding (and the resources that come with embodied cognition) is laminated onto the more formal and symbolic world of traditional computer simulations, and where students are supported in moving fluidly between the two.

Methods

Our analysis is grounded in the tradition of cognitive ethnography (Hutchins, 2003). Video tapes of a single lesson of second-grade students engaged in learning about friction were used to inductively examine how the conceptual blending framework applied to our data. The activity itself brings together students, teachers, physical materials, abstract symbols, and live video in an augmented reality simulation focused on modeling an object’s trajectory through different types of friction. The class session occurred within a larger 15-week unit on basic physics. In this analysis, we attempt to trace what resources were being mapped together (composition), what inferences or computations were being made about the speed of a ball under different conditions (completion), and the way that the publicly available blend was elaborated through collaborative activity (elaboration). The case study student chosen for analysis, Marissa (a pseudonym), was fairly typical of the class as a whole. Most important for the present paper, her qualitative answers on the topic of friction showed that on the post test she understood the mechanism for friction, but still had difficulty in conceptualizing low or no friction environments. This was typical of our results for the intervention as a whole. In Enyedy et al. (2012), we reported that only 16 of 43 (37%) of the students received significantly higher scores on a question that addressed friction during the posttest than on the pre-test ($Z = 2.38$, $p = 0.02$). For example, when asked why friction slows and stops an object, Marissa explained: “Because the grass has a hard friction...It’s bumpy and it sticks up to the ball, have to fight to get over it.” However a little further into the question Marissa talks about what happens when the ball rolls onto ice: “It will go faster. Because it’s just smooth surface.” In this way,

Marissa fits the profile of many of the students in the class in showing a promising but incomplete understanding of friction.

Findings

In an activity aimed at having students explore the effect of friction on motion, the teacher lays out a life-sized game board on the floor—in reality a long strip of paper marked off into several squares. The students take turns ‘playing’ the ball and deciding how the speed of the ball changes as they receive force cards or friction cards. At the same time, an overhead camera records the play space and projects a live feed on a white board mounted to the wall, and overlays the friction and force cards with symbols on the video feed. That is, the student, force cards, and friction cards on the carpet space in turn appear in the video space as a black ball, forward-facing red arrows, and backward-facing red arrows (see figure 1a and 1b).



Figure. 1a and 1b

In this first section, we demonstrate *composition*, how disparate resources from distinct spaces in the classroom become mapped together to create the life-sized board game environment. That is, we show how the discourse between students and instructors in addition to the material anchors—despite being spread out over time and in the classroom—fuse together or join side-by-side into a board game blend.

Composing the Floor space. The first space established in the activity is the floor space—a 10-foot long, rectangular sheet of white paper marked off into a dozen 10” x 10” squares. Three squares have real sheets of flooring material—linoleum, carpet, and an outdoor welcome mat. Researcher 1 notifies students that they will need to place cardboard patterns in the appropriate places so that the computer knows the correct amount of force or friction in each game square. The floor space is marked with multiple resources: cardboard patterns, paper, students’ bodies, and floor materials.

Composing the Narrative space of playing a game. Researcher 1 helps to establish the overarching narrative structure. Researcher 1 sits down with the students on the carpet and initiates a whole-class discourse that explicitly maps the conventions of game playing onto the physical floor space. She makes a sweeping gesture from the start to the finish of the paper board game, showing the spatial trajectory typical of a board game.

Composing the overhead Live Feed space. As the activity unfolds, students quickly orient towards a live feed from a camera mounted directly above the carpet space and pointing downward. The camera feed is projected onto the white board. That is, if students look toward the white board, they can see live video of the carpet (and themselves moving around) seen from a bird’s eye perspective. This creates a mapping between students’ first-person perspectives and the camera’s third-person perspective.

Composing the Video Symbol space. Researcher 2 hands Marissa one of the flat cardboard symbols and says, “Marissa, do you want to hold the ball while you walk?” This interaction blends Marissa’s first-person experience of her body, the video image of her body, and the animation of a ball into one object. Other cardboard pieces appear as colored symbols in the live feed space, floating on top of the carpet. The ball symbol appears on screen as a black ball and the 2-force cards appear on screen as two horizontal red arrows (see figure 1b).

Composing Math. The final input space involves simple mathematics: adding two integers. The math space projects structure into the blend during multiple moments of the activity. We detail the completion and elaboration of these input spaces in the sections below.

Episode 1: *Completing the blend of narrative, game board, and sensory experience*

The activity begins with Marissa and Researcher 1 standing at the start of the game board. After Marissa draws a “force of 2” card, she takes two steps forward and pauses at the second square. Marissa’s small steps are a somewhat trivial completion of the very complicated blend that has been collaboratively constructed. She has blended together several of the available resources to compute the number of squares she is supposed to move on the game board. The number of squares moved in turn is used to represent the constant speed she would be

traveling. The narrative space of the board game—the game piece, board, dice, movement along a track, and event cards—offer the conceptual framework that structures movement of the body (see Figure 2). The carpet space offers the elements of Marissa’s whole body, a white rectangular paper, squares on the paper, bits of paper, and cards with information about force numbers. Marissa, the instructors, and other students experience the fusion of the narrative and carpet space. The blend is now publicly available for others to comment on, elaborate, or re-mix. In this episode, Marissa and Researcher 1 discuss Marissa’s speed after she lands on the second square, which contains a symbol for 1 force.

Researcher 1: Well, what did you start with? (pointing at Marissa)
 Marissa: (Turning her shoulders to look back at the start square) Two...three
 Researcher 1: So you’re going two and then you’re going three because...
 Marissa: (Turning her shoulders again to look back at the start square) I st—I had two.
 Researcher 1: (Pointing to the start square) You had two (and then pointing to the second square) and then
 you landed on a...
 Marissa: Three
 Researcher 1: (Leaning in to take a closer look at the second square). A three?
 Marissa: A one.

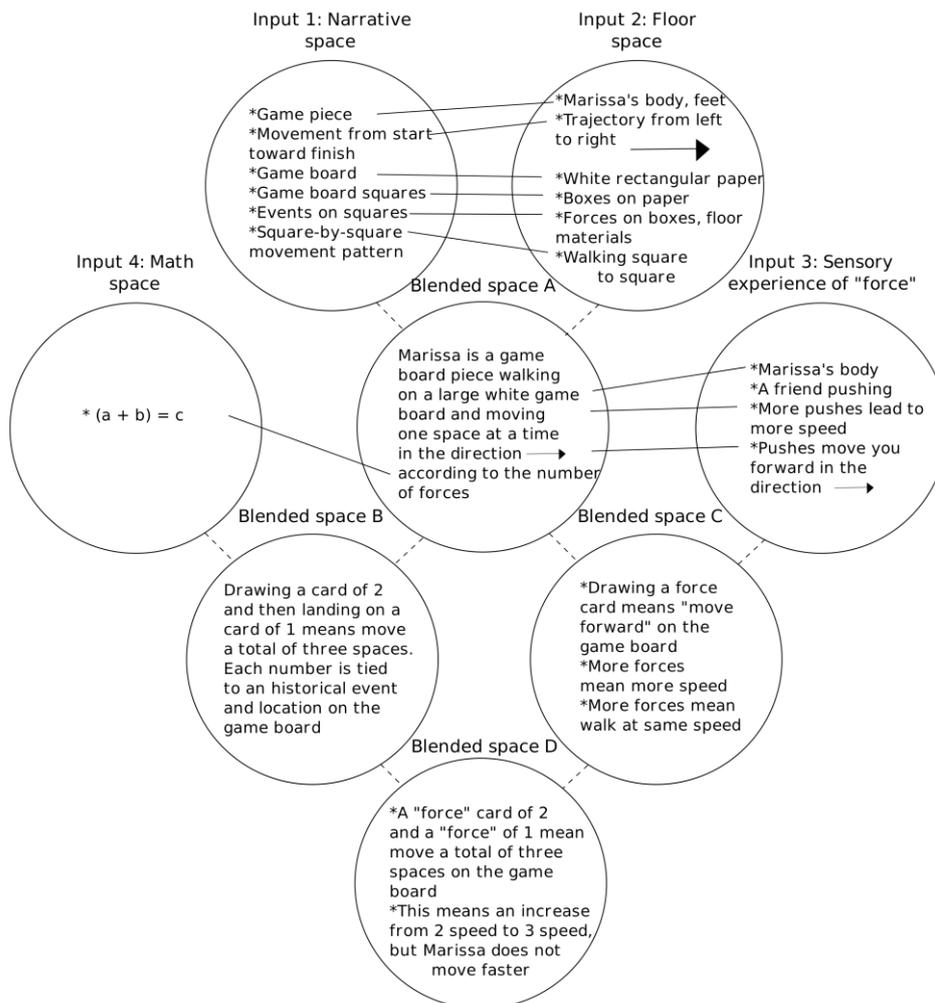


Figure 2. A string of blends that combine the narrative space, floor space, math space, and sensory memory space to produce a numerical representation of speed.

Researcher 1 and Marissa’s discussion of speed involves mathematics rooted both in the physical resources in the room and in the narrative structure of the game board. Marissa has a chance to provide a description of her speed within the context of the blended narrative and floor spaces. The math input space $(a + b)$ becomes an additional tool to evaluate the events in the narrative and the floor blend. There is a 2-force symbol that advanced Marissa from the first square and there is a 1-force symbol on Marissa’s current square. In the blend,

Marissa can combine these two moments in the journey—the initial 2-force and the 1-force—to tally the total forces accrued.

Importantly, the numerical total represents units of force tied historically to specific events in the narrative, both conceptually and physically. In the same way, Fauconnier & Turner (1998) note that, “In the blend, but not in the original inputs, it is possible for an element to be simultaneously a number and a geometric point” (p. 147). Marissa’s reasoning, in this context, incorporates integers, forces, historical moments in the game, and specific spaces on the game board. Speed, in turn, is construed in terms of the (history of the) game board narrative and in terms of the physical semiotic structures of the game board. The concept of speed becomes housed in numbers and in locations on the game board, not in the actual speed with which Marissa moves her body between squares.

Episode 2: *Elaborating the blend to reason about friction*

After landing on the force square in the previous episode, Marissa prepares to advance three squares, where she will land on the linoleum slab used to represent a low friction surface. She walks slowly from one square to the next, and when she steps on the linoleum, Marissa, who is wearing socks, slips slightly forward with her right foot. Researcher 1 initiates a dialogue with a question about what will happen next:

Marissa: I slip? (followed by three exaggerated motions swiveling on the linoleum back and forth in socks)

Researcher 1: Oh, we have an interesting situation...

Marissa: I’m SLIPPING! (co-timed with feet intentionally sliding on the linoleum; arms raise up slightly)

Researcher 1: Marissa is going speed 3, and then she landed on the linoleum. So she says she might slip. What’s that going to do to your speed?

Marissa: Make it faster.

...

Marissa: Because, because, if there’s a 3, and I’m going very fast (steps back one square and faces forward), I would land on this and I would slide (walks forward and slides her feet forward in a controlled way on the linoleum; then returns to standing on the linoleum tile), because it’s slippery.

In this episode, the experience of placing feet on actual linoleum causes the blend to be remixed and the computation to produce an unexpected answer. Marissa’s initial slip, and her memories of slipping on linoleum in socks (an event she later describes as “freaky” and “scary”), leads her to the conclusion that her speed will increase. This inference emerges from an interaction between blends that draw on different source inputs.

The game board blend described in the prior episode fused the number of forces with the number of game board squares. In the blend, greater forces instantiate as greater numbers of squares traveled on the board. Speed, similarly, is represented as the number of squares Marissa can traverse given the forces on the game board. Marissa, then, despite receiving a strong force of three, can slowly walk from square to square; the kinesthetic response to force is never projected into the blend. However, as Marissa walks slowly from one square to the next in the current slipping episode, she steps on the linoleum and encounters a new input to the blend: the kinesthetic experience of actual slipping. The slip happens fast relative to the deliberate, slow pace of walking. Marissa simulates her prior slip several times on the linoleum square and cries out in jest, “I’m slipping!” From a blending perspective, Marissa integrates two embodied representations of speed—the slip on the linoleum and the slow walk—despite that one of those representations was an incidental representation of speed (see Figure 3). The walk between squares represented a default walking speed that never increased with increases in force. That is, even though Marissa had increased her “numerical” speed on previous turns, she never walked any faster. The distance traveled on the game board represents speed in the blend, not how fast the body moves. The result is that Marissa construes the linoleum as increasing her speed even though the increase is relative to the red herring *walking* experience of speed. The numerical depictions of speed are selectively projected into the blend as independent entities. After Marissa draws the conclusion that she will speed up in an embodied sense, she decides to bump up the numerical representation of speed from 3 to 4. This episode reveals that blends can be used to produce both normative and non-normative inferences depending on how the emergent structure is elaborated.

Episode 3: Comparing the computer’s blend to Marissa’s blend

At multiple points throughout the activity, Researcher 1, Researcher 2, and Marissa establish a mapping between Marissa’s journey through the floor space game board and the ball’s journey through the live feed space projected on the white board. Researcher 1 notes early on that the cardboard symbols in floor space are “for the computer” and will appear as symbols in the live feed space. Researcher 2 both hands Marissa the flat cardboard square for the ball, asking “Marissa, do you want to hold the ball while you walk?” and asks Marissa,

“Can you bring me the ball?” upon which Marissa brings over the cardboard square. The ball, in other words, becomes synonymous with the cardboard square symbol and also takes the same journey as Marissa, albeit seen from an overhead view on the classroom wall instead of on top of the white paper on the carpet.

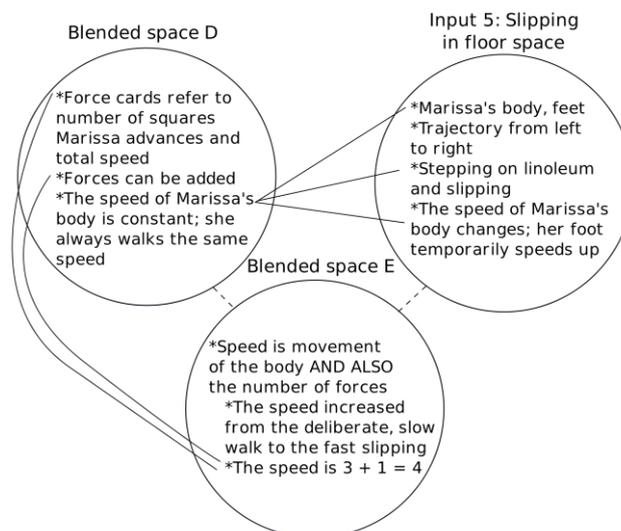


Figure 3. Blending the quantitative representation of forces and the physical experience of slipping.

As shown above, interaction and collaboration is used to establish a direct and public blend between Marissa, the narrative journey, and the image of the ball. In the blend, cardboard and arrow depictions of forces move Marissa and the image of the ball. The participants work to align the elements in the floor space, live feed space, and symbol space according to the narrative structure of the board game. With this blend firmly established, Researcher 2 organizes a comparison between Marissa's journey and the computer's depiction of the ball's journey:

Researcher 2: Let's try to see if the computer agrees with her (Marissa's) prediction.

...

Researcher 2: So the question is, when we run this, is it going to speed up or is it going to slow down when the ball hits the linoleum, right? (moves the cursor in the live feed space to point to the linoleum square). So, Marissa, you said, when the ball get's here, it's gonna get faster, right?

Marissa: Where?

Researcher 2: Right here (moving the mouse up and down)

Teacher: Look at the screen, Marissa.

Marissa: Yeah.

If the fusion between Marissa and the ball was implicit before, the mapping now becomes public and explicit. Researcher 2 refers to “Marissa's prediction” of what happens “when the ball get's here,” while pointing with the cursor to the live feed space. Marissa's early movements with her own body on carpet space are collectively realized as a prediction of how the computer will show the ball moving in live feed space. Marissa, at first, does not realize that Researcher 2 is pointing toward the live feed space. Up until this point, the journey had been extremely focused on the carpet space; cardboard symbols were merely “for the computer.” Now, the spaces have become fully integrated, and Marissa quickly agrees that her earlier embodied prediction corresponds with how the ball will interact with linoleum as determined by the computer.

Despite that the inputs to the computer blend remain completely hidden—there is no mention of how the computer generates the simulation—Marissa is strongly impacted by the computer's prediction. The computer shows the ball rolling across the game board in the live feed space and then slowing down at the linoleum (the opposite of her earlier prediction). Marissa, after agreeing that the ball did slow down on the linoleum, maps the experience “back to the input spaces” (Fauconnier & Turner, 1998) of her earlier movement. She introduces a caveat to her earlier prediction: “If I go on this (walking to stand on the linoleum square), I could slip (acting out the slipping with her right foot) and then I would fall and then it would make me go slower because I would slip.” Marissa introduces the event of falling on the linoleum—which would slow her down—in order to match the computer's prediction of the ball's journey across the game board. Marissa and the ball have been fused to such an extent that the computer's prediction invites Marissa to backtrack and revise her own prediction. Importantly, she revises her prediction by adding the event of falling rather than changing her inferences about linoleum friction.

Discussion

In these three episodes, we see mathematics and physics rooted in a game board narrative, a physical game board, bodies, and augmented-reality symbols. Toward the end of the activity, the computer simulates the normative model of *the ball encountering friction* using the representations Marissa had already put into action, which leads Marissa to revise the description of *how her own body encounters friction*. The augmented reality activity establishes a liminal world blend between Marissa and the ball that allows for a dialogue between Marissa's first-hand experiences and classical physics simulations. Importantly, the computer receives high epistemic credibility as a source of how balls move on linoleum. This finding begs for the study of interactions between social others (e.g. teachers and peers) and the cognitive spaces that people blend to produce inferences.

The liminal blend allows continuity between past and present sensory experiences and the ball's classical response to force and friction. Once the ball and Marissa become coupled in their trajectory through the game board, Marissa comes to believe that the events that the ball encounters according to the computer in the live feed space need to match how she moves through the floor space. The blend simulating the journey of Marissa/ball call for Marissa to look back at the inputs to her own blend and think about her experience in new ways. However, this integration does not happen in a vacuum. The kinesthetic experiences are read into a narrative and into semiotic infrastructure that creates two contrasting roles for the body. Is the body enacting the movement of the game board player or an interaction with the physical surface? Is speed the mathematical total of forces or how the body responds to walking and slipping? The blend combines these inputs, making predictions based on the resources in this environment problematic. Conceptual blending, in this way, shows how resources gather meaning against the ground of other resources, and how accounts of learning need to consider integration across these resources.

References

- Coulson, S. & Oakley, T. (2000). Blending basics. *Cognitive Linguistics*, 11(3/4), 175-196.
- Dudis, P. G. (2004). Body partitioning and real-space blends. *Cognitive Linguistics*, 15(2), 223- 238.
- Enyedy, N., Danish, J. A., Delacruz, G., & Kumar, M. (2012). Learning physics through play in an augmented reality environment. *International Journal of Computer Supported Collaborative Learning* 7 (3) DOI: 10.1007/s11412-012-9150-3.
- Fauconnier, G. (1985). *Mental spaces. Aspects of meaning construction in natural language*. Cambridge, MA: MIT Press-Bradford Books.
- Fauconnier, G. & Turner, M. (1998). Conceptual integration networks. *Cognitive Science*, 22, 133-187.
- Greeno, J., & Hall, R. (1997). Practicing representation: Learning with and about representational forms. *Phi Delta Kappan*, 78, 361-367.
- Hutchins, E. (2003). Cognitive ethnography. Plenary address for the 25th annual meeting of the Cognitive Science Society, Boston.
- Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics*, 37(10), 1555–1577.
- Klopfer, E. (2008) *Augmented Learning: Research and Design of Mobile Educational Games*. MIT Press, Cambridge, MA.
- Liddell, S. K. (1998). Grounded blends, gestures, and conceptual shifts. *Cognitive Linguistics*, 9, 283–314.
- Ochs, E., Gonzalez, P., Jacoby, S. (1996). When I come down, I'm in a domain state: Grammar and graphic representation in the interpretive activity of physics. In E. Ochs, E. A. Schegloff, & S. Thompson, (Eds.), *Interaction and grammar* (pp. 328-369). Cambridge, UK: Cambridge University Press.
- Parill, F., & Sweetser, E. (2004). What we mean by meaning: Conceptual integration in gesture analysis and transcription. *Gesture*, 4(2), 197-219.
- Sidnell, J. (2011). The epistemics of make-believe. In T. Stivers and J. Steensig (Eds.), *The Morality of knowledge in conversation* (pp. 131-156). Cambridge, UK: Cambridge University Press.
- Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, 8(1), 3-19.

Understanding Collaborative Practices in the Scratch Online Community: Patterns of Participation among Youth Designers

Deborah A. Fields, Michael Giang*, & Yasmin B. Kafai†

Utah State University, 2830 Old Main Hill, Logan, UT 84321

*Mount St. Mary's College, 12001 Chalod Rd., Los Angeles, CA 90049

†University of Pennsylvania, 3700 Walnut Street, Philadelphia, PA 19104

Email: deborah.fields@usu.edu, mgiang@msmc.la.edu, kafai@upenn.edu

Abstract: Most research in massive online youth communities has focused on understanding patterns of participation and collaboration in games, social networks, and virtual worlds. Few studies have examined the nature and dynamics in amateur design communities where youth contribute content they have designed themselves. In this paper, we examine quantitative trends of participation in a youth design site focused on programming. Scratch is an online community with over 1 million registered youth designers 11-18 years of age. Drawing on a random sample of 5,000 youth programmers and their activities over three months in early 2012, we examined log files that captured the frequency of their contributions and comments on the site, making visible distinct classes of users who engaged in different sets of practices that support design on a collective scale. In the discussion we discuss implications for the design of collaborative spaces, tools, and communities.

Introduction

A growing body of research in massive online communities has sought to understand patterns of participation through collaboration in online sites, games, social networking sites, and virtual worlds. Research into these sites has provided insights into how people develop collaborations in massive communities within and beyond the designed structures, for instance developing fluid social networks for information gathering and gameplay (Williams, Contractor, Poolec, Srivastad, & Cale, 2011), building trust in long-term relationships that promote more effective teamwork (Chen, 2012), and engaging in knowledge sharing and problem solving in game forums (Steinkuehler & Duncan, 2009). Our own studies illustrated how younger users (twens) drew on many social resources to learn secret commands in a virtual world, navigating relationships in-person as well as with familiar and unfamiliar people in the online community (Fields & Kafai, 2009).

Increasingly however, interest is growing in online communities where users contribute the main content through collaborative or “cooperative” work (Benkler, 2006). In such communities, often organized by community members rather than companies, the content is generated by members themselves. Research into these communities begins to reveal the motivation behind such volunteer collaboration as well as mechanisms for holding together both small and massive collaborative work. Finding fellow collaborators to work on a project and hold that collaboration together in an amateur design site can be a challenge. Indeed, some studies have noted in these cases that most groups fail (e.g., Luther, Caine, Zigler & Bruckman, 2010). Communication by leaders both to coordinate work (Luther et al, 2010) and to maintain socio-emotionally sustaining personal and social discussions (Aragon, Poon, Monroy-Hernandez, & Aragon, 2009) is key in nurturing collaborative design work. This importance of high levels of communication has also been noted in larger scale studies of native “in the wild” collaborations. For instance, Benkler (2006) noted the importance of recognition and communication by key leaders in sustaining vast unpaid volunteer contributions in Wikipedia and Linux. Yet it is unclear what other site-wide practices and design structures might support collaborative learning in amateur design sites. In addition, studies are rare for *youth* amateur design communities that are the focus of this paper. Understanding youth amateur design communities writ large can allow us to make more informed design decisions on how to sponsor collaborative learning at a collective level as well as which users may need scaffolds in participating in collective design communities.

In this paper, we tackle two challenges related to understanding collaboration in massive websites: (1) to understand *site-wide* group dynamics and behaviors that depict and promote collaboration at a collective level and (2) to study a *youth* programmer community. We examine broad trends of participation in the Scratch online community (<http://scratch.mit.edu>), with over 1 million registered youth predominantly aged 11-18 years who share creative programming projects. Drawing on a random sample of 5,000 active designers and their activities over a three-months time period in early 2012 we examined log files that captured the frequency of different kinds of contributions and comments on the site and addressed the following research questions: What are the profiles of users on Scratch.mit.edu and what is their participation over time? How do users engage in the social and/or creative aspects of Scratch.mit.edu? Are there users who engage in one versus another? What role does gender play in users' participation in the website? In the discussion we review our approach to profile analysis and outline implications for the design and study of collaborative online spaces and tools.

Background

Prior CSCL work has focused on understanding various dimensions of smaller group work including different group arrangements (e.g., Engelmann & Hess, 2010), scaffolds for promoting group work (e.g., van der Pol, Admiraal, & Simons, 2006), and interactions between online and offline collaborations (e.g., Birchfield & Megowan-Romanowicz, 2009). With some exceptions (e.g., Fields & Kafai, 2009; Rick & Guzdial, 2006), there is one assumption about collaboration underpinning many of these efforts, which is the idea that collaboration happens in small groups, often of dyads and triads, as they engage in computer-supported collaborative tasks. Research is now starting to examine collaboration in *collective* levels found in massive websites, as can be found in amateur design communities that are the focus of this paper. This may involve the smaller enterprises of individual collaboratives who work together on shared projects as well as the broader dynamics of participation in amateur design communities. Although there are growing numbers of such communities where youth share art (e.g., Deviant Art, Bitstrips), mods of games (e.g., Little Big Planet, the Sims), or stories (e.g., Fanfiction.net, Storybird), we don't know much about who is participating in these productions or who engages in which aspects of or combinations co-designing, sharing and commenting.

We do know from observational studies that a number of collaborative practices such as small group collaborative design and remixing, broadly dispersed constructive feedback, and social incentives for design have sprung up on amateur design sites. Co-design in user-created small groups is closest to the type of small group collaboration often studied in CSCL. In Scratch, many groups of kids or "companies" gather together to work on creating games (Aragon, Poon, Monroy-Hernández, & Aragon, 2009) or interactive stories (Brennan, Valverde, Prempeh, Roque & Chung, 2011). This is similar to the collaboration amongst small groups (collabs) of adult Flash video designers in Newgrounds studied by Luther and colleagues, but as mentioned earlier, most collabs fail to produce a final product (e.g., Luther et al, 2010). Another collaborative practice is remixing, where users download designs made by others, edit them, then re-post. Remixing plays a double role as a collaborative practice. Users can learn by studying and editing others' projects and they also create social links through the traces left by remixing one another's projects, sometimes creating networks of thousands of remixes from just one generative project (Monroy-Hernandez, 2012). Finally, social feedback is type of collaborative practice that supports design in online design communities. These can take the forms of peer reviews on fanfiction sites (Black, 2008), constructive criticism as well as "flames" with negative feedback (Brennan, 2011). Successful small scale efforts have been made to educate select website members to leave more positive, constructive criticism in targeted design challenges (e.g. Roque, Kafai & Fields, 2012) but without measures for how widespread certain practices are (i.e., who leaves comments and who does not) it is difficult to measure change with any certainty.

In developing a framework for parsing computer supported collaborative learning at multiple scales, we need to analytically bring together different practices that support collaborative design on a collective scale, from creating to remixing to commenting, and investigate who engages in these practices, in what combinations of activity, and for what duration. This means that we need to focus not only on the artifacts of networked collaborations but also on the "networking residues" meaning the traces left on projects or profiles such as "love-its," friend requests, "favorites," "likes" and even gifts are types of that show that users have viewed and appreciate projects (Grimes & Fields, 2012). Networking residues can even become a type of commodity as they elevate the virtual presence of a person or project through signs of popularity. In Scratch, members leverage networking residues to support user-created design contests, offering projects, illustrations, love-its, and friending as prizes (Nickerson & Monroy-Hernandez, 2011). Yet while we see evidence of a range of collaborative practices supporting design in amateur online communities, we cannot judge how widespread or distributed these practices are across a full range of users on these sites, nor what patterns of activities users take up over time. To understand connections between practices of creating and sharing that traditionally have been seen as the cornerstone of collaborative design together with collective practices that create the underlying social fabric that encourages and supports continuing and iterative design practices, we examine participation patterns through log files collaborative activities (such as designing, remixing, commenting) of a random sample of users in the Scratch online community, thus complementing prior case studies of individuals, groups of designers (collabs), or common practices of activities (e.g., remixing).

Context & Methods

Scratch.mit.edu is an online massive community where participants, mostly youth ages 11-18 years share their computer programs (Resnick et al., 2009). Kids who share an interest in programming post animations, games, stories, science simulations, and the interactive art they have made in the visual programming environment of Scratch. Launched in May 2007 out of the MIT Media Lab, the Scratch site has grown to more than 1.2 million registered members with nearly 1500 Scratch projects uploaded everyday. As a type of social networking site, activity centers around sharing user-created projects. User profiles are portfolio based, showing individuals' created projects, "favorite" projects, and links to user-created galleries (collections) of projects and recent "friends" on their home page. While there are small spaces for a thumbnail picture and city/country information,

projects dominate the user profile: one gets to know others through the quality of their projects or the comments they leave. Networking residues show up in comments, inclusion in someone's "favorites," and descriptive statistics listed under a project, including the number of views, taggers, "love-its," remixes, downloads, and the user-curated galleries in which the project is located. Projects with more views, comments, and love-its may eventually make it to the front page of Scratch through categories like "Featured Projects," "What the Community is Loving," "What the Community is Viewing," and other sections. The front page is a prized area for Scratchers; having one's project on the front page (or linked from the front page) means getting more views, more feedback, and more visibility. Yet even though the Scratch site is primarily project-based, project creation and social networking are deeply intertwined and the site allows for a number of forms of participation.

Data Collection and Analyses

To understand the group dynamics and behaviors of Scratch users, we used latent class analysis (LCA) to identify communities of similar Scratch users based on their participation behavior (for more details, see Muthen & Muthen, 2000). LCA identifies the maximum number of latent classes (groups of similar individuals) based on a set of observable categorical and/or continuous variables that can be observable online activities. This process can uncover different patterns of activity in various "classes" of users in addition to casual users, social users, and hard-core users that have been identified with more traditional cluster analyses (Giang, Kafai, Fields & Searle, 2012). To do so, LCA relies on model fitting statistics and theoretical interpretation of each class to identify the optimal number of latent classes. This approach avoids the risk of identifying classes with only a few users or a class that is generally similar to another except for minor statistical differences in specific observed activity. The second advantage of LCA is its ability to create unique profiles for each latent class. When dichotomous categorical data, each latent class contains a probability of answering 'yes' to questions about participation. For example, Scratch users can choose to participate by posting a bulletin board comment. The probability of answering 'yes' to this might be 80% for the first class, 20% for another, and 10% for the third class. The first class may consist of social users, while other classes may be different types of casual users who do not socialize.

We used LCA to examine the patterns of relations amongst a set of six Scratch participation variables: 1) *Remixing*: downloading, editing and reposting a project that someone else originally posted. 2) *Downloading*: clicking on a project to download it. This is indicative of looking into the inner workings (i.e., programming) of a project, since all projects can be played online without downloading them. 3) *Commenting*: leaving a comment on a project or a set of projects (a gallery). 4) *Favorites*: clicking "favorite" on a project. Favored projects show up on the profile of the user who clicked "favorite," meaning that others users can see an individuals' "favorite" projects by others. 5) *Love-its*: clicking "love-it" on a project, which leaves a heart on the project. 6) *Friend Request*: sending a friend request to a user. Friend requests on Scratch are unidirectional – they do not have to be reciprocated. Friends' latest projects show up on a users' homepage, making it a way to keep track of favorite designers.

Participants included a random sample of 5004 users from amongst more than 20,000 users who logged into Scratch during the month of January 2012. This sample reflected the broader population on Scratch in regard to self-reported gender and age. As there are no definitive indicators for the correct number of latent classes, both statistical and substantive criteria were used to identify the best model fit for each wave of analyses – one for each of three months (January, February, and March 2012). However, prior to conducting LCA, we had to transform the continuous variables as they were highly skewed, with many participants engaging in no or very few instances of a practice and a few participants engaging in hundreds of instances of a practice. For instance, 4101 users left no comments, 163 users left one comment, 104 users left 2 comments, but six individuals left more than 1000 comments. As a solution, we dichotomized each variable to indicate no activity (0) or activity (1).

Findings

This section reveals different aspects of participation in collaborative practices: (1) the impact of project creation in participation, (2) variability in participation over time, and (3) differences in participation between user groups.

Project Creation Influences Participation

The first discovery we made was that making a project was a gateway for other forms of visible participation on the website. Previous studies have found that Scratch project creation and commenting are not equally distributed amongst the users. Only about 29% of Scratch site participants, primarily male users, share projects. Of these, about half contribute only to a single project (Grimes & Fields, 2012). Some Scratchers prefer activities like commenting, live role-playing or forum posting over project creation (see Brennan, 2011). Our analyses revealed that creating at least one project in a given month was a gateway to all other activities. For instance, in the month of January, there were no users who posted comments who did not create at least one

project, whereas there were many users who created projects but did not post comments. This meant that users who did not create projects also did not participate in any other activities (social or otherwise) represented by the variables, making that group difficult to study with the data available. From our sample of 5004 users, 1379 created an original project in one of the three months (January – March 2012), 533 created a project in each of two months, and 313 created a project in all three months. Thus, 2225 users (67% boys, 33% girls) who created at least 1 project across a three-month period formed the new sample from which all further analyses reported in this paper are drawn. This sample represents about 44.5% of the initial random sample of users.

Participation Varies over Time

We then examined participation in the other collaborative practices (remixing, downloading, commenting, favorites, love-its, and friend requests) for each of three months, January, February, and March. To unpack these rather complex analyses, we take first a look at the findings for January (see Figure 1) that resulted in a 5-class model.

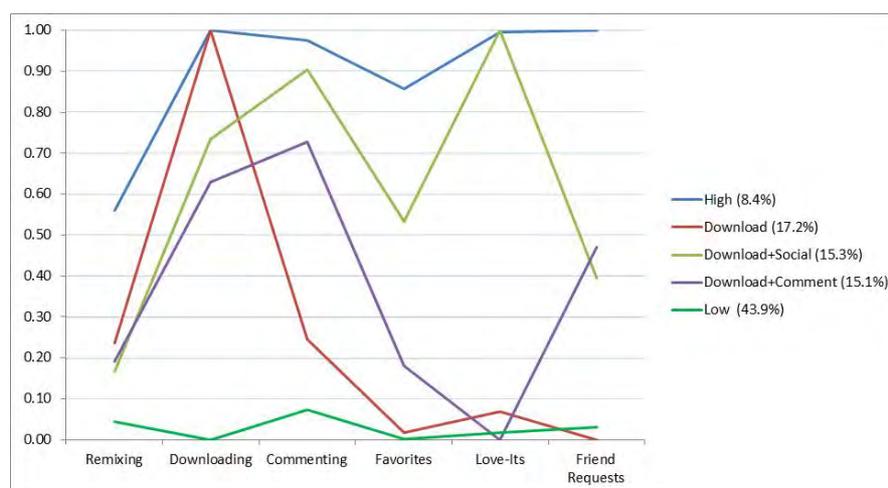


Figure 1: The 5-Class Model for January and the percentage of users for each class.

Here we see a “*high class*” profile of users (8.4% of all users) who are likely to participate in nearly every type activity studied. They have a 55% chance of posting a remix, a 100% chance of downloading a project, and very high (above 85%) chances of commenting on a project, favoriting a project, and 100% chance of loving a project and making a friend request. It is worth noting that the existence of this high class profile is consistent across all three months, as we will show later. A second profile we nickname the “*download class*” (17.2%) because there is a 100% chance that they will download projects from the Scratch site. There is also a profile of users who are likely to participate in most of the social networking activities available on Scratch a “*social + download class*” (16.3%), including commenting, favoriting, loving, and less likely friending. They are also very likely to download projects (75%). A fourth profile is the “*download + comment class*” (16.1%) who has a strong likelihood of participating in downloading project, commenting and projects, and friending, with low likelihoods of favoriting or loving projects. The fifth class is the “*low-level class*” (class 5, 43.9%) who are unlikely to do anything except post a project during the month.

The high class in particular is interesting, because it suggests that the core users on the Scratch site are the ones most likely to participate in remixing projects and in commenting, favoriting, loving, and issuing friend requests. In addition, the next class of users (download+social) most likely to participate in the social activities on the website are also very likely to download projects (75% likelihood). Indeed, against our expectations, there was no class of individuals who were likely to participate in commenting, favoriting, loving, or friending others without also having a strong likelihood of downloading projects, an activity which suggests that kids are not just playing projects but investigating and looking into them. In other words, besides posting a project, downloading a project is a second gatekeeper to social activity on the Scratch site. Although leaving networking residues such as favorites, love-its, and friend requests would seem to have the lowest bars for participating in a social networking website, this actually appears to be a practice in which only those who are most involved in a full range of practices on the site participate.

Another one of our key findings is that rather than having the same class model for each month, the statistical analyses suggested a different class model for each month (for details of the LCA analysis, see Appendix A). In the models for February and March we see the continuation of a high class of users who were very likely to participate in all forms of activities we studied, from remixing to friending. The percentage of users in this class stayed steady from January (8.4%) to February (8.6%) and increased into March (11.6%),

perhaps because of the decrease in number of classes in the model. A class of downloaders+commenters also continued across all three months and remained relatively consistent (from 15.1% in January, 12.5% in February, and 11.4% in March), and the low class of users who were only likely to post one project also remained present (increasing from 43.9% to 70.1% to 77.1% in each subsequent month). This supports the idea that clicking love-it or favorite on projects or sending friend requests are high-end activities among a relatively small group of users compared to practices like downloading projects and commenting on projects. It also suggests that users who were active in one month (our sample was drawn from active users) are likely to drop off in activity in subsequent months.

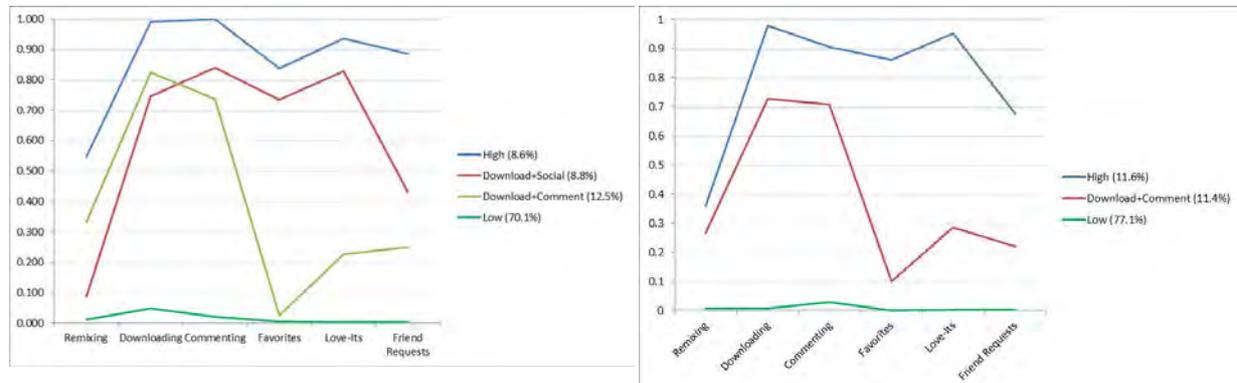


Figure 2. Four Class Model for February (left) and Three Class Model for March (right).

Absence of Gender Differences in Participation

Girls only represent one-third of all registered members on the Scratch site. The gender distribution in our overall sample reflected the distribution of self-reported gender on the Scratch site: 67% male, 33% female. Given this prior knowledge about differential membership in Scratch community, we tested whether gender was proportionately represented in each of the latent class. The distribution of gender within each class model was generally insignificant with only two exceptions: a higher proportion of girls in the high class in the month of January and a higher proportion of boys in the download+comment class in the month of March (for more detail on statistical analyses, see Appendix A). These analyses suggest that while males dominate the population of Scratch at large, within class profiles gender differences are minimal, an interesting finding for a youth amateur design site focused on programming.

Discussion

This paper examined broad quantitative trends of participation in a youth amateur design site focused on programming, making visible sets of practices that support design on a collective scale. Prior qualitative studies have documented that many of the activities on the Scratch site are collaborative in nature and support youths' programming designs through social supports (love-its, favorites, friend requests, comments) and constructive criticism as well as through opening up youths' designs to each other through the opportunity to play, download, and remix others' projects (e.g. Brennan, 2011; Roque et al, 2012; Aragon et al, 2009; Kafai et al, 2012). Before this study we had little idea of what kinds of participation patterns users exhibited in the variety of collaborative design practices common on Scratch. Our findings suggest that there are several classes of users in Scratch and that making a project and downloading others' projects are gateways for other more prominently social activities. This is a surprising finding given our own prior assumptions that clicking "love-it" or "favorite" on a project were the lowest bars of activity. Indeed only a small percentage of the active Scratch population engaged in those activities: those already engaged in creating, downloading, and commenting on projects. Instead, active Scratch users prioritize designing projects and downloading others' designs, a finding that suggests studying social networking forums focused on design may be qualitatively different from the social networking sites discussed more prominently in popular media (e.g. Facebook, MySpace) that center on relationships.

Our results suggest several future steps for deeper analyses. First, the unavailability of data such as views of projects, home page, and notifications make the activities of a large proportion of users (55%) in the study invisible. The development of Scratch 2.0 is designed to capture these kinds of data, which should help illuminate the activities of this hidden group of participants. In addition, we need better ways of documenting the relative richness and sophistication of projects and comments at a quantitative scale. This will allow us to differentiate between types of project creators and comment posters, for instance between users who post many relatively simple projects or users who share one or two highly sophisticated or complex projects. Further, although the current class models of participation we presented here are static, in future analyses we will investigate which users transition between classes in each month's model, seeing what proportion of high-class

(and other) users remain in that class or move to other classes. We also plan to investigate the role of experience in Scratch in seeing which users move between classes as well as analyzing whether other users' activities (clicking love-it on a project, commenting on a project, etc.) have an effect on users moving between classes. These analyses would provide richer information on who is participating in Scratch and what kind of users' activities may influence the likelihood of their changing their involvement on the site. Establishing models of participation in Scratch was an important precursor to these other analyses.

The larger goal of this research is to illuminate collaborative practices in massive communities that support learning and design, to see who is participating in those activities, and to evaluate how to sustain those types of activities. In the end we want to engineer online sites to more productively encourage computer supported collaborative learning on a collective level. Analyses presented in this paper are a first step toward a broader understanding of who participates in which activities over time in Scratch. Future interventions at a website-design level and at various local levels will be able to more intentionally target certain classes of users to involve them in more collaborative kinds of work. Some models for this already exist albeit in a small scale, for instance supporting highly involved designers to learn to post more positive, constructive criticism or helping other users connect to each other for collaborative co-design through semi-structured design challenges (see Roque, Kafai & Fields, 2012; Kafai, Fields, Roque, Burke, & Monroy-Hernandez, 2012). Yet these types of interventions reach out primarily to the "high class" of users who already engage in most opportunities on the site. Analyses in this paper suggest opportunities to target other classes of users, for instance reaching out to those in a "downloading" class to engage them in one additional collaborative practice such as commenting. The importance of design to this community also suggests that interventions should be targeted on youths' designs, perhaps by drawing attention to new users' designs who might benefit from feedback or making designs more easily searchable so that users can connect with others who have common interests. More broadly, further research is needed on site-wide analyses of other youth amateur design sites to see if similar trends prevail.

Appendix A

The LCA analysis resulted in a 5-class in January, a 4-class in February, and a 3-class model in March based on the indicators for each model fit (Hagenaars & McCutcheon, 2002; Muthen, 2002). For each participant, LCA generates probabilities for membership into each class, and generally one class has the highest probability of members. For instance, results show the students classified into their highest probability latent class had a 76-94% probability of being in that class, and a 0-24% of being in the other classes. Among the latter, 2 of the 20 classification were above .19. Taken together, these results support a 5-class model. LCA results for February indicate that a 4-class model provides the best fit for the data (see Figure 2). This is supported by a non-significant LMR p-value at the 5-class model ($p = .08$). In addition, aBIC value dipped to its lowest point at the 4-class model. Further, the substantive interpretation of the four model provided greater clarity. Additional statistical support can be provided by the average most likely latent class membership probability. For the 4-class model, students classified into their highest probability latent class had a 75-97% probability for being in that class, and a 0 to 21% of being in the other class (with only 2 of the 12 other class probabilities being above 10%). LCA results for March indicate that a 3-class model provides the best fit for the data (see Figure 2). This mainly supported by lowest BIC and aBIC values for that class. In addition, the most likely class probabilities ranged from 85.5% to 98.0%, and 0 to 8.7% for the other classes. Interpretation of the 3-class model provided a better, more parsimonious interpretation of the data. A 4 class model was substantively rejected because one class consisted of a small number of users (1.4%) and it was only distinguished to another class by 1 of the 6 indicator items. Due to space constraints we have not put in the model fit indices table as we did for January.

Table 1: Model Fit Indices for January.

DICH	January: N = 2225 (Includes project creators for any of the 3 months)						
	likelihood	free par	BIC	aBIC	LMR p-value	Entropy	AIC
1	-6456.55	6	12957.97	12938.91	N/A	N/A	12925.10
2	-5527.54	13	11152.31	11111.01	0.0000	0.828	11081.09
3	-5477.94	20	11105.44	11105.44	0.0176	0.681	10995.87
4	-5448.24	27	11098.41	11012.63	0.0000	0.818	10950.48
5	-5426.66	34	11107.59	10999.58	0.0000	0.804	10921.31
6	-5416.54	41	11139.71	11009.46	0.0466	0.773	10915.08

Note. The Lo-Mendel-Rubin (LMR) likelihood ratio tests whether the current model is an improvement over the model with one less class. For instance, the p-values indicate the 5-class model provides the best fit. The non-significant p-value at the sixth class (.1129) suggests that it is not an improvement over a 5-class model. In addition, information criteria (i.e., Akaike

Information Criterion (AIC), Bayesian Information Criterion (BIC), sample size Adjusted BIC (aBIC)) were used to compare models, wherein models with additional model parameters were penalized in the search for the most parsimonious model and information criteria with the lowest values indicate the best model fit. As values begin to level off (especially if there values do not increase), substantive interpretation and selection criteria take a greater role. For this analysis, the lowest value is at the 5-class model, wherein BIC and aBIC values begin to increase. In addition, a 5-class model provides a more meaningful and distinct interpretation of the class than models with fewer or great classes.

For assessing the distribution of gender, we performed multiple chi-square tests for independence analyses. For January, 2 (gender) x 5 (latent classes) chi-square analyses initially revealed that gender was distributed differently across the 5 latent classes, $\chi^2(4) = 9.635$, $p = .047$. However, upon close inspection of standardized residual scores comparing difference between the observed and expected, only 1 marginally significant difference emerged ($z > 1.96$). That is, there were more women in the class 1 (high class) than expected, $z = 2.030$. For February, chi-square test for independence suggest that each latent classes had a similar proportion of boys and girls, $\chi^2(3) = 5.613$, $p = .132$. For March, chi-square test for independence initially revealed that gender was distributed differently across the 3 latent classes, $\chi^2(2) = 10.040$, $p = .007$. However, standardized residual comparisons revealed showed only one significant difference within Class 1 (download+comment), $z = -2.152$; that is, there were fewer girls found in this group than expected.

References

- Aragon, C., Poon, S., Monroy-Hernandez, A., & Aragon, D. (2009). A tale of two online communities: Fostering collaboration and creativity in scientists and children. *In Proceedings of the Creativity and Cognition Conference* (pp.9–18). Berkeley, CA. New York: ACM Press.
- Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. New Haven and London: Yale University Press.
- Birchfield, D. & Megowan-Romanowicz, C. (2009) Earth science learning in SMALLab: A design experiment for mixed reality. *International Journal of Computer Supported Collaborative Learning*, 4 (4), 403-421.
- Black, R.W. (2008). *Adolescents and online fiction*. New York: Peter Lang.
- Brennan, K. (2011). Mind the gap: Differences between the aspirational and the actual in an online community of learners. *Proceedings of the International Conference of Computer Supported Collaborative Learning*.
- Chen, M. (2012). *Leet noobs: The life and death of an expert player group in World of Warcraft*. New York: Peter Lang.
- Engelmann, T. & Hesse, F.W. (2010) How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *ijcscl* 5 (3), pp. 299-320.
- Fields, D. A. & Kafai, Y. B. (2009). A connective ethnography of peer knowledge sharing and diffusion in a tween virtual world. *International Journal of Computer Supported Collaborative Learning*, 4(1), 47-68.
- Giang, M. T., Kafai, Y. B., Fields, D. A., & Searle, K. A. (2012). Social interactions in virtual worlds: Patterns and participation of tween relationship play. In J.Fromme & A. Unger (Eds.), *Computer games and new media cultures: A handbook on the state and perspectives of digital games studies*. New York: Springer Verlag. 543-556.
- Grimes, S. & Fields, D. (2012). *Kids online: A new research agenda for understanding social networking forums*. New York. The Joan Ganz Cooney Center at Sesame Workshop. Available online at <http://www.joanganzcooneycenter.org/reports-38.html>.
- Hagenaars J, & McCutcheon A. (Eds). (2002). *Applied Latent Class Analysis*. Cambridge, UK: Cambridge University Press.
- 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.
- Monroy-Hernandez, A. (2012). Designing for remixing: Supporting an online community of amateur creators. *Unpublished dissertation*. MIT.
- Muthen B. 2002. Statistical and substantive checking in growth mixture modeling. Retrieved January 2007 from http://www.gseis.ucla.edu/faculty/muthen/full_paper_list.htm.
- Muthen B, Muthen L. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcohol Clinical Experimental Research*, 24 (6), 882–891.
- Nickerson, J.V., Monroy-Hernández, A. (2011). Appropriation and Creativity: User-initiated Contests in Scratch. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS '11)*.

- 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.
- Rick, J. & Guzdial, M. (2006). Situating CoWeb: A scholarship of application. *International Journal of Computer-Supported Collaborative Learning*, 1(1), 89-115.
- 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.
- Steinkuehler, S., & Duncan, S. (2009). Scientific habits of mind in virtual worlds. *Journal of Science Education and Technology*, 17(6), 530–543.
- van der Pol, J., Admiraal, W. & Simons, P. R. J. (2006). The affordance of anchored discussion for the collaborative processing of academic texts. *International Journal of Computer-Supported Collaborative Learning*, 1(3), 339-357.
- Williams, D., Contractor, N., Poolec, M. S., Srivastad, J., & Cale, D. (2011). The Virtual Worlds Exploratorium: Using Large-Scale Data and Computational Techniques for Communication Research. *Communication Methods and Measures*, 5(2), 163-180.

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, St. Mary's College, or the University of Pennsylvania. Special thanks to Anant Seethalakshmi for help with gathering data and to the Scratch Team for providing feedback on analysis.

Incentives in educational games: A multilevel analysis of their impact on elementary students' engagement and learning

Michael Filsecker, Duisburg-Essen University, Forthausweg 2, 47058 Duisburg, Germany,
michael.filsecker@uni-due.de

Daniel Thomas Hickey, Indiana University, 201 North Rose Avenue, Suite 4000
Bloomington, Indiana 47405-1006, dthickey@indiana.edu

Abstract: The effects of incentives on engagement and learning were analyzed at multiple levels in an immersive videogame for elementary science. One group of fifth-graders was offered incentives and another group was not offered incentives. The feedback afforded by the videogame was expected to mitigate predicted negative effects of incentives. No significant motivational effects of incentives were found across engagement levels: immediate (engagement with resources), close (participation in drafting in-game reports), proximal (self-reported situational interest) or distal (gains in self-reported personal interest). Nearly all of the differences that were found favored the incentive condition. Students in the incentive condition showed significantly larger gains in conceptual understanding (proximal) and non-significantly larger gains in achievement (distal). These results suggest that the predicted negative consequences of extrinsic incentives may be addressed or even reversed in this new generation of learning environments, and point to value for a multi-level model of assessment and engagement.

While most commercial videogames offer players some form of *incentives* (such as points or “levels”) to motivate their progress, incentives remain controversial in education. Cognitive theorists assume that incentives undermine intrinsic motivation and subsequent engagement via the *overjustification* effect (Deci, Ryan, & Koestner, 2001, Lepper, Greene, & Nisbett, 1973). This occurs when an extrinsic incentive is introduced for activity which was previously intrinsically interesting. After the introduction of the incentive (e.g., a prize or a certificate) the individual subsequently attributes the basis for the activity to the extrinsic reward. Hundreds of studies have shown that “extrinsic” incentives direct attention away from intrinsically motivated learning, leading to diminished engagement once incentives are no longer offered (Tang & Hall, 1995). Reflecting the antithetical relationship between cognitive and behavioral theories of motivation, analyses of the same body of studies by behaviorally-oriented theorists support the conclusion that the negative consequences of incentives are limited to specific easily-avoided situations (Cameron & Pierce, 1994).

Sociocultural Perspectives on Incentives

Newer sociocultural theories of knowing and learning offer a different way of thinking about incentives and motivation that might move this debate forward. In their groundbreaking paper on *cognitive apprenticeship*, Collins, Brown and Newman (1989) suggested that the corrosive educational effects of competition (which is typically fostered by incentives) may be more the results of impoverished learning environments that lacked opportunities to improve and the formative feedback needed to do so. Most of the prior studies of incentives were conducted in highly structured laboratory settings or very traditional classrooms. This suggests that the newest generation of educational videogame incentives might have positive consequences that outweigh or even eliminate any negative consequences. Furthermore, the rich interactive narratives in the latest generation of immersive videogames and the participatory culture of many networked learning environments might counter or even reverse the overjustification effect via what Gresalfi, et al. (2009) called *consequential engagement*.

The meaning of educational engagement is bound to views of learning. Prior scholars have advanced notions such as *mindfulness* (Salomon & Globerson, 1987), *intentional learning* (Bereiter & Scardamalia, 1989) and *committed learning* (diSessa, 2000). As Dewey put it a century ago “...the educational significance of effort, its value for an educative growth, resides in its connection with a stimulation of greater *thoughtfulness* not in the greater strain it imposes” (Dewey, 1913, p. 58). Sociocultural approaches highlight Dewey’s *thoughtfulness* as the process by which students engage in an activity, interact with each other and use resources and tools purposefully. Engel and Conant’s (2002) notion of *productive disciplinary engagement* highlights (a) the number of students making substantive disciplinary contributions, (b) the number of disciplinary contributions made in coordination with each other, (c) students attending to each other and making emotional displays, and (d) students spontaneously reengaging. In this characterization, the role of discourse is key to supporting any claim concerning engagement.

Multi-Level Assessment Model

The model in this study emerged in prior design-based research involving *GenScope* program for learning genetics. The first cycle analyzed and fostered learning at three levels, including the *immediate-level* enactment of the GenScope activities, *close-level* informal assessments and “feedback conversations,” and a *proximal-level* performance assessment (Hickey, Kindfield, Horwitz, & Christie, 2003). The second cycle added a fourth *distal-level* external test that documented significant achievement gains on targeted standards without resorting to expository instruction (Hickey & Zuiker, 2012). Subsequently in the context of immersive games, this study explored the usefulness of the model for providing valid inferences of the translation of the intense engagement with videogames to academic subject matter (Roschelle, Kaput, & Stroup, 2000). The difficulty of such translations lies, in part, on the unique affordances of educational games (i.e., formative feedback and numerous low-stakes opportunities to improve). While the formative assessment functions of these features enhance learning, they can compromise evidential validity of assessments used to examine engagement and learning in videogames. This study assumes that doing so calls for assessments along different “levels” of learning outcomes (Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002). The current study assumes that using different learning outcomes across levels means that formative feedback at one level does not directly coach or prepare students for the outcomes at the next level. This provides a tractable way of controlling for the construct-irrelevant variance (Messick, 1994) that occurs when students are given feedback for solving problems that are similar to the problems that appear on an assessment (Hickey & Anderson, 2007; Hickey et al., 2006). This maximizes consequential validity (i.e., the formative function of assessment) at one level while preserving evidential validity at the next level (the summative function). Doing so across three or more levels promises to overcome the complexities of assessing learning outcomes from educational games (e.g., the concerns over assessment sensitivity raised by Annetta, Minogue, Holmes, & Cheng, 2009, p. 79).

This study extended the multi-level assessment model as it had emerged in design studies of Quest Atlantis’s *Taiga* ecology game (Barab, et al., 2011) to the study of incentives and their impact on engagement. Learning was conceptualized in terms of the four levels shown in Table 1. Generally speaking, these levels were pragmatically informed by the three “grand theories” of learning outlined in Greeno, Collins, & Resnick (1996). First, a situative/sociocultural perspective was used to conceptualize (1) the immediate-level enactment of sequences of inquiry-oriented game activities and (2) close-level participation among the player, teacher, and non-player characters in writing and revising written “quests” after those activities. The model then uses a cognitive/rationalist perspective to frame learning in terms of (3) proximal-level conceptual understanding assessed with a curriculum-oriented performance assessment. Finally, the model uses a more behavioral/associationist perspective to frame learning in terms of (4) distal-level achievement measured with a multiple-choice test. This means that the collected evidence of close, proximal, and distal learning (a) were increasingly removed from the enactment of the Taiga inquiry activities, (b) were increasingly oriented towards a broader curricular scope, and (c) used increasingly abstract representations of the targeted knowledge.

Table 1: Multilevel assessment model

LEVEL (Orientation)	Assessment Format	Learning Outcome	Relationship to Curriculum	Feedback Timescale	Primary Summative Functions	Primary Formative Functions
IMMEDIATE (Action/procedures)	Analysis of log-files, “live” discourse and social interaction	Enactment of actions and procedures, communal discourse	Same content and context	Minutes	Actual enactment of QA activities	Foster discourse and intentional learning
CLOSE (Activity)	Analysis of content of quest submissions	Interactive discourse & intentional learning	Same content and context	Hours-Days	Enactment of preceding QA activities	Foster individual understanding
PROXIMAL (Curriculum)	Open-ended problem solving assessment	Individual understanding of targeted concepts	Same content in similar context	Weeks-Months	Understanding of concepts targeted	Refine curriculum and compare versions
DISTAL (Standards)	Externally-	Aggregated	Same &	Months-	Measure	Inform broad

	developed test items	achievement of targeted standards	different content in new contexts	Years	impact on achievement	audience of curricular impact
--	-------------------------	---	--	-------	--------------------------	-------------------------------------

Each level of analysis has potential summative and formative functions. For example, the close-level analysis of the questing activity has a summative function relative to the game activities but has a formative function relative to the understandings that individuals take away from that interactive writing. Aligning learning across levels reveals the presence or absence of “echoes” across levels. This distinguishes the actual consequences of design features from random variation. For example, when marginally significant distal outcomes are correlated with larger, statistically significant proximal outcomes, designers should be less inclined to dismiss the distal outcomes as having occurred by chance (see also Schaffer & Serlin, 2004). This is useful when attempting to ensure that refinements to specific curricula are consistently impacting learning on distal measures, even when working with small numbers of learners typical of early-stage design studies.

This study attempted to extend the multi-level assessment design model to the issue of incentives by building on emerging situative/participatory approaches to motivation (Greeno et al., 1998; Hickey, 2003). At the close level, we examined students’ written quests as evidence of their success while participating in the interactive practice of drafting a quest. While this method lacked the attention to context or a more interpretive discourse analysis, it seemed the most appropriate way of capturing participation at this level in a way that could be interpreted in the quasi-experimental comparison. At the proximal level, we examined individual players’ self-reported motivational states during that same quest. At the distal level, we examined players’ more enduring self-reported personal interest towards the kinds of problems they were solving in the game. This relationship between assessment and motivation are explored in more detail in Hickey & Schaffer (2006).

Methods

This study was the third in a series of annual design studies of the 15-hour Taiga curriculum with the same teacher and population of students. In the previous year, new formative feedback resources (e.g., teacher rubrics for reviewing and giving students feedback) and new cut-scenes with useful information were added to help students complete the crucial second quest. These new resources substantially increased gains in understanding and achievement, but only for those students who accessed them (Hickey, Ingram-Goble, & Jameson, 2009). This suggested that using strategies to motivate students to access more of the resources and do so more meaningfully should further enhance learning outcomes. Incentives seem to be a promising strategy.

A quasi-experimental design was conducted to examine the effect of providing incentives on students’ engagement and learning science. For two of the classrooms in this study, the teacher’s acceptance of a written quest at one of three increasingly accomplished levels (proficient, expert, or wise) was rewarded with a corresponding badge that players could affix to their in-game virtual avatar (Figure 1a). Additionally, students in this Public Recognition (PR) condition were invited to move a paper version of their avatar up and across a physical “leader board” that was prominently placed in the room (Figure 1b). In two other classrooms taught by the same teacher in the same semester, students in the Non Public Recognition (NPR) condition were not offered badges or a ready means to communicate their level of progress to the other students and in-game information on incentives was replaced by messages encouraging players to work hard to save the park and become more capable apprentices (Lepper & Malone, 1987). The study tested the following hypotheses:

Hypothesis 1: Students in the PR condition will engage more deeply in the process of drafting and revising their quests, use more relevant scientific formalisms, and use those formalisms more correctly than students in the NPR condition.

Hypothesis 2: Students in the PR condition will exhibit significantly larger gains in conceptual understanding of the targeted science concepts and achievement of the targeted science standards than students in the NPR condition.

Hypothesis 3: There will be no difference between the PR and NPR conditions in self-reported intrinsic motivation during the second quest, and no differences in impact of the game on personal interest in learning to solve these types of scientific problems.

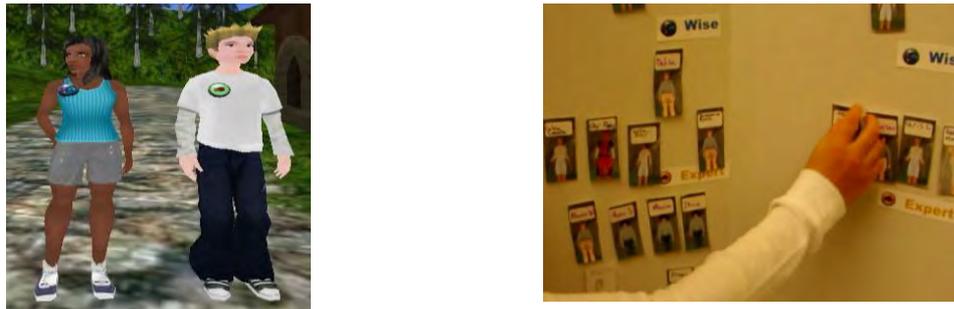


Figure 1a and 1b: Avatar Badges and Leader Board Displaying Levels.

Participants and Materials

This research was conducted at a public elementary school in a medium-sized city in the Midwestern US. As is typical of university communities, the students were predominantly Euro American and most came from well-educated professional families. In this study, average grades from prior work were used to identify pairs of similar achieving classes, and one class in each pair was assigned to the Public Recognition (PR) and the Non Public Recognition (NPR) condition. Consent to participate in the study was obtained from almost every student, resulting in 106 participants (56 females and 60 males).

Instrumentation and Procedures

Learning and engagement were assessed simultaneously at the immediate and close levels and separately at the proximal and distal levels. At the immediate level we analyzed the number of screens of formative feedback that students accessed by accessing the log files generated during gameplay as in the previous studies (Hickey et al., 2009). This reflected our tentative assumption that choosing more pages represented more intentional engagement in the structured discourse of the revision process. To assess learning and engagement at the close level, we analyzed the quality of the initial and final submissions of crucial Quest 2 (scored by researchers) using a 14-point scale rubric which assigned six points for summarizing the water quality indicators, four points for explaining what the processes were (i.e., erosion and eutrophication), and four points for describing the dynamic relationship between indicators and processes.

While this rubric could capture the students' right or wrong answers to Quest 2, it could not capture the meaningful appropriation of concepts in the domain discourse. For example, one student could say *dirt from Site B got into the river*, while another one could say *the sediment from Site B is eroded into the river*. By using the 14-points rubric, both students would have earned one point, without distinguishing the nuances such as the difference between *dirt* and *sediment* and between *got into* and *eroded*. In a sense we were aiming at the *disciplinary* engagement pointed out by Engle and Conant (2002). Therefore, we quantified the verbal data (Chi, 1997) to capture this domain-specific or *disciplinary* discourse around students' Quest 2 submissions (n=106). Initial and final submissions in Quest 2 were coded in terms of the meaningful appropriation of nine relevant scientific concepts. The text of the submissions of all students (n=106) was coded using the NVivo qualitative analysis software program. We were interested in capturing students' engagement with the content in a progressive knowledgeable way as a result of the incentive manipulation, instead of students' actual representation of knowledge (e.g., Chi, 1997), scientific argumentation (e.g., Kelly, Drucker & Chen, 1998; McNeil et al., 2007) or knowledge construction (e.g., Weinberg & Fischer, 2006).

Table 2: Examples of a few scientific concepts coded as accurate, non-accurate or other.

Scientific concepts	Examples of evidence for coding		
	Accurately	Non-accurately	Other
Dissolved Oxygen	Warm temperature takes out the DO in the water so the fish suffocate... (Student ID 118412)	The Temperature is effected by Do level in the river (Student ID 118517)	There is too much of every thing except DO and it is way to hot. (Student ID 118306)
Turbidity	Turbidity is caused some by erosion. (Student ID 118221)	without turbidity the sun will get through the water and then the plants can't grow. (Student ID 118509)	At site A and C, the Turbidity is in between. (Student ID 118504)

Table 2 provides examples of actual students' answers (misspellings in original responses) and the category in which they were coded. The meaningful appropriation of the concepts has to do with (1) identifying the right level of the indicators displayed in the charts, (2) the concept being used to establish a valid relationship with other concepts, (3) relating the concept to a relevant activity or event, and (4) identifying the concept as being the cause/effect of another concept or associated activity or event. A non-meaningful appropriation situation has to do with (1) establishing the concepts in an invalid relationship with one another (2) the concept being used to explain the wrong ecological process (e.g., erosion or eutrophication), or (3) the concept being used incorrectly as a cause or effect of an event or another concept. The category "other" was used when the student's response was too ambiguous to discern the appropriate category.

To examine engagement at the proximal level, we developed a scale to assess players' situational motivation regarding the Quest 2 activity. The scale consisted of 4 or 5 Likert-type items (strongly disagree, disagree, neutral, agree, or strongly agree) for each of the following subscales of the motivational states that prior research has shown to be diminished by incentives: *interest* in the activity, *value* for completing the activity, *perceived competence* during the activity, and *effort* completing the activity. So long as the individual scores for each set of items are internally reliable, scores on each scale are presumed to be indicative of various aspects of students' cognitive engagement during the tasks (see Fredricks, Blumenfeld, & Paris, 2004). Once their Quest 2 submission was accepted, students completed the brief survey. The survey asked students, "How did you feel while completing Quest 2?" The survey also encouraged students to respond honestly and assured students that their responses were confidential.

To examine engagement at the distal level, we measured changes in personal interest in solving the types of problems students were learning to solve in Taiga. One of the main concerns with incentives is that they may supplant existing intrinsic motivation towards activities with the extrinsic motivation offered by the incentive - the "overjustification effect" (Lepper et al., 1973). Hundreds of prior studies in laboratories or traditional classrooms showed that extrinsic incentives lead to decreased free choice engagement in the incentivized activity. Many of those studies also examined self-reported interest in the activities (and sometimes instead of) free choice engagement. To this end, we measured students' self-reported personal interest in the three types of problems that they were learning to solve in Taiga: *water ecology* problems, complex scientific problems, and controversial socio-scientific problems. An 18-item survey was created consisting of six Likert-scale items for each type of problem and was administered before and after students played the game.

To examine learning gains at the proximal level, we used the *Lee River* performance assessment developed in the prior design cycles. The assessment was "curriculum-oriented" in that it asked students to solve similar problems as in Taiga but in a somewhat different context. The assessment had been created alongside extensive refinements to Taiga the previous year and was designed to be highly sensitive to different enactments of the curriculum. It involved another fictional watershed and a range of stakeholders who had similar (but not identical) effects on the ecosystem. For example, both Taiga and Lee River involve stakeholders with different land use practices who are arranged along a river. The stakeholders from both scenarios impact their ecosystems by doing things that cause erosion and eutrophication; however, erosion is caused by loggers in Taiga and by construction in Lee River. To capture a range of understanding at the pretest and the posttest, the items covered included a broad range of difficulty. It included several multi-part items that started out with simple tasks that most students would be able to answer without instruction, and proceeded to a few complex items that focused on the nuances of scientific hypotheses, the relationship between social issues and scientific inquiry, and the relationship between water quality indicators such as dissolved oxygen and processes like eutrophication. A 21-point scoring rubric was used to score completed assessments, with a subset of assessments scored by two researchers to establish reliability.

To examine learning gains at the distal level we used the same 20-item test that had been created the previous year by random sampling from pools of items aligned to the four targeted content standards, but independent of the Taiga curriculum. Such standards-oriented tests are necessary to support claims of impact on externally-developed achievement measures and to compare the impact of different curricula that target those standards. Such tests are not particularly sensitive to specific interventions and represent a relatively ambitious target for innovative curricula like Taiga.

Results and Findings

For engagement and learning at the immediate level, analysis of the log files found no significant difference in the number of feedback pages accessed for the PR ($M=8.69$, $SD=6.91$) and the NPR ($M=9.24$, $SD=5.98$) conditions [Mann-Whitney $U=1285$, $n_1=51$, $n_2=55$, $p=.452$]. At the close level improvement scores for the initial and final Quest 2 submissions (using a 14-point scale; inter-rater reliability = .85) did not reach statistical significance between conditions [Mann-Whitney $U=1099$, $n_1=47$, $n_2=51$, $p=.475$]. In addition, a one-way MANOVA was conducted to compare the effects between conditions on the meaningful appropriation of the scientific concepts as enlisted during the drafting of Quest 2. The analysis of the coded initial and final submissions revealed higher levels in the PR condition, but the difference did not reach statistical significance

[Wilks' Lambda = .973, $F(1,102)=2.797$, $p=.097$] Therefore, strictly speaking, we found no evidence of negative consequences of incentives engagement in the written discourse around Quest 2.

Concerning proximal engagement, all four self-reported assessments of situational motivation during Quest 2 revealed high internal reliability (all alphas over .85). This was crucial, since unreliable measures could have masked consequences of incentives in random variance. A one-way between subjects ANOVA was conducted to compare the effects of the incentive and non-incentive conditions on perceived interest, value, competence, and effort. There was no significant effect on any of the variables [$F(1,106)=.826$, $p=.366$; $F(1,106)=.051$, $p=.821$; $F(1,106)=.467$, $p=.496$; $F(1,106)=.321$, $p=.575$, respectively]. While none of the four differences reached statistical significance, the fact that slightly higher scores were observed for all four of the aspects in the PR condition argues strongly against the predicted negative consequences from the incentives.

For distal engagement the scales of interest in solving the three different types of problems showed acceptable levels of reliability (alphas over .80) at both administrations. A one-way repeated measures ANOVA was conducted to compare the effects of incentives on three indices of interest. None of the tests yielded significant difference between conditions [Wilks' Lambda = .99, $F(1,102)=.442$, $p=.508$; Wilks' Lambda = .99, $F(1,101)=.703$, $p=.404$; Wilks' Lambda = .99, $F(1,101)=1.026$, $p=.314$], supporting our initial hypothesis that the "overjustification" is unlikely to occur in contexts such as QA. These results suggest that the introduction of incentives in this environment did not undermine personal interest (or presumably subsequent free-choice engagement) in these times of scientific investigations.

For proximal learning, a one-way repeated measures ANOVA tested the effects of incentives on students' gains in conceptual understanding. Students in the PR condition gained significantly higher levels of understanding than students in the NPR condition [Wilks' Lambda = .946, $F(1,99)=5.6$, $p=.02$]. As shown in Figure 2, this represented the difference between 1.4 and 1.1 SD gain, given the pooled standard deviations across the score points. Importantly, the differences in gains between the two classes in each condition were not statistically significant ($F < 1$). Thus, the students in the incentive condition developed significantly greater understandings of the concepts, topics and processes associated with solving scientific and socio-scientific problems involving water quality.

For distal learning, the achievement tests revealed strong internal consistency, and showed that students in the PR condition gained 5.44 points compared to 4.02 points for the other students. Given the variance within the scores, this was a difference between gains of 1.1 and 0.8 SD. A one-way repeated measures ANOVA revealed that this difference in gains did not reach conventional criteria for statistical significance [Wilks' Lambda = .972, $F(1,114)=3.234$, $p=.075$, gains between classes within groups was again $F < 1$]. However, such a gain seems highly unlikely to have occurred by chance given the corresponding significant difference in gains in proximal understanding. This is an example of the aforementioned "echo" and an illustration of the advantage of assessing learning outcomes across multiple increasingly formal levels.

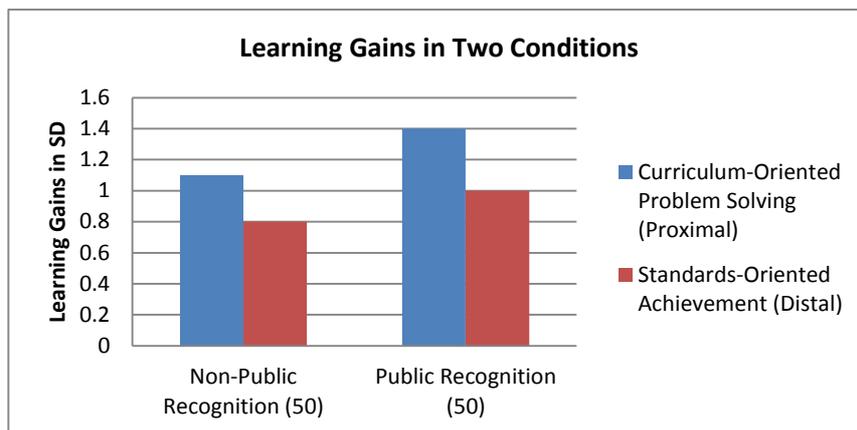


Figure 2: Proximal and Distal Learning Outcomes by Implementations and Conditions (* = $p < .05$; *** = $p < .01$).

In summary, the incentives as enacted in this study were not shown to motivate students' engagement with the learning activities such as drafting and revising Quest 2 and using the resources embedded in the game. Therefore, Hypothesis 1 was not supported. However, results showed a significant larger gain in understanding of ecological processes (proximal), and a non-significant differential gain in achievement (distal) both in favor of the PR group. Therefore, Hypothesis 2 was partially supported. Finally, examination of engagement at the three levels failed to uncover any of the potential negative consequences of the incentives, supporting our third hypothesis.

Implications and Significance

These findings lend initial support to the argument advanced by Collins, Brown, and Newman (1989) that the negative consequences of competition may be more indicative of impoverished learning environments and the lack of feedback and opportunity to improve, than of a fundamental consequence of competition. Likewise, the study provides some initial empirical support for the speculations about sociocultural theories of engagement in Hickey (2003) and Hickey & Shaffer (2006). Rather than (a) using incentives haphazardly or (b) attempting to prove their fundamental impact, we believe that designers should ask about the motivational design features concerning their impact on immediate-level and close-level engagement in learning. While there are likely many ways of doing so, we believe that this more process-oriented and contextual analysis offers a helpful starting point. We also believe that this study shows some initial value in extending the multi-level model of assessment used in past studies to consider engagement and motivation as well.

Arguably, the multilevel assessment model applied in this study begins to address a core validity issue that has long plagued the assessment of individually-oriented motivational interventions (see Adelman & Taylor, 1994). Just as with our learning outcomes, our formative efforts to refine engagement at one level do not undermine the evidential validity of the engagement outcomes at the next level. In other words, there was nothing about close-level motivational intervention (i.e., incentives and competition) that might have directly encouraged students to characterize that activity as more interesting or engaging on the proximal-level survey items. At the same time, we indirectly examined the consequences of incentives and completion on student's self-reported cognition during those activities and in changes in self-reported interest towards those activities. This seems like a promising way around the obvious dilemma facing many motivational interventions: programs that focus directly on changing behavior may deliver behavioral change, but fail to impact cognition, while programs that focus directly on cognition may indeed impact cognition but fail to deliver enduring changes in behavior. Likewise, the model represents an extension and may well complement current analytical strategies based on discourse and video analysis (e.g., Azevedo, 2006; Engel & Conant, 2002) by introducing performance and achievement levels together with self-reported motivational states. In summary, while protecting the validity of outcome claims, the model also emphasizes the assessment of the process of engagement and learning encompassing the "hybrid research methodologies" characteristic of multidisciplinary fields such as CSCL (Stahl, Koschmann, & Suthers, 2006). Thus, the model provides a promising solution to the assessment of learning beyond sociocultural perspectives on teaching and learning.

Finally, the increased learning outcomes across the three design cycles demonstrates the broader value of this assessment driven multi-level model of assessment. While the present study focused on the impact of incentives, numerous other refinements had been made to the Taiga curriculum that were informed by evidence obtained at the various levels. Of course, some (but certainly not all) of the increased gains were due to teachers learning. Most innovators who have attempted to impact valid measures of external achievement know how difficult it is to obtain gains of this magnitude without resorting to expository direct instruction. In addition to offering a way forward on enduring design controversies like incentives, the multi-level model appears to be a promising way to deliver the evidence of achievement impact that is demanded by many educational stakeholders without compromising the more authentic learning supported by innovations like Quest Atlantis.

References

- Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M. T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education*, *53*(1), 74–85.
- Azevedo, F. S. (2006). Personal excursions: Investigating the dynamics of student engagement. *International Journal of Computers for Mathematical Learning*, *11*(1), 57–98.
- Barab, S. A., Gresalfi, M., & Ingram-Goble, A. (2010). Transformational play using games to position person, content, and context. *Educational Researcher*, *39*(7), 525–536.
- Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L. Resnick (Ed.), *Cognition and instruction: Issues and agendas* (pp. 361-379). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*, *64*(3), 363-423.
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: a practical guide. *Journal of the Learning Sciences*, *6*(3), 271–315
- 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: Erlbaum.
- Deci, E. L., Ryan, R. M., & Koestner, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin*, *125*(6), 627.
- Dewey, J. (1913) *Interest and Effort in Education*. Boston, MA: Riverside Press.
- diSessa, A. A. (2000). *Changing Minds: Computers, Learning and Literacy*. Cambridge, MA: MIT Press.

- Engel, 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.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59.
- Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 15-46). New York: Macmillan.
- Gresalfi, M., Barab, S., & Siyahhan, S. (2009). Virtual worlds, conceptual understanding, and me: designing for consequential engagement. *On the Horizon*, 17(1), 21-34.
- Hickey, D. T. (2003). Engaged participation vs. marginal non-participation: A stridently sociocultural model of achievement motivation. *Elementary School Journal*, 103 (4), 401-429.
- Hickey, D.T., Kindfield, A., & Horwitz, P. (2003). Integrating curriculum, instruction, assessment, and evaluation in a technology supported genetics learning environment. *American Educational Research Journal*, 40(2), 495-538.
- Hickey, D. T., Zuiker, S. J., Taasoobshirazi, G., Schafer, N. J., & Michael, M. A. (2006). Three is the magic number: A design-based framework for balancing formative and summative functions of assessment. *Studies in Educational Evaluation*, 32, 180-201.
- Hickey, D. T., & Anderson, K. (2007). Situative approaches to student assessment: Contextualizing evidence to transform practice. In P. Moss (Ed.), *Evidence and decision making, the 106th yearbook of the National Society for the Study of Education (Pt. I)* (pp. 264-287). Chicago: National Society for the Study of Education.
- Kelly, G.J., Drucker, S. & Chen, K. (1998). Students' reasoning about electricity: combining performance assessment with argumentation analysis. *International Journal of Science Education*, 20(7) 849-871.
- Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28(1), 129-137.
- Lepper, M. R., & Malone, T. W. (1987). Intrinsic motivation and instructional effectiveness in computer-based education. *Aptitude, learning, and instruction*, 3, 255-286.
- McNeill, K.L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In M. Lovett & P. Shah (Eds.), *Thinking with data: Proceedings of the 33rd Carnegie Symposium on Cognition* (pp. 233-265). New York: Taylor & Francis.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-18.
- 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). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Ruiz-Primo, M. A., Shavelson, R. J., Hamilton, L., & Klein, S. (2002). On the evaluation of systemic science education reform: Searching for instructional sensitivity. *Journal of Research in Science Teaching*, 39(5), 369-393.
- Salomon, G., & Globerson, T. (1987). Skill is not enough: The role of mindfulness in learning and transfer. *International Journal of Educational Research*, 11, 623-637.
- Shaffer, D. W., & Serlin, R. C. (2004). What good are statistics that don't generalize? *Educational Researcher*, 33(9), 14.
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 409-426). Cambridge, UK: Cambridge University Press.
- Tang, S. H., & Hall, V. C. (1995). The overjustification effect: A meta-analysis. *Applied Cognitive Psychology*, 9(5), 365-404.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, 46(1), 71-95.

The Joint Action Theory in Didactics

A case study in videoconferencing at primary school

Gruson Brigitte, Sensevy Gérard, CREAD (Research Centre on Education, Learning and Didactics), IUFM de Bretagne-UBO, CS 54 310 - 153, rue Saint-Malo, 35043 Rennes Cedex, France.
brigitte.gruson@bretagne.iufm.fr, gerard.sensevy@bretagne.iufm.fr

Abstract: In this paper, we aim to show how the Joint Action Theory in Didactics (JATD) may contribute to the construction of a science of instructional practice. In this perspective, we propose two concepts: the student's twofold semiosis and the teacher's equilibration work, which allow us to understand the instructional practice and the learning process in the same conceptualization. We illustrate these concepts and our video research methodology with a case study in videoconferencing at primary school (Second Language Learning). We attempt to clarify how this technology-enhanced learning environment enables a reconstruction of the learning situation in a specific inquiry process, requiring a particular teacher's equilibration work and a specific student's entitlement. We eventually argue that the Joint Action Theory in Didactics could contribute to the development of Learning Sciences research.

Introduction

In this paper, we address three major issues.

First, we concentrate our analysis on practice, and particularly on *instructional practice*. Sharing Koschmann's question "How... do we begin to construct a real science of instructional practice?" (Koschmann, 2011, p. 6), we try to provide some elements to contribute to this construction. In that perspective, we sketch a theory, the Joint Action Theory in Didactics (Sensevy & Mercier, 2007; Sensevy, 2011a; Sensevy, 2011b) that we describe according to its main lines, offering a glimpse into its structure and some of its elements. This theory aims at linking teaching and learning and postulates that one cannot understand learning practices without understanding related teaching practices.

The second issue of our paper concerns the way the teaching/learning relationship can be effectively mediated by technology. In order to work out this issue, we present how a specific technology-enhanced learning environment, videoconferencing, can foster teachers and students' activity. We give some elements of a comparative analysis between an "ordinary class" and a "videoconferencing class", where English is studied as a Second Language at primary school, grounded in the categories of the Joint Action Theory in Didactics previously described.

The first two issues are interrelated in so far as the purpose of this paper is both to emphasize videoconferencing as a method to study and document didactic transactions in relation with JATD.

The third issue of our paper seeks to show that French Didactics Concepts, and particularly the Joint Action Theory in Didactics, could contribute to the Learning Sciences research.

The Joint Action Theory in Didactics: origin and main principle

The Joint Action Theory in Didactics (JATD) stems from French Didactics (in particular, Brousseau, 1997; Chevallard, 2007; Laborde, Perrin-Glorian & Sierpienska, 2005). In this perspective, we will refer to what could be termed the first theoretical principle of French Didactics. This principle states that in order to understand a didactic activity (i.e. an activity where someone teaches and someone learns), one needs to understand a *system*, the *didactic system*, which is a system of three subsystems, the subsystem of Knowledge (the piece of knowledge in question), the subsystem of the Teacher, and the subsystem of the Student. By arguing that the didactic system is an *undividable* system, we emphasize the fact that all theorisation in didactics rests on the fact that one cannot understand the teacher's behaviour without simultaneously understanding the student's behaviour and knowledge structure and function. It suffices to replace the word "teacher" with the words "student" or "knowledge" in the preceding sentence to obtain three assertions that constitute the core meaning of this framework.

The Joint Action Theory in Didactics: an epistemological background

In order to understand the meaning-making process in practices, we have to understand the logic of practice on which people base their behaviors. In our conception, acting according to the logic of a practice is to be able to master a specific language-game in a particular life-form (Wittgenstein, 1997). In order to master this language game, one has to be able to produce and decipher signs of various kinds in an appropriate way. From that perspective, Mead (1934) provides us with a remarkable conception. Indeed, he considers that a *social act* is an act in which certain features of a participant's conduct are taken as "stimuli" by her partners, and reciprocally. The joint action thus rests on what we may call the *semiosis of others*, the deciphering of actions – verbal as

well as bodily – that other persons carry out in a given situation. In a similar way, Dewey argued that “things gain meaning by being used in a shared experience or joint action’ (Dewey, 1916/1985, p. 20)”. We contend that the teaching-learning process could contribute to the paradigm of joint action (Clark, 1996, Sebanz et al., 2006).

The Joint Action Theory in Didactics: some theoretical tools

The fundamental grammar of the didactic game

We describe the didactic transactions (1) (Dewey and Bentley, 1949) between the teacher and the students as a game of a particular kind, a *didactic* game. What are the prominent features of this game? It involves two players, A and B. B wins if, and only if, A wins, but B must not give A the winning strategy directly. B is the teacher (the teaching pole). A is the student (The studying pole). This description allows us to put forward that the didactic game is a collaborative game, a *joint* game, within a joint action. But what kind of joint action? If we look at a didactic game more carefully, we see that B (the teacher), in order to win, has to lead A (the student) to a certain point, a particular “state of knowledge” which enables the student to play the “right moves” in the game, which can persuade the teacher that the student has built the right knowledge. At the core of this process, there is a fundamental condition: in order to be sure that A (The student) has really won, B (The teacher) has to be *reticent*.

Brousseau (1997) coined the “Topaze Effect” idea that emphasizes this peculiar point. In a dictation, concerning the sentence “les moutons...”, the teacher Topaze wants the students to accurately write down the plural of the noun “mouton” (*sheep*), which is marked by a final “s” in French. So the teacher clearly pronounces the “s” (moutonssse) at the end of the word “moutons” (in the normal French pronunciation, this final “s” is not pronounced), and the students automatically write down the “s”. One can argue that the teacher cheated in this didactic game, and that the students did not really win the game of writing down the plural of the noun “moutons” accurately. A fundamental rule of the didactic game has not been followed: the teacher has to be tacit, *reticent*, in order to let the student build *her* proper knowledge. The student must act *proprio motu*, the teacher’s scaffolding must not enable the student to produce the ‘good behaviour’ (i.e. to put an “s” at the end of the word “mouton”) without calling on the adequate knowledge (i.e. the rules of plural agreement). This *proprio motu* clause is necessarily related to the *reticence* of the teacher. We argue that in all kinds of teaching (i.e. direct instruction or inquiry-oriented teaching), the teacher has to be reticent in order to be sure that the *proprio motu* clause is respected, that the students’ behaviors are grounded on actual knowledge, the knowledge involved in the teaching/learning process. Thus, we consider that the *proprio motu* clause and the teacher’s reticence compose the general pattern of didactic transactions and give them their strongly asymmetrical nature.

In order to characterize the teaching-learning process more deeply, we use the notion of Learning Game. A Learning Game is fundamentally a joint game, in that it refers to the teacher’s game on the student’s game as it occurs *in situ*. In the following, we focus on some of the concepts we use to describe a Learning Game and to better understand the instructional practice and the learning process in the same conceptualization.

The didactic contract and the milieu

In order to understand joint action in the didactic game, we have to identify the game *thought style*, which functions as a background to the didactic transactions. Concerning the concept of thought style, we rely on Fleck’s contention (1979) according to which a thought style is viewed as “a readiness for directed perception”. In our framework, these thought-style properties are taken into account in the notion of *didactic contract* (Brousseau, 1997), which is a system of habits between the teacher and the students. These habits entail particular expectations (either from the teacher or the students), with each agent attributing some expectations to the other(s).

A good example of the cognitive strength of the didactic contract can be found in the research paradigm called “the Captain’s age” (Schubauer-Leoni and Perret-Clermont, 1997). Researchers submitted “absurd problems” to students in primary education (i.e. a boat has 3 veils and 42 crew members, how old is the captain?). Surprisingly (at least for the researchers), a great proportion (80–95 %) of the students gave an answer (by using the numbers given in the problem) and only very few of them replied that it was impossible to respond. This example makes us understand what a didactic contract is. Every time they had to solve a problem, students had to use the same set of procedures. First, they had to give an answer. Second, in order to ‘produce’ this answer, they had to use the numbers provided in the particular problem. Third, these numbers had to be ‘associated’ thanks to the last ‘means of calculation’ (addition, or subtraction etc.) they had learnt.

It is important to note that the didactic contract, as a system of expectations, is largely implicit: it functions as a common background fostered in everyday didactic joint action, a thought style attached to the ‘problem solving’ game. In a didactic institution, such an institutional contract has to be modified all along the learning-teaching process.

We can theorize this change by introducing the notion of *milieu* (Brousseau, 1997), which can be described as a system of physical and symbolic objects that is elaborated to constitute an *antagonistic system* to

the previously taught system. In that perspective, the didactic contract can be seen as a way of assimilating, in the Piagetian sense (Piaget, 1975), the didactic experience. In our previous example (the Captain's age), the contract, as background knowledge, 'enables' the students to answer the problem. But when new elements that the joint action cannot directly assimilate are brought into the milieu, they bring a kind of *resistance* to the joint action, which entails that the didactic contract has to be changed. In that case, the assimilation of the new milieu requires the accommodation (*Ibid.*) of the contract.

As can be understood from the previous description, the relationship between contract and milieu holds a prominent position in JATD. Indeed, we argue that in order to characterize the didactic joint action, one has to identify how the students orient themselves, either by enacting the didactic contract habits or by establishing epistemic relations with the milieu. It means that empirical studies have to reveal what kind of dialectics is built between the "contract-driven students' orientations" and "the milieu-driven students' orientations". We use the expression *didactic equilibration* to designate the search for an adequate equilibrium between these two kinds of orientations in the transactional didactic process, and we see the teacher's action as the production of such an equilibration work.

The student's twofold semiosis process and the didactic equilibration

In JATD, the student's activity is mainly thought of as a semiosis process, of two kinds. First, the student has to produce a first semiosis, which corresponds to the deciphering of the signs of the milieu that are non-intentional signs. But this first kind of semiosis is enacted against the background of previously taught knowledge as it has been structured in the preceding didactic joint action, the didactic contract background. However, in the joint action, the teacher may orient the student's action in the milieu in a more or less reticent way. The signs produced by the teacher to orient the student in the milieu refer to what we call the second semiosis. In the *first semiosis*, the student has to decipher the non-intentional signs of the milieu. In the *second semiosis*, she has to decipher the intentional signs that the teacher provides in order to orient her in the milieu.

With respect to the didactic equilibration we refer to, we can therefore determine two main ideal-typical relationships, which one can consider as two poles of a gradient.

In the first typical relationship, *the contract can be seen as an auxiliary to the milieu*. For example, in a reading session, the teacher may want the students to be able to read the text accurately, in a text-centered process, in which the students discover the meanings of the text in a first-hand relationship. In order to foster such an inquiry, the teacher can rely on the didactic contract, but as a kind of orientation through the milieu. In this kind of relationship, the semiosis process is mainly of the first sort. The students have mostly to decipher the non-intentional signs of the milieu.

In the second typical relationship, *the milieu can be seen as an auxiliary to the contract*. For example, in a reading session, the teacher may want the students to read a certain text, but this text, as a milieu, is merely a way of presenting some meanings the teacher intends to make the students recognize. The core of the didactic process does not lie in an inquiry concerning the text, but in a system of meanings that the teacher tries to reinforce within the contract. The most important thing in this learning-teaching process is the teacher's discourse or conduct, to the extent that the text (milieu) could be removed from the transactions without compromising the learning at stake. In a way, the milieu is a pretext. In this kind of relationship, the semiosis process is mainly of the second sort. The students have mainly to decipher the teacher's intentional signs.

An actual didactic process generally mixes these two ideal-typical relationships, according to the characteristics of the knowledge in question. But it is possible to draw a line between contract-oriented transactional systems and milieu-oriented transactional systems, and, in doing that, to acknowledge different strategic systems in the teacher's action.

An outline of the methodological approach: Video Research

As we have seen, our theoretical viewpoint focused on the analysis of the joint action between the teacher and the students gives a major importance to the production and the deciphering of signs, in the semiosis process, which is partly a semiosis of others. The teacher has to identify the signs provided by the students, and, in what we have called the second semiosis, students have to recognize the teacher's intentional signs. Moreover, the knowledge involved is most of the time constituted by a symbolic system the students have to acknowledge and practice. For these reasons, we ground our enquiries in the video recording of teaching/learning practices, and we include our approach (Tiberghien & Sensevy, 2012) in the methodological paradigm of the Video Research (Goldman et al., 2007). As we have argued elsewhere (Tiberghien & Sensevy, 2012), video recordings enable to keep the *analogical* dimension of the situations with their specificities and their infinity of information (Dretske, 1981). Above all, as Goldman & McDermott (2007) state, it makes communication visible, and makes it possible to account for embodied *instructional communication*.

An empirical study: comparing an ordinary situation to a situation involving synchronous exchanges via videoconferencing

The empirical study provided here is based on the comparison of two episodes extracted from two different corpora used to analyze the way English as a Second Language (ESL) is taught and learnt in primary schools in France. The first episode comes from a five-lesson teaching unit implemented in an ordinary class composed of eighteen 5th graders. The second episode comes from a videoconferencing session during which four French students interact with four English students at lunchtime under their teacher's supervision.

In both situations, the students play a familiar game similar to the well-known "who is it" game. This famous game is a two-player game in which each player secretly and simultaneously selects one character and then tries to be the first to guess who her/his opponent picked. In both classes, this game has been turned into a "what monster is it" game as the secret character to uncover is not a person but a monster. For the "monster" game, each teacher has developed specific documents and slightly changed the original rules. However, beyond these differences, the general knowledge involved is very similar in each class. It corresponds to the students' ability to ask and answer questions so as to discover the secret monster, which is the monster chosen by another student.

In the ordinary class, the game is organized as follows: one student - here William - plays against the whole class. It means that William chooses a monster his friends have to identify, asking him questions in English he has to answer in English as well. In this class, the French students have to be able to produce and understand questions and answers in the target language to play the game. In doing so they will respect an overarching didactic contract that strongly characterizes ESL classes at primary level: the "use of the target language" contract (Gruson, 2009), thus situating themselves in the communicative paradigm.

For the videoconferencing session, the French and English students play against each other. In the selected episode, one player, a French student called Emma, asks questions in French to Sophie, an English student, who replies in English. In contrast with the ordinary class, we note that both French and English are being used in the videoconferencing session. Indeed, one vital ingredient of the use of videoconferencing, which transforms teaching and learning a foreign language, is that, during videoconferencing sessions, both French and English students have to learn the target language. Consequently, these sessions are in essence bilingual sessions during which both languages are supposed to be used equally to solve the problem in question – in this case to guess the secret monster. These bilingual sessions are based on a fundamental principle that could be termed the "equal benefit" principle meaning that both partners have to benefit equally from the exchanges (Gruson, 2010). This principle, as we will show, has a strong impact on the teachers' and the students' actions: it not only modifies the didactic milieu but also the type of didactic contracts that occur in the ESL class.

To conclude the presentation of the two situations, we have to keep in mind that, for the selected episodes, the language activities involved in each class are different. In the ordinary class, the students will have to practice both receptive and productive activities in the target language whereas, during the videoconferencing session, the main knowledge in question consists in understanding the target language. If both teachers agree on focusing on a receptive activity in that episode, it is because they want to take advantage of the native speaker students' "expertise" in their own language and "use" them as authentic references for the target language.

What happens *in situ*

In the ordinary class

This episode takes place at the beginning of the game just after William designated Linda to ask him a second question.

Table 1: Transcripts and screen captures, episode 1.

<p><i>Linda: how many nose PT1 (2): how many noses has it got ? + how many noses has it got? Linda: how many noses has it got?</i></p>	<p><i>William: it's one</i></p>	<p><i>PT1: it has got William: it has got one</i></p>	<p><i>William: nose PT1: nose William: nose</i></p>
			

<p>The teacher makes a gesture to encourage Linda to repeat the whole sentence and correct her mistake. PT1's gaze is directed towards Linda (not visible on the photograph).</p>	<p>William looks at the document on the board to check how many noses the monster he has chosen has. He answers with an incorrect sentence limited to the necessary information: the number.</p>	<p>The teacher provides William with the right form, and makes a gesture to encourage William to correct his sentence. PT1's gaze is directed towards William. William repeats the verbal expression and the number.</p>	<p>PT1 touches her nose to encourage William to complete his sentence. PT1's gaze is directed towards William. William produces the word "nose", which PT1 repeats to produce a correct phonological model. William reproduces PT1's pronunciation.</p>
---	--	--	---

The way Linda and William produce their sentences clearly shows that they limit their productions to signifying elements not bothering to utter complete sentences. In doing so, they comply with the milieu, which doesn't make the production of complete sentences necessary. Indeed, in everyday life, youngsters playing the same kind of game would most likely do the same: ask a three word questions "how many noses?" and answer with a number: one. However, we observe that William tries to produce a longer answer "it's one". He does so because he is familiar with the expectations of his teacher whose gestures and reformulations remind her students twice that she wants them to produce long and correct sentences.

In the current learning game, we observe that the teacher's reticence is quite low: PT1 provides her students with the verbal expression "has got", which they repeat although she does not ask them to do so explicitly. In fact, they repeat PT1's words or sentences as the result of the way they interpret the signs she produces. As a consequence, the students do not act *proprio motu*: their oral productions are strongly oriented by PT1's verbal and non-verbal actions; they are the result of the system of habits fostered previously. In that episode, as we explained before, the milieu is not resistant enough: using "has got" is not necessary to play the "what monster is it?" game. As PT1 has given the winning strategy directly, she cannot know if the students have really understood the use of "has got". The milieu being too weak, the didactic contract allows the students to assimilate it.

Consequently, in this first episode, we can say that the milieu is first and foremost an auxiliary to the contract and that the on-going semiosis process is mainly of the second sort. In order to win the learning game, the students have to orient themselves mainly by deciphering the contract signs afforded by the teacher. In that case, we can state that the teacher's equilibration work is contract-driven.

In the class using videoconferencing

This episode takes place in the middle of the session. Emma produces her first question after two other questions have been asked.

Table 2A: Transcripts and screen captures, episode 2.

<p>Emma: est-ce que ton monstre a cinq oreilles ? (<i>Does your monster have five ears?</i>)</p>	<p>PTE: cinq oreilles++ combien ça fait ? cinq oreilles ? (<i>five ears++ how much is it? Five ears?</i>) Sophie: does your monster have five ears?</p>	<p>Sophie: no our monster hasn't got five ears</p>
		
<p>Emma asks the next question. She is looking straight at the camera so as to establish eye contact with her English peers. She pronounces her</p>	<p>PTE repeats the sentence key-elements, "five ears", twice to help her students understand. Sophie touches her ear and then produces</p>	<p>Sophie then looks at the camera to answer Emma's question. As Emma, she makes efforts to articulate and pronounce clearly to</p>

question slowly and clearly.	the question, which she addresses to her teacher to check its validity. Her gaze is directed towards PTE whose knee is visible in the left side corner of the screen.	make sure Emma will understand her answer. Emma is concentrated on the document on her lap.
------------------------------	---	---

Table 2B: Transcripts and screen captures, episode 2.

<p>Emma: j'ai pas compris (<i>I didn't understand</i>) PTF: tu le dis (<i>Say it</i>) Emma: repeat please PTF: je crois qu'ils ont pas bien entendu Emma vas-y (<i>I think they haven't heard Emma go on</i>)</p>		<p>In spite of Sophie's efforts, Emma hasn't understood the answer, which she first tells PTF. She turns towards her teacher who is standing next to her and whispers to him. The latter does not repeat the answer but tells her to ask Sophie to repeat, which she does.</p>
<p>Sophie: no our monster hasn't got five ears PTF: he hasn't got five ears +++ à vous ! (<i>your turn!</i>)</p>		<p>While Sophie repeats the answer for the second time, Emma is looking intently at the screen to make sure she will understand this time.</p>

In this episode, the on-going learning game is quite different. The milieu is much more complex as the students have not only to interpret the signs produced by their teachers but also those produced by their peers in the same room and across the Channel. However, as in the first episode, answering the yes/no question formulated by Emma "Does your monster have five ears?" can be reduced to the production of one isolated word, here, 'yes' or 'no'. Yet, if we examine the answer produced by Sophie, we observe that she utters a long answer "no our monster hasn't got five ears" even if Emma only needs to understand the first word to get the required information and make progress in solving the problem. In a very different way from what happens in the ordinary class, the production of long and correct sentences is not enforced by the teachers' expectations but by the English student's desire to provide the French students with a model on which they will be able to rely when they have to produce their sentences in English.

Concerning repetitions, we can observe that if the question is repeated here it is because Emma asks Sophie to do so as she was unable to pick up the clues (no / hasn't) included in the milieu (the answer) to get the sentence meaning. In videoconferencing sessions, repetitions are usually due to technical problems that often disrupt virtual communication. This can also explain why both teachers, PTE at the beginning and PTF at the end, repeat after the students.

If we look closely at Emma and Sophie, we observe that they really enact a first-hand relationship to the knowledge involved. First, as shown in the second picture, even if Sophie tries to find some support from her teacher to make sure she understood the word "oreille" correctly, she produces it by herself. As for Emma, pictures 1 and 5 clearly illustrate how concentrated she is and willing to decipher the signs of the milieu: the words produced by Sophie. In this second episode, PTF's reticence is high when he encourages Emma to find in her own linguistic resources the word to ask Sophie to repeat. In that case, the assimilation of the new milieu requires the accommodation of previous knowledge: the contract has to be changed.

To conclude, we can first put forward that, in this second episode, the students' actions are mainly oriented by the milieu and by a semiosis process of the first sort. Secondly, we can argue that by providing another kind of didactic game, videoconferencing sessions have a strong impact on the dialectics between the "contract-driven students' orientations" and "the milieu-driven students' orientations". As a consequence, the teachers' equilibration work is very specific. It consists mainly of enabling the student to confront the milieu, within an appropriate guidance, which does not rest on the unveiling of the teachers' epistemic expectations.

Discussion

It seems to us that the case studies we theorize in this paper may lead to some exploratory conclusions and implications.

The need for a theory of instructional practices

We argue that there is a need for a transactional theory that holds together teaching and learning. In our opinion, such a theory cannot ontologically separate teaching from learning. If we describe the teaching/learning process

as a way of building a common background, relating to the piece of knowledge involved, between the teacher and the students, and renewing it as this process unfolds, we have to see learning and teaching in transactions. That is to say that, if there is a fundamental asymmetry between the teacher and the students, due to their different relationship to the subject matter, the way students learn fundamentally depends on the teacher's instructional moves. In what we termed "Equilibration work" a teacher may obtain dramatically different learning depending on whether she carries on a "contract-driven" teaching process or a "milieu-driven" teaching process. Thus a theory of instructional practices has to account for these differences, and show how a "milieu-driven" instructional practice may enact a kind of Deweyan inquiry (Dewey, 1938/2008), in which students' certainty will be based on a first-hand relationship to the milieu and the knowledge in question. In that way, we agree with Koschmann's contention according to which "Deweyan inquiry, when successfully carried out not only effects a change in the problem solver (what psychologists treat as "learning") but also leads to a reconstruction of the problematic situation that led to the inquiry in the first place" (Koschmann, 2011, p. 12). We argue that we can see such "a reconstruction of the problematic situation" in our videoconferencing case study, in which the teachers' equilibration work allows the students to deal with the foreign language in a way they would not be able to achieve in the initial problematic situation. We account for this reconstruction by using our concept of learning game, which makes us see learning as the consequence of the teacher's game on the student's game in the milieu.

In that way, JATD, by considering learning and teaching as the process and the outcome of a joint action, does not ontologically separate the transactional process that occurs between the teacher and the students from the epistemic process within which the knowledge involved is made available. Knowledge itself is provided in a transactional way. It seems to us that such a conception could meet the distinction operated by Greeno (2001, 2011) between the systemic and semantic aspects of interaction. Systemic principles "involve ways in which students are positioned in interaction" and semantic principles "involve ways of achieving coherence of information, including alignment of the situation with the goal of the task (Greeno, 2011, p. 48). As Greeno put it, "these two kinds of principles are inherently interactive" (2011, p. 49). It seems to us that inquiring how the teacher's work (in particular what we refer to as "equilibration work") may enact a relevant relationship between systemic and semantic aspects in learning could be a promising avenue of research. It could be one of the goals of JATD, within a transactional perspective.

Technology-enhanced learning environment and JATD

We argue that technology-enhanced learning environment such as videoconferencing may offer regular opportunities to modify the very logic of teaching/learning practices. The reconstructed situation allows the students to position themselves as knowledgeable persons, entitled to produce relevant sentences in their mother tongue. We hypothesize that this positioning (Greeno, 2001, 2011; Harré & Van Langenhove, 1998) has a strong influence on the students' epistemic conduct.

One could assert that it has been done by augmenting the "degree of authenticity" of the situations. But the question is not so much that of the degree of authenticity as that of the *nature* of this authenticity. According to us, the empirical case study we provided in this paper is instructive from this point of view. By uttering *complete* sentences in their mother tongue, students do not meet the features of an "everyday situation", in which *only some words* could be sufficient to win the "monster game". But winning the "monster game" does not mean winning the didactic *learning game*, in that it does not afford relevant opportunities to learn. In the case study in question, as each student is entitled to provide other "foreign" students with accurate linguistic forms, the completeness of these forms can be seen as a warrant for better learning. In that way, the "authenticity" of the learning game is not an everyday life authenticity, it's a *didactic* authenticity – that one could see as the most rewarding authenticity in a learning situation.

Consequently, we argue, on the one hand, that technology-enhanced instructional practices have to enact counterfactual learning games, as we have seen in our case study, in which the logic of the situation itself puts the students in didactically authentic and relevant situations, that the teacher's equilibration work may concretize. On the other hand, our second case study clearly illustrates that videoconferencing can greatly enhance the transactional possibilities of the learning situations. Obviously this small-scale study represents a modest contribution to the debate about CALL research and the question of the effectiveness of videoconferencing. Multiple factors such as the learners' level, the teachers' expertise, the setting and the communicative situations have a significant effect on students' learning. Consequently, more research on videoconferencing and language learning in primary education needs to be conducted so as to investigate the specificity of the primary context and identify examples of good practice. Whatever could be the results of these investigations, we strongly believe that studying technology enhanced learning will lead us to reconsider the potentialities of JATD.

Finally, as we argued at the beginning of this paper, JATD stems from French Didactics and relies on it for its main concepts. But it seems to us that JATD represents an emerging paradigm for the educational sciences that could contribute to current research in the Learning Sciences Community, and more particularly with this

research, to CSCL. This research, according to us, leads to a new vision of the relationship between teaching and learning, a new vision of what could be a science of instructional practices grounded in a deep understanding of what learning is.

Endnotes

- (1) In our work, we use the word "transaction" as Dewey elaborated it in his late work with Bentley (Knowing and the Known), in which they explain that the notion of transaction differs from the notion of interaction in the way it emphasizes the fundamental dynamics of activity.
- (2) In the transcriptions, PT1 designates the teacher in the first episode, PTF the French teacher and PTE the English teacher in the second episode.

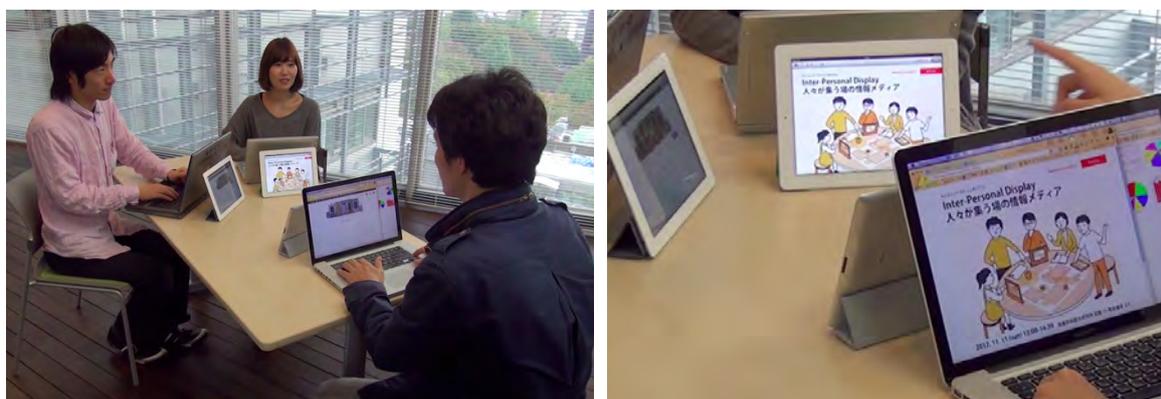
References

- Brousseau, G. (1997). *The theory of didactic situations in mathematics*. Dordrecht: Kluwer.
- Chevallard, Y. (2007). Readjusting didactics to a changing epistemology. *European Educational Research Journal*, 6(2), 131–134.
- Clark, H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Dewey, J. (1916/1985). *Democracy and Education*. In J. A. Boydston (ed.), *The Middle Works*, Vol. 9. Carbondale: University of Southern Illinois Press.
- Dewey, J. (1938/2008). *John Dewey The Later Works, 1925-1953: 1938: Logic: The Theory of Inquiry*. Chicago: Southern Illinois University Press.
- Dewey, J. & Bentley, A. (1949). *Knowing and the known*. Boston: Beacon Press.
- Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge: MIT Press.
- Fleck, L. (1979). *Genesis and development of a scientific fact*. Chicago: University of Chicago Press.
- Goldman, R., Pea, R., Barron, B., & Derry, S. J. (2007). *Video Research in the Learning Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Goldman, S., & McDermott, R. (2007). Staying the course with video analysis. In R. Goldman & R. Pea & B. Barron & S. Derry (Eds.), *Video research in the learning sciences* (pp. 101-113). Mahwah New Jersey: Lawrence Erlbaum Associates.
- Greeno, J.G. (2001). *Students with competence, authority and accountability: affording intellectual identities in classroom*. New York: The College Board
- Greeno, J. G. (2011). A Situative Perspective on Cognition and Learning in Interaction. In T. D. Koschmann (Ed.), *Theories of Learning and Studies of Instructional Practice* (pp. 41-72). New York: Springer.
- Gruson, B. (2009). Etude de la dialectique contrat-milieu dans l'enseignement-apprentissage de l'anglais en CM2 et en 6è. *Revue Suisse Des Sciences De l'Éducation*, 31, 641-659.
- Gruson, B. (2010). Analyse comparative d'une situation de communication orale en classe ordinaire et lors d'une séance en visioconférence. *Distances et Savoirs*. 3, 8, 395-423.
- Harré, R., & Langenhove, L. van. (1998). *Positioning Theory: Moral Contexts of Intentional Action*. Oxford: Blackwell.
- Koschmann, T. D. (2011). *Theories of Learning and Studies of Instructional Practice*. New York: Springer.
- Laborde, C., Perrin-Glorian, M.-J., & Sierpiska, A. (2005). *Beyond the Apparent Banality of the Mathematics Classroom*. New York: Springer.
- Mead, H. G. (1934). *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Piaget, J. (1975). *L'équilibration des structures cognitives : Problème central du développement*. Paris : PUF.
- Schubauer-Leoni, M.-L., & Perret-Clermont, A.-N. (1997). Social interactions and mathematics learning. In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics. An international perspective* (pp. 265–283). London: Psychology Press.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, 10(2), 70-76.
- Sensevy, G. & Mercier, A. (2007) *Agir ensemble. L'action conjointe du professeur et des élèves dans le système didactique*. Rennes : PUR.
- Sensevy, G. (2011a). Sensevy, G. (2011). Overcoming fragmentation : Towards a joint action theory in didactics. In B. Hudson & M. Meyer (Eds.), *Beyond Fragmentation : Didactics, Learning and Teaching in Europe* (pp. 60-76). Opladen and Farmington Hills : Barbara Budrich.
- Sensevy, G. (2011b). *Le Sens du Savoir. Eléments pour une théorie de l'action conjointe en didactique*. Bruxelles : De Boeck.
- Tiberghien, A., & Sensevy, G. (2012). Video studies : Time and duration in the teaching-learning processes In J. Dillon & D. Jorde (Eds.), *Handbook "The World of Science Education" volume 4*(pp. 141-179). Rotterdam/Boston/Taipei : Sense Publishers.
- Wittgenstein, L. (1997). *Philosophical Investigations*. Oxford: Blackwell.

Inter-Personal Browsing: Supporting cooperative web searching by face-to-face sharing of browser pages

Tomoko Hashida, The Univ. of Tokyo, 7-3-1, Bunkyo-ku, Hongo, Tokyo, 113-0033, hashida@nae-lab.org
 Koki Nomura, Domus Academy, Via Mario Pichi, 18, 20143 Milano, nomura@nae-lab.org
 Makoto Iida, The Univ. of Tokyo, 7-3-1, Bunkyo-ku, Hongo, Tokyo, 113-0033, iida@nae-lab.org
 Takeshi Naemura, The Univ. of Tokyo, 7-3-1, Bunkyo-ku, Hongo, Tokyo, 113-0033, naemura@nae-lab.org

Abstract: We are studying a framework that facilitates the gathering and sharing of information in group work situations where several people are working on multifaceted problem-solving or emergent tasks by cooperating and sharing the workload. In particular, we are developing a framework for the smooth sharing of public information to stimulate debate while maintaining the focus of the group's web searching activities and confining private information on the PC screens of individual users. In this paper, we report on the implementation of Inter-Personal Browsing - a framework where each user has a small shared display in addition to the PC display, and can use this to share web browser screens with other users. Having introduced this system into a mechanism for practical learning of group work methodology, we introduce the results of studying its effectiveness from the viewpoint of information sharing and communication.



(a) A user's iPad automatically displays the window of the web browser running on the user's own PC. (b) A user can receive pages by touching other's iPad.

Figure 1 Concept of Inter-Personal Browsing

Introduction

In the fields of education and business, there has been renewed recognition of the importance of group work, where several people work on multifaceted problem-solving or emergent tasks by cooperating and sharing the workload. Recently, there have also been greater opportunities for people to bring laptops or PCs into group work situations, but in this case the attention of users is drawn towards their individual computer screens, thus depriving them of the communication opportunities that group work is supposed to provide. One possible reason for this is that the screens of other people's laptops and PCs are highly likely to show private information, which can make people feel apprehensive about perusing someone else's screen. However, having people direct their attention entirely toward their own computer screens is undesirable with regard to facilitating the exchange of information in group work situations. We have therefore been looking into ways of addressing this situation by using such means as projector displays to produce large-screen public displays that are shared by all the members of a group, or by using terminal displays whose orientation can be changed, such as iPads. However, this sort of approach can lead to problems such as private information being displayed on a large screen, or information only being publicized to a subset of the group members, or to users having to go to the trouble of changing the orientation of their terminal screens.

Against this background, we considered that a new framework is needed to facilitate the sharing of public information that can stimulate debate in group work situations while keeping private information confined to the computer screens of individual users, without the need for intentional operations to be performed by the users. In particular, we focused on web browser windows as a source of information to be made public in group work performed using PCs. Recently, web searches have been often used in group work on PCs, especially at the investigatory stages. According to Morris (2008), 97.1% of people have taken part in joint web

searches, but over 80% of them used inefficient methods for sharing their search results, such as looking at the screen over the other person's shoulder, sending search results by email, or sharing the screen display by turning it towards the other person. We therefore used iPads and browser extensions to implement a mechanism that satisfies three conditions: (i) only the browser window is made public as shown in Figure 1(a), (ii) it is easy for users to access the search results of other users in their own PC as shown in Figure 1(b), and (iii) this framework can be set up instantly without having to introduce any special equipment. Since this framework supports website browsing between multiple individuals, we called the proposed system Inter-Personal Browsing.

In this paper, we first present a detailed introduction to the Inter-Personal Browsing system that we implemented from this viewpoint. Then, by using this system in a group work situation, we compare its performance with group work performed using existing information sharing tools such as notes written on paper and Google Documents shared via PCs, and we study how this system changes the nature of the group work, such as the ease with which information can be shared, and the degree to which it stimulates communication. We conducted verification trials to test these effects in a practical work group lecture.

Related Works

Systems that support joint web searching have mainly used either a shared display such as a computer screen, or a tabletop device.

Systems implemented using a shared display include SearchTogether (Morris & Horvitz, 2007) and CoSearch (Amershi & Morris, 2008). SearchTogether is software that allows multiple users in remote locations to use a single shared display to perform searches, and includes features such as a text chat function, individual user search history displays, and a search results display window. CoSearch is an enhanced version of SearchTogether, but unlike SearchTogether it assumes that the users are all situated around a single computer. Although shared display systems make it easy to share information and provide information with better clarity, they have drawbacks such as being unable to secure adequate workspace for each user, resulting in poorer individual work efficiency.

Tabletop systems include WeSearch (Morris et al., 2010) and WebSurface (Tuddenham et al., 2009). These systems exploit the space and interactivity offered by a tabletop workspace. An experimental study comparing the execution of collaborative tasks on personal devices and on a tabletop device (Heilig et al., 2011) has also confirmed that people are more likely to interrupt and engage in the actions of other users' actions when using a tabletop device. Tabletop systems eliminate the drawbacks of shared displays, but their high cost and lack of portability have impeded the introduction of these systems.

New trends have also emerged recently. These include frameworks that facilitate the integrated use of heterogeneous devices, and PCs equipped with double-sided displays. The former approach makes it easier for individuals to obtain the information they need in a system where it is possible to use tabletop displays and mobile displays in combination with each other (Doring et al., 2010; Seifert et al., 2012). However, it fails to address issues such as the drawbacks and high cost of tabletop displays. On the other hand, the Asus Taichi (a laptop PC that went on sale recently) is equipped with a double-sided full HD IPS display. This offers an inexpensive way of displaying information on the front (facing the user) and back surfaces of a display, but does not allow the user to select specific information on the screen and have just this information displayed on the back surface. This makes it difficult to share information safely.

The Inter-Personal Browsing system proposed in this paper tackles these issues from a different viewpoint by employing a simple configuration whereby users install a browser extension on the PCs they use every day, and they each have their own iPad for use as a public display.

Inter-Personal Browsing

This paper summarizes the Inter-Personal Browsing framework, where private information is kept on the user's individual PC, while public information that can stimulate discussions in the group work is displayed on a separate web browser screen. First, after describing the concept of an individualized public display, we present a detailed discussion of Inter-Personal Browsing.

Individualized public display

When a projector or the like is connected to a PC's external display terminals, it is generally operated either in a "mirror" mode where the PC's screen (main display) and the shared screen (external sub display) display the same content, or an "extension" mode where they each provide separate displays. For example, if private information is shown in a mail application and public information is shown in a browser application, then the content of the mail application would be displayed to others on the shared screen in the mirror mode. Also, in the extension mode, when a browser is displayed on the shared screen, the user is forced to constantly work while looking at the shared screen. Instead, this paper proposes a method where the mail and browser applications continue to be displayed on the PC screen in the normal way, while a mirror of just the browser window is displayed on the shared screen. These relationships are summarized in Table 1.

Table 1 Information presentation concepts in the proposed method

	PC screen (Main display)	Shared screen (External sub Display)
Mirror	Mail, Browser	Mail, Browser
Extension	Mail	Browser
Proposed method	Mail, Browser	Browser

In this paper, a shared screen is owned by each individual user. This is called an individualized public display. By using the individualized public display to show only a mirror of the browser window, users can check each others work status without having to exercise restraint, resulting in group work with smoother communication.

System overview

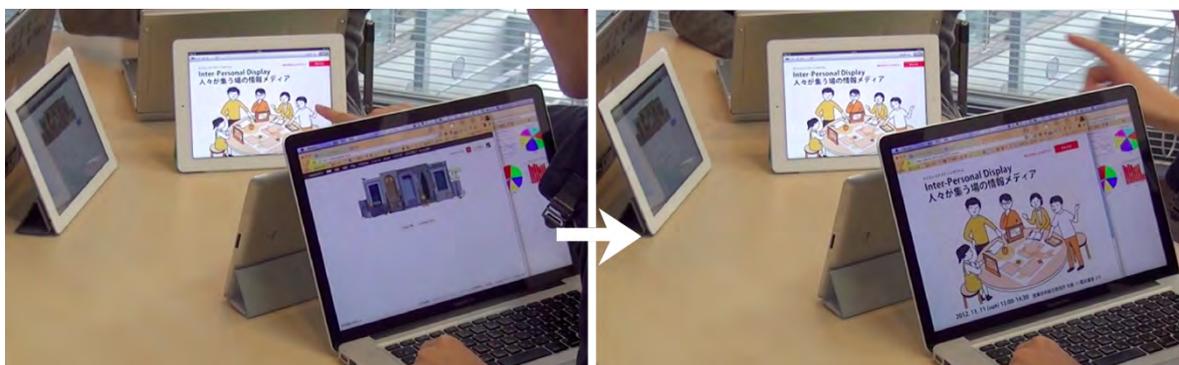
In this paper, we used iPads to implement the individualized public displays discussed above. Figure 2 shows the configuration of the Inter-Personal Browsing system.

A user’s iPad automatically displays the window of the web browser running on the user’s own PC, and is automatically made available for sharing at the location of the group work. However, as shown in Figure 2(a), this is not implemented as a simple screen copy, but by running a separate browser in the iPad to show the content of the same URL. Next, as shown in Figure 2(b), an iPad is placed beside each user in addition to the user’s own individual PC so that it can be seen by other users. After making it possible for users to keep track of each other in this way, the next important requirement is to make it possible for users to interact with the system in order to transfer information from another user’s public display to their own PCs. To allow group work to proceed smoothly while sharing information, it is important to make it easy for users to transfer pages to their own PCs where they can view them in more detail without interrupting the work of other users. For this purpose, we implemented a mechanism whereby a user can use the iPad’s touch input functions to receive pages without anyone else’s involvement. Specifically, as shown in Figure 2(c), this is done by touching the other user’s public screen (iPad) while holding down a control key. To implement this function, we require a framework for specifying the PC of a specific user out of multiple other users. It is also essential that it is implemented in a way that does not require users to use a specific platform or operating system.



(a) Sharing a URL on an individualized public display

(b) System arrangement



(c)Turning pages by touch input

Figure 2 System Configuration of Inter-Personal Browsing

System implementation

In this system, mirror displays on iPads and the transfer of information to other users is achieved by transmitting URLs between the iPads and web browsers. We used browser extensions to allow the system to be used in a wide range of PCs and operating systems, and we also implemented an iPad application to work with these extensions. We set up a server to centrally manage the flow of information between PCs and iPads to facilitate many-to-many communication, and the communication was all routed via this server. The display of web pages on the iPads was implemented by transmitting a URL to the iPad every time the browser loads a new page. With a browser extension API, we capture an event whenever a new page is loaded in the browser, whereupon the URL of the new page is transferred to the iPad. Communication is implemented by HTTP between the browser and server, and by socket communication (TCP) between the server and iPad. The procedure is as follows.

- (1) Fetch the URL of the web page displayed by the browser.
- (2) This URL is sent to the server, and is transferred from the server to the iPad.
- (3) The same page is then displayed on the iPad.

The web page transfer function is implemented so that when the receiving user touches the iPad, the URL displayed by the iPad is transferred to the user's PC. For this to happen, it is necessary to identify which user touched the iPad. In this paper, as an interaction method for implementing this, we provided a framework whereby the user holds down the Ctrl key while touching the iPad in order to designate himself or herself as the receiving party. In this way, it is possible for multiple users to receive information simultaneously with a single touch operation. The processing procedure is as follows.

- (1) The browser extension is used to report to the server when a Ctrl key is pressed.
- (2) When an iPad is touched, its URL is transmitted to the server.
- (3) The server identifies the users that are pressing their Ctrl keys.
- (4) The URL information is sent from the server to the browsers running on the individual PCs of these users.
- (5) On receiving this URL information, the browsers open the URL in a new tab.

The browser keypress events are implemented using the browser extension API, and the iPad touch events are implemented using the iOS API. For the communication of URLs, it is necessary for them to be transmitted from the server to the clients. In the communication systems currently available in web browsers, support for server-to-client communication is provided by Comet and WebSocket web technologies. Comet uses HTTP communication to achieve artificial bidirectional communication. Although WebSocket is faster and generates lower server loads, it is still relatively new and its specifications are not yet fixed. We therefore decided to use Comet in this system.

Verification of concept

To examine how the quality of group work is changed by using Inter-Personal Browsing proposed and implemented in the above way, we performed group work tests using Inter-Personal Browsing in a lecture course that we are currently teaching. This lecture course is called "Groupwork of the future - future classroom technology", and is being taught as part of the 2012–3 masters course at Tokyo University. Its aim is to provide practical education in the use of digital tools and group work methodologies, and it comprises 13 lessons targeting 9 undergraduate students.

The specific aim of practical research through this course is to use Inter-Personal Browsing in a comparative study to see if it changes the quality of group work with regard to communication and the sharing of information. For comparison, we also provided group work based on notes written on paper (paper-based group work) and group work using existing digital information-sharing tools like Google Documents (<https://www.google.com/intl/en/drive/start/index.html>) and Dropbox (<https://www.dropbox.com/home>) running on PCs (PC-based group work). The PC-based group work differed from paper-based group work in that it was able to handle larger quantities of information and made it easier to share resources such as the results of the group work. However, it also resulted in less communication because the users' attention was directed toward their computer screens, and it was difficult for them to share work processes. Our study focused mainly on the issue of whether the introduction of Inter-Personal Browsing, where information can be shared easily with the browser window, causes any changes in communication or information sharing (especially during the investigation process). The details of these discussions and our findings are summarized below.

Overview of the proof-of-concept experiment

In this experiment, group work was performed using three different tools, and a combination of behavioral observation and questionnaires was used to clarify the level of communication activity and the ease with which information could be shared during the examination phase of the group work. Since this experiment was performed in tandem with lectures, it was difficult to maintain absolute control over the tasks, and the tasks performed using each tool were also different. This approach was adopted because our purpose here is to clarify

the general feasibility of Inter-Personal Browsing (IPB) through practical experience. The first tool was paper, and users were given the task of creating a billboard for the Nikkei newspaper. In this task, the users had to decide which articles were worthy of inclusion, summarize these articles, and create the final layout. This is called method 1. The second tool was a combination of paper and PCs, and users were given the task of creating a time-line of a celebrity's career. In this task, it was necessary to determine which aspects of the celebrity's career should be included in the time-line (e.g., awards won by the celebrity), gather data on the celebrity's name and active time period, and then assemble this information in a time-line where it can be easily understood. Here, the PCs were used to access Google Documents as an online document creation tool and Dropbox as an online storage tool. This is called method 2. The third tool was a combination of paper, PCs and IPB, and users were given the task of designing the interior of a shared house for female college students. In this task, the users had to clarify their target residents (which area they live in, what their interests are, etc.), put forward concepts that these residents might like, create a collection of interior furnishings needed to realize these concepts, and produce summaries of these concepts together with costs and images. This is called method 3. In each system, the group work was performed over three sessions of approximately 90 minutes, including time for practice, presentation, review and so on. The nine students were split into two groups, and the members of these groups were replaced for each method.

To perform a comparative study of the quality of group work in each of these methods, we first observed the state of communication in each group by having on average about two observers constantly watching each group. The observers concentrated on group characteristics such as the amount of conversation, lines of sight and seating arrangements (the students were instructed to arrange the desks and chairs and move around as much as they wanted in order to facilitate communication). The last ten minutes of the lecture was designated as a review period in which the students were asked to complete questionnaires. As shown in Table 2, the questionnaire included two questions where students were asked to reflect on the level of activity in the discussions. Next, in methods 2 and 3, we used the separate questionnaire shown in Table 3 to compare the ease of sharing information during the examination phase with and without Inter-Personal Browsing.

Results and discussion

Communication

To examine the effectiveness of communication in the group work, the students in methods 1, 2 and 3 were asked to complete the two questions shown in Table 2.

First, with regard to the all questionnaire of Table 2, in methods 1 and 3 the students responded that they had all actively participated in the group work and that the group itself facilitated active discussions, suggesting that active communication had taken place. On the other hand, in method 2 where group work was performed using PCs, some students responded that they had been unable to participate actively in the group work or play an active role in discussions (38% answered "No" to Q1-1 "Did you actively participate in the group work?" and 63% answered "No" to Q1-2 "Were there active discussions in your group?"). There were students who gave this response in each group.

From photographs of the discussions, we can see that in method 1 (Figure 3(a)), the students took steps such as moving the desks around so they could speak more easily with the other students, and in method 3:IPB (Figure 3(c)) they were also seen to lean across the desk while in conversation. The amount of conversation itself was found to be large. However, in method 2 (Figure 3(b)), the students continued to direct their gaze at their computer screens even when carrying out the actual work itself, and there was less conversation than in methods 1 and 3.

These findings suggest that the use of Inter-Personal Browsing allows for more communication than existing PC-based group work, and a high level of activity close to that of traditional paper-based group work.

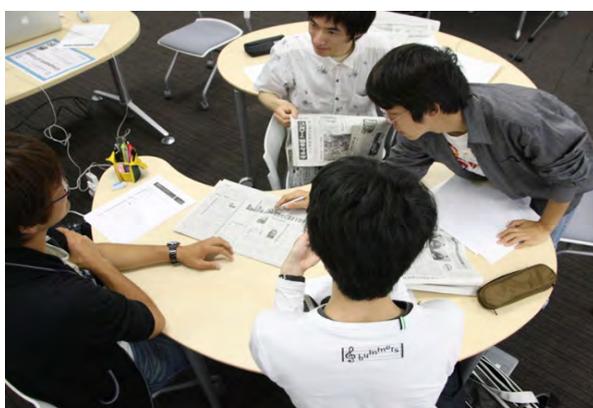
Sharing of information

For a more detailed consideration of the sharing of information in the examination phase of the group work, the students in methods 2 and 3 were asked to complete the three questions shown in Table 3.

First, in method 2, most of the students were concerned about what the other students were examining during the task (88% answered "yes" to Q2-1 "Were you concerned about what the other students were examining?"). On the other hand, in the question relating to how students ascertained what the other students were doing (Q2-2), we received a number of similar responses such as "I just looked at the comments on shared sheets in Google Document". Moreover, in the question about issues experienced when sharing what the students had examined (Q2-3), we received responses such as: "It was difficult to find out what the other members were studying, or how much, so I didn't know how far I should go with my own study", "We were all working at the same time, and ended up studying the same things", "It took longer than I expected to share things, and it was hard to ascertain straight away what the others were doing". That is, in group work conducted as in method 2 where information is shared via services such as Google Documents and Dropbox, it is difficult

for people to share what they are studying in real time, so people do not know how much progress the others are making, it is difficult to split tasks up because people are unaware of how much progress the others are making, and a number of the students felt that the task took longer than it should have. A possible explanation of how this situation arose is that when people used Google Documents, they were focused on the screen so much that they did not notice what the others were doing. This would have made it harder to know who had provided each bit of information, thereby depriving the users of opportunities to talk to each other.

In method 3, as in method 2, most of the students were concerned about what was being studied by the other students during the task (88% answered “yes” to Q2-1 “Were you concerned about what the other students were examining?”). In the question relating to how students ascertained what the other students were doing (Q2-2), we received a number of similar responses, including “When someone found an interesting site, they said so, and the other members used their iPads to display and assess this site on their own computer screens”. Moreover, in the question about issues experienced when sharing what the students had examined (Q2-3), we received many responses such as “No, nothing in particular”, but there was also a response that pointed out a problem with the system: “The page sharing sometimes didn’t work properly”. Consequently, in group work using Inter-Personal Browsing as in method 3, users can share information indirectly with one another, and can communicate verbally if there is something they want to convey to the others or something important they need to know. It is thus inferred that the students themselves were able to figure out a way of actively participating in the group work while switching between different channels of communication.



(a) Method 1: paper



(b) Method 2: paper & pc



(c) Method 3: IPB

Figure 3 The appearance of group work in each method

Table 2 Results from questionnaire 1 (average ratio of participating students who answered affirmatively)

	Method 1	Method 2	Method 3
Q1-1 Did you actively participate in the group work?	100%	63%	100%
Q1-2 Were there active discussions in your group?	100%	44%	100%

Table 3 Results from questionnaire 2

	Method 1	Method 2	Method 3
Q2-1 Were you concerned about what the other students were examining?	–	88% answered “yes”.	88% answered “yes”.
Q2-2 How did you ascertain what the other students were examining?	–	75% just looked at the comments on shared sheets in Google Documents	88% used Inter-Personal Browsing and taught each other orally.
Q2-3 Did you have any difficulty sharing what you had examined?	–	50% had difficulty finding out what the other members were studying.	63% did not have no difficulty in particular.

Conclusion

We have proposed a framework called Inter-Personal Browsing that facilitates the sharing of information by displaying a web browser window on an individualized public display, and we have performed a proof-of-concept test as part of a practical lecture course in group work, in which we performed a comparative study of group work using this framework and other tools. Our results suggest that the proposed Inter-Personal Browsing facilitates more active communication than group work using ordinary PCs alone, and allows task processes to be mutually shared more smoothly. However, the proposed Inter-Personal Browsing still has several issues. For example, its framework currently only allows web pages to be shown to other users. On the other hand, when diverse tasks are performed by a group, there may be many other different combinations of applications that users want to keep private and applications that users would want to show to other people. In the future, we intend to provide greater freedom in the choice of applications and introduce a framework where users can take part in group work while diversely switching between information they want to show (public information) and information they want to keep to themselves (private information).

References

- Morris, M. R. (2008). A survey of collaborative web search practices, In Proceeding of the twenty-sixth annual CHI conference on Human factors in computing sys- CHI '08, pp1657-1660, Italy
- Morris, M.R. and Horvitz, E. (2007). SearchTogether: An Interface for Collaborative Web Search, In Proceedings of the 20th annual ACM symposium on User interface software and technology - UIST '07, pp.3-12, USA
- Amershi, S. and Morris, M.R. (2008). CoSearch: A System for Co-located Collaborative Web Search, In Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08, New York, pp.1647-1656, Italy
- Morris, M.R., Lombardo, J. and Wigdor, D. (2010). WeSearch: supporting collaborative search and sensemaking on a tabletop display, In Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10, pp.401-410, USA
- Tuddenham, P., Davies, I. and Robinson, P. (2009). WebSurface: An interface for co-located collaborative information gathering, Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS '09, pp. 181-188, Canada
- Heilig, M., Huber, S., Gerken, J., Demarmels, M., Allmendinger, K., and Reiterer, H. (2011). Hidden details of negotiation: the mechanics of reality-based collaboration in information seeking. In INTERACT'11, Springer-Verlag, pp.622–639, Berlin
- Doring, T., Shirazi, A. S., and Schmidt, A. (2010). Exploring gesture-based interaction techniques in multi-display environments with mobile phones and a multi-touch table. In AVI '10, ACM, pp.419–419, New York
- Seifert, J., Simeone, A.L., Schmidt, D., Reinartz, C., Holleis, P., Wagner, M., Gellersen, H. and Rukzio, E. (2012). Mobi-surf: Improving Co-located Collaboration through Integrating Mobile Devices and Interactive Surfaces, In Proceedings of ACM International Conference on Interactive Tabletops and Surfaces, pp.51-60, Boston
- ASUS TAICHI™ <http://taichi.asus.com/jp/#TEXTURE>

Acknowledgments

This project is supported by JST (Japan Science and Technology Agency)-CREST (Core Research for Evolutionary Science and Technology) project “Harmonized Inter-Personal Display Based on Position and Direction Control”. We would like to thank Shohei Takei, Tomoko Ohtani, Motohide Hatanak and Yasuaki Kakehi for helping out with the lectures.

Learner-Support Agents for Collaborative Interaction: A Study on Affect and Communication Channels

Yugo Hayashi, University of Tsukuba, Faculty of Library, Information and Media Science, 1-2, Kasuga, Tsukuba, Ibaraki, 305-8550, Japan, hayashi.yugo.gp@u.tsukuba.ac.jp

Abstract: This study investigated if and how a conversational agent facilitates better explanations from students in a computer-based collaborative activity. Pairs of students enrolled in a psychology course performed a task where they attempted to explain to their partners the meanings of technical psychological terms. During the task, they interacted with an affect-based conversational agent, which was programmed to provide back-channel feedback and metacognitive suggestions through visual and/or audio output. The study compared students' performance after using this agent with their performance after using an agent without audio output or affective expressions. Our findings suggested that the use of multiple communication channels for feedback facilitates collaborative learners' understanding of concepts. This provides implications for designing pedagogical agents for effective collaboration.

Introduction

Past studies on collaborative problem solving in cognitive science have revealed how concepts are understood or learned through interaction. Researchers have shown that asking reflective questions for clarification to conversational partners is an effective interactional strategy to better understand a problem or concept (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Miyake, 1986; Okada & Simon, 1997; Salomon, 2001). The use of strategic utterances such as asking for explanations or providing suggestions has been found to stimulate reflective thinking and metacognition involved in understanding concepts. Playing different roles during an explanation could also help problem solvers construct external representations and develop their understanding of concepts (Shirouzu, Miyake, & Masukawa, 2002). These studies suggest that one's ability to explain is key to understanding and learning concepts. However, effective explanation often fails if retrieving and associating relevant knowledge required for the explanation is difficult. These difficulties typically rise among novice problem solvers (Coleman, 1998; King, 1994). Additionally, people cannot learn effectively if they cannot communicate fluently with each other, such as when they have little conversation experience or have conflicting perspectives (Hayashi & Miwa, 2009).

One way to help collaborative problem solvers could be introducing a third-party facilitator who could provide suggestions or back-channels. An important breakthrough has been the use of computer-based technology, such as pedagogical conversational agents, for this purpose. Pedagogical agents can serve not only as multimedia extensions, but also as social entities that become learning companions (Moreno & Mayer, 2002). The present study investigated the use of pedagogical conversational agents that could facilitate effective explanations.

Pedagogical conversational agents as learning advisers

The effects of affective feedback

Recently, studies have shown that conversational agents that act as educational companions or tutors can facilitate learning (Baylor & Kim, 2005; Holmes, 2007). Many computer-based tutoring systems use conversational agents (Graesser & McNamara, 2010), but it is not fully understood what kinds of support from such agents could improve collaborative learners' performance. Collaborative activities are difficult, especially for new learners who are not used to expressing their thoughts or understanding others' viewpoints. It is assumed that learners would have high cognitive loads during explanation, and paying attention to both their partners and third parties (computer agent) could be too difficult. Holmes (2007) indicated that learning pairs ignored the presence of an agent and conducted the learning activities on their own. Therefore, investigating how to design and use pedagogical agents for effective enhancement of collaborative activities is very important. These agents should be designed using concepts based on human collaborative activities. One point that must be considered in studies of human performance is emotion. Emotions affect individuals' performance in both negative and positive ways and these effects are especially important for learning activities (Baylor & Kim, 2005). For example, Bower and Forgas (2001) revealed that positive mood can increase memory performance. Mayer (2001) also demonstrated that a positive state of mind can improve text comprehension. Moods may also affect people's verbal and non-verbal performance on activities. Kim, Baylor, and Shen (2007) examined how positive and negative comments from conversational agents can affect learning performance. They programmed a pictorial image of an agent to project a textual message to the participant; in the positive

condition, a visual avatar produced a short comment such as “This task looks fun.” In the negative condition, it produced a comment such as “I don’t feel like doing this, but we have to anyway.” The results showed that the conversational agents that provided comments in a positive mood increased participants’ motivation to learn. In addition, it is important that students acquire confidence, even if it is only the “illusion of knowing.” This phenomenon describes how students sometimes do not accurately acquire knowledge, but still become confident that they have learned something. This confidence is important for facilitating students’ motivation towards learning.

Hayashi (2012) examined how the expressions of positive and negative pedagogical conversational agents could facilitate explanations between collaborative pairs. Participants were required to explain a concept taught in a university psychology course to others through a computer chat system. During the task, they interacted with a conversational agent, which was programmed to provide back-channel feedback and metacognitive suggestions to encourage and facilitate conversational interaction. Their results suggested that positive affective feedback from these conversational agents facilitates explanation and better learning. Thus, conversational agents can play a role in pedagogical tutoring and may help trigger deeper understanding of a concept that students are trying to explain. The above studies have suggested that explanation performance will also be enhanced if suggestions are given in a positive mood either verbally or through visual feedback. Unfortunately, further analysis of the dialogue process shows that learners sometimes do not listen to agents’ comments or suggestions. This indicates a need to further investigate how to provide such affective feedback more effectively. However, there are several difficulties in providing feedback during interactions among learners, such as (1) timing of feedback, (2) quantity/quality of information, and (3) communication channels. Appropriate communication channels are especially important for new learners, as it helps them avoid cognitive overload from the information provided by both agents and learners. In the next section, discuss are made on such communication channels.

Communication channels during pedagogy in online tasks

While the issue discussed above concerns the nature of the information in collaborative tasks, it is more concerned with identifying the communication channels that could improve learners’ comprehension of unfamiliar concepts when affective feedback is presented

A related issue is how information is processed (and how it is processed in the most efficient manner), which has been a long-considered topic in psychology. For example, Baddeley and Hitch (1974), in their “working memory” model, argued that there are two subcomponents of information processing that handle different types of information (visual and verbal). Thus, human cognition comprises different components or modules for different kinds of information, allowing us to predict that people perform better when information is presented more efficiently and economically. That is, when people perform two kinds of activities at the same time, they may do better if information is presented through different communication channels rather than a single channel. Moreover, previous research has suggested that information in working memory may overload processing when the same modality is used for various types of information, thus making it harder for learners to understand (Chandler & Sweller, 1991; Mousavi, Low, & Sweller, 1995; Sweller, Chandler, Tierney, & Cooper, 1990). Mayer (2001), who extended this view to multimedia learning, reported that students who viewed an animation depicting the formation of lightning while listening to a narration explaining this process generated more useful solutions on a subsequent problem solving transfer test than did students who viewed the animation with the narration as subtitles. In the case of collaborative problem solving tasks, learners may perform better if agent feedback is provided through voice messages (i.e., auditorily) than by text message, since they are already engaged in reading a description through a computer monitor (i.e., visually).

Another important point to consider in relation to communication channels is the problem of attention during tasks, which is related to the efficiency problem discussed above. This is generally referred to as the “split attention effect,” which is often observed in poorly designed instructional materials where, for example, the same modality (e.g., visual) is used for different types of information within the same display. In such cases, learners’ attention may be split between different materials, making it difficult to use the materials effectively. When individuals are engaged in a challenging task, they may be preoccupied with information that is more directly related to the task and pay less attention to information, which although useful, is only indirectly related to the task.

For collaborative learning via online communication, possible ways to communicate with one another include using text chats (visual) and/or voice messages (auditory). Unlike recent web-based tutoring and e-learning systems, the most widely used technology is text-based communication (e.g., Hayashi, 2012; Holmes, 2007). In this case, attention may be paid primarily to the exchange of textual information on the computer display, since explanation is performed via online visual communication, with students describing concepts by typing. It is likely that learners, especially novice learners, are able to pay sufficient attention to affective feedback from a computer agent, even if this feedback is given in the same modality (visually). However,

collaborative learners may be able to pay more attention if they are given the feedback in a different modality (auditorily).

Research Goal and hypothesis

The goal of this study was to experimentally investigate if and how conversational agents facilitate students' understanding and learning of concepts. It was hypothesized that the use of affective feedback could facilitate collaborative learners' understanding of concepts. In addition, explanation performance would be enhanced if positive feedback from computer agents were given through multiple communication channels (i.e., visual or auditory modalities). This paper analyzes the performance of student pairs who performed an explanation activity on two types of technical terms.

Method

Experimental task and procedure

The experiment was conducted in a room where the computers were all connected by a local area network. Participants were given two technical terms presented on the screen. They were "schema" and "long-term memory," and had been previously introduced in the psychology class. Along with the key terms, a brief explanation of the concept was provided. Participants were asked to describe these concepts in their own words. After this pretest, participants logged into computers and used the chat program installed from a USB flash drive (see the next section for details). Pairs of participants then communicated through this program, with one participant explaining to another the meaning of each word presented on their computer screen. After one concept had been explained, the partners switched roles and the other partner explained the remaining concept. This was then repeated, so that both students had a chance to explain both terms. All participants received the same suggestions from the agent on how explanations should be given and how questions should be asked about the concepts. They then took a posttest, which consisted of the same material as the pretest. Participants' descriptions of the concepts in the posttest were compared with those of the pretest to determine if the participants had gained a deeper understanding of the concepts after the collaborative activity.

Experimental system

In the experiments, a computer-mediated chat system was set up through computer terminals connected via a local area network and the participants' interactions during the activity were monitored. The system used in the experiments was programmed in Java (Hayashi, 2012 a).

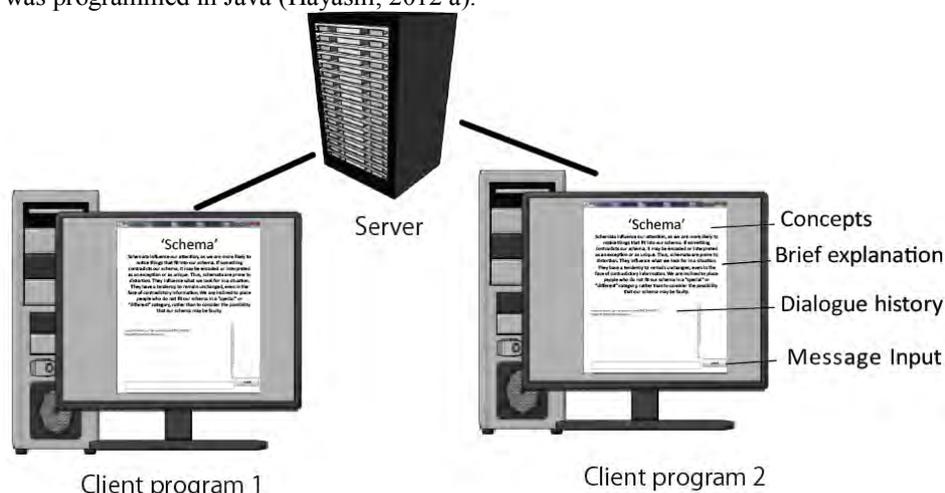


Figure 1. Interface of the task.

The system consisted of three program modules, including: (1) server, (2) client, and (3) agent, all of which were simultaneously activated. The server controlled all communication traffic during the task, and the client programs were installed in the students' computers to activate the graphical user interface (GUI) for communication (Figure 1). Multithreads were used so that the server program could send all messages to the clients' chat systems and the agent simultaneously.

The pedagogical agent used in this study was a simple rule-based production system composed of the following three different modules: (1) semantic analyzer, (2) generator, and (3) motion handler. First, the textual input of all of the participants' conversational exchanges was sent to the semantic analyzer in the agent. The agent monitored all the chat dialogues and messages were analyzed according to a list of keywords via the "bag of words" method (e.g., "I think that a **schema** is some kind of **knowledge** that is used based on one's own

experience” (detected keywords are shown in bold italics). The list of keywords was stored in a database (the “Dictionary Database”) in the semantic analyzer. Thirty different keywords were registered in the database. These keywords were selected according to a past study (Hayashi, 2012 a). Next, the extracted keywords were sent to the “working memory” in the generator and processed by a rule base, where various types of rule-based statements such as “if X then Y” were stored to generate prompt messages (if there are several matching statements for the input keywords, a simple conflict resolution strategy is utilized). When the matching process was completed, prompt messages were selected and sent back to the working memory in the generator. The messages consisted of information about the goals and the achievements of the task, and some initial suggestions on how to give good explanations to others. This point was designed according to the method used in Holmes (2007). The basic response rule was that if too many keywords were detected in the system, then prompts were generated asking students to use different words. For example, if the learner simply copied and pasted the words used by the system, it would provide messages such as “You should use more original words in your explanations.” When the system detected some keywords such as technical words, it generated messages such as “Good! You are explaining the concept with some unique words. Keep on going!” When the system detected combinations of technical keywords and questions, it provided messages addressing those combinations. For example, “It seems you have difficulty answering this. But you use good keywords!” Each output message was presented in text on the computer display.

The messages generated by the rule base were also sent to the motion handler module to activate an embodied conversation agent, a computer-generated virtual character that produced human-like behaviors such as blinking and head-shaking. The types of affective expressions used were based on an affective model developed in a preliminary study (Hayashi, 2012 b). These expressions were created using the 3D-image/animation-design tool Poser 8 (www.e-frontier.com).

Participants and conditions

In this study, 58 participants (23 men and 35 women; mean age = 19.78) participated in pairs. Participants were all undergraduate students taking a psychology course, who participated as a part of coursework. They were randomly assigned to three conditions, which varied according to how the suggestion prompts were presented and how the conversational agents were used (see the section below for details). In all conditions, the participants were given positive suggestions, which were synchronized with the facial expressions of the embodied agent. The messages were given through chat dialogue and the virtual character moved its hands and lips while the participants chatted on the computer. Furthermore, in one condition, a male voice was generated using the Microsoft speech platform while the agents produced facial expressions.

Three conditions were used to test our hypothesis (see Figure 2). In the “no suggestion” condition (Group SSA, $n = 18$), participants were given no suggestions without any affective expressions. In the text suggestion condition (Group SSA+T, $n = 20$), participants were given suggestions via textual prompts with affective feedback according to the affective model. In the text and voice suggestion condition (Group SSA+TV, $n = 20$), participants were given no suggestions via textual prompts but rather via audio output. In the SSA+TV condition, participants wore headphones to listen to the responses from the agent.

In the pretest and posttest, participants were asked to describe the meaning of the same technical words. As in Hayashi (2012 a), the results of the pretest and posttest were then compared to find out how the different conditions facilitated participants’ learning of the concepts. In the comparison, descriptions were scored in the following ways: one point was awarded for a wrong description or no description, two points for a nearly correct description, three points for a fairly correct description, four points for an excellent description, and five points for an excellent description with concrete examples. It was judged that the greater the difference in scores between the two tests, the greater the effect of the explanation activity. Two coders then coded the results, and their correlations were 0.67. The coders discussed their results before making any final decisions. The pretest and posttest scores were used to assess the degree of learning performance.

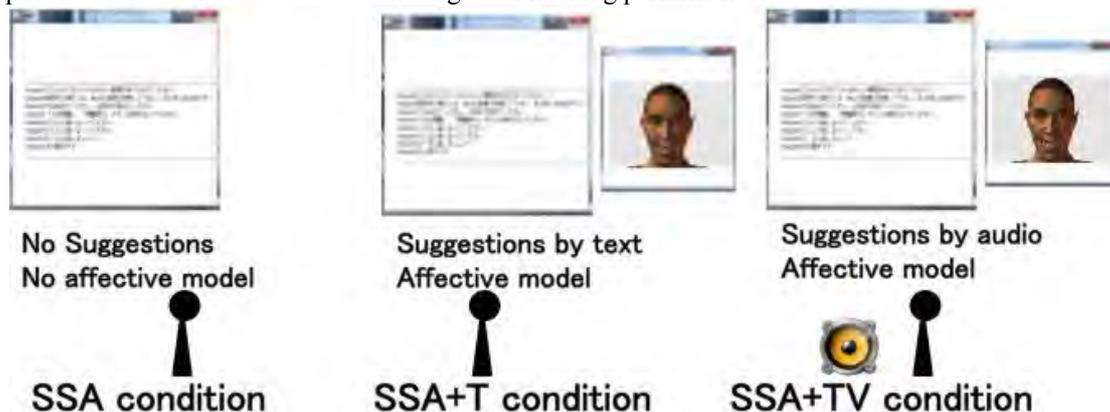


Figure 2. Experimental Conditions.

Results

Figure 3 shows the results of the pretest and posttest for the term “schema.” The vertical axis represents the average scores of the tests for the three groups at the two different test times. A statistical analysis was performed using a 2 (evaluation test: pretest vs. posttest) \times 3 (conversational agent condition: SSA vs. SSA+T vs. SSA+TV) mixed-factor analysis of variance (ANOVA).

There was a significant interaction between the two factors ($F(2, 55) = 12.457, p < .01$). First, an analysis of the simple main effects was done on each level of the conversational agent factor. In the SSA, SSA+T, and SSA+TV conditions, average scores on the posttest were higher than on the pretest ($F(1, 55) = 99.604, p < .01$; $F(1, 55) = 53.616, p < .01$; $F(1, 55) = 8.899, p < .01$, respectively). These results showed that the explanation activity had an effect on learning.

Next, simple main effects were analyzed according to evaluation test time. In the pretest, there were no differences between conditions ($F(2, 110) = 0.48, p = .95$), indicating no differences between participants before conducting the experiment. On the other hand, differences between conditions were found in the posttest ($F(2, 110) = 23.599, p < .01$). Post-hoc analysis on the posttest was conducted through Ryan’s method. Results indicated that the average score of the SSA+TV condition was higher than that of the SSA+T and SSA conditions ($p < .01$ for both), and the average score of SSA+T was higher than that of the SSA ($p < .01$). The difference in scores between the SSA+TV and SSA+T conditions in the posttest indicated that using different communication channels to explain a concept to a partner and receiving learning prompts from a PCA are useful to facilitate participants’ understanding of the concepts, compared with using the same communication channels.

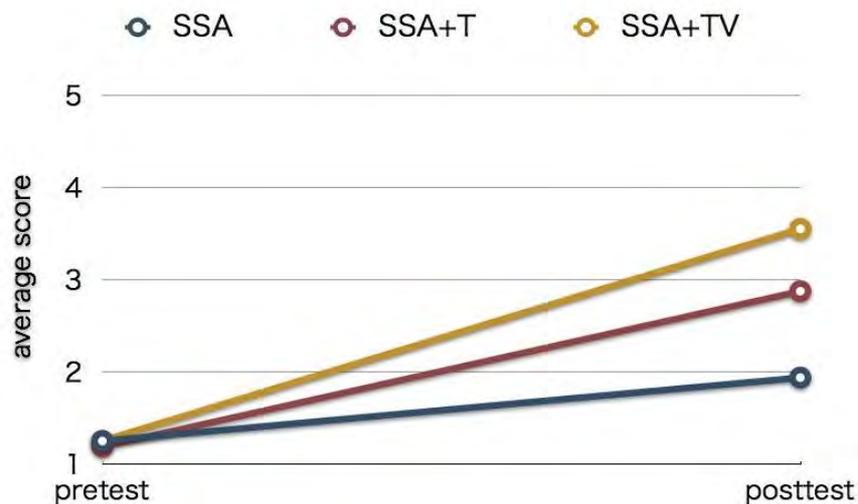


Figure 3. Results of ‘schema’.

Figure 4 presents the results of the pretest and posttest for the term “long-term memory.” The vertical axis represents the average scores of the pretest and posttest for the three conditions. The same type of statistical analysis was performed as above.

There was a significant interaction between the two factors ($F(2, 55) = 10.143, p < .01$). First, simple main effects were analyzed according to conversational agent. In the SSA+T and SSA+TV conditions, the average scores on the posttest were higher than on the pretest ($F(1, 55) = 62.662, p < .01$; $F(1, 55) = 23.673, p < .01$, respectively). In the SSA condition, the average scores did not differ ($F(1, 55) = 0.127, p < .01$). The increases shown in the SSA+T and SSA+TV conditions show that the explanation activities produced a learning effect related to affect.

Next, simple main effects were analyzed according to evaluation test. In the pretest, no differences were found between conditions ($F(2, 110) = 0.022, p = .97$), indicating no differences between participants before conducting the experiment. There were differences between conditions in the posttest ($F(2, 110) = 16.535, p < .01$). Post-hoc analysis on the posttest was conducted using Ryan’s method. The results indicated that the average score of the SSA+TV condition was higher than that of the SSA+T and SSA conditions ($p < .01$ for both), and the average score of SSA+T was higher than that of the SSA ($p < .01$). This result was consistent with the results of the analysis of “schema.”

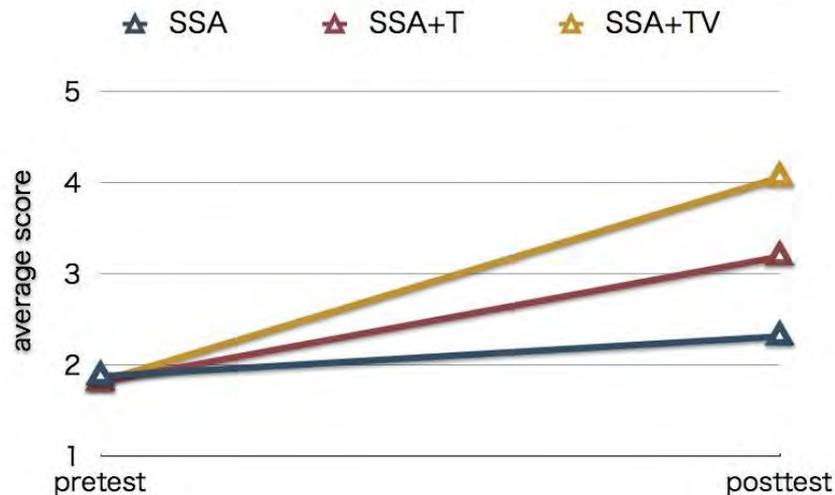


Figure 4. Results of 'long-term memory'.

Discussion

Affective factors and communication channels

The results of this study have several implications on the advantages of using pedagogical agents and providing feedback through different communication channels. First, the SSA+TV and SSA+T conditions had greater influences on students' performance than the SSA condition. This indicates that the use of an affective model has a strong effect on the performance of learning activities. This result is consistent with the results of a previous study by Hayashi (2012 a). In that study, agents with positive expressions had greater influences on performance compared with agents with no expression. However, because Hayashi (2012 a) did not conduct a direct comparison between positive affective agents and text-only prompts, the present study shows results that are more reliable on the advantages of using affective conversational agents during learning activities. Furthermore, the present study used more sophisticated affective models in the conversational agents. Although pedagogical agents have great potential, they should be modeled with parameters that are more detailed and based on human cognition.

The experimental results also showed that participants performed better in the SSA+TV condition (receiving verbal comments from the agent) compared with the other two conditions. According to the split attention method and the multimedia design model, using different communication channels for the learner's conversations and the suggestions from the agent is a good pedagogical method. Our results supported the notion that using different communication channels enables learners to pay more attention to agents' suggestions, and encourages them to consider the terms according to its comments. Post-hoc analyses showed that agents' comments facilitated students' use of related conceptual terms, and allowed them to understand the target keyword from a different perspective (see Table 1 for an example). Furthermore, learners—especially those not trained to give effective explanations—confronted their difficulties in giving explanations by asking appropriate questions. Therefore, they were concentrated on their partner's text messages, which occupied their visual attention. Table 2 shows some examples of participants' dialogue, where one student failed to pay attention to the pedagogical conversational agent.

In previous pedagogical agent studies, some attempts have been made to understand how an environment rich with multiple communication channels could facilitate the learning process. Unfortunately, some of these investigations showed that pedagogical agents had no significant learning effects. Moreno and Mayer (2002) conducted an experiment using a pedagogical agent named "Herman the bug," and presented information to learners via a desktop computer or head-mounted display. The head-mounted display condition was used to examine whether virtual reality could lead to better learning results, as it may encourage learners to engage in more active cognitive processing. However, the results yielded no difference between the virtual reality conditions (using a head-mounted display while walking or sitting) and the desktop computer condition in performance on retention or transfer tasks (Moreno & Mayer, 2002). These results are likely because virtual reality distracts the learner from the learning task. Therefore, additional features such as more technically complex learning environments do not necessarily facilitate learning. Another explanation is that the way the information was presented resulted in cognitive overflow. The present results indicated that if information is provided in a cognitively economical way, such as using multiple communication channels and splitting attention, computer-based learning with agents can be a powerful learning tool.

Table 1: Example dialogue of participants interacting frequently in SSA+TV condition.

Speaker	Dialogue
Participant A	“In the case of short term memory, the amount of the information affects memory, right?”
Conversational Agent	“Good! You’re using important keywords such as ‘short-term memory’!”
Participant B	“Oh. So, long-term memory is the opposite and we can remember more.”
Participant A	“I’m not sure, but I think so...”
Conversational Agent	“Nice, keep on it guys! Our goal in this activity is to reach a better understanding of this term by explaining it to each other.”
Participant B	“Then how about length of time...by using long-term memory, can we remember information for longer?”

Table 2: Example dialogue of participants not paying attention to the agent in SSA+T condition.

Speaker	Dialogue
Participant B	“Okay, so please continue explaining the term ‘schema’.”
Participant A	“They influence what we look for in a situation.”
Conversational Agent	“Good job! However, a couple of times you used the same words that were written in the example. Try to use your own words!”
Participant A	“We are inclined to place people who do not fit our schema in a special or different category....”
Participant B	“I think it’s better to not copy the sentences from the examples.”
Participant A	“Oh, I missed the instructions.”
Participant B	“The computer agent said that.”
Participant A	“I wasn’t paying attention to it at all. Oops...”

Conclusion and future work

The present study investigated the effectiveness of a conversational agent in a collaborative activity, where paired students explained to each other the meaning of several psychological terms to improve their understanding. The agents were used to encourage and facilitate students’ interactions through both verbal and visual output. The experimental results suggested that affective conversational agents using multiple communication channels can help trigger a deeper understanding of a concept when attempting to explain that concept. This gives us a new perspective on how to design pedagogical agents for collaborative activities such as giving explanations to others.

The present study used conversational agents that exhibited only positive affective expressions; future studies could use expressions that are more specific or according to personal preferences. Those personal preferences may be based on social constructions such as gender and culture. In future studies, these preferences for affective expression should be investigated and implemented into the system to produce more effective learning. For example, Kim, Baylor, and Shen (2007) found that learners had positive impressions of male agents with positive expressions than of female agents. This indicates that social stereotypes in the real world are applied to the agent-learner relationship.

Acknowledgments

This work was supported,(in part) by 2012 KDDI Foundation Research Grant Program and the Grant-in-Aid for Scientific Research (KAKENHI), The Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXTGrant), Grant No. 25870910

References

- Baddeley, A. D., & Hitch, G.J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 8. London: Academic Press.
- Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education*, 15(1), 95–115.

- Bower, H. G., Forgas, P. J. (2001). Mood and social memory. In J. P. Forgas (Ed). *Handbook of affect and social cognition* (pp. 95–120). New Jersey: LEA.
- Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of Learning Sciences*, 2(2), 141–178.
- Bruner, J. (1999). Postscript: Some reflections on education research. In E. C. Lagemann & L. S. Shulman (Eds.), *Issues in education research: Problems and possibilities* (pp. 399–409). San Francisco: Jossey-Bass Publishers.
- Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332.
- 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, 145–182.
- Coleman, E. B. (1998). Using explanatory knowledge during collaborative problem solving in science. *The Journal of Learning Sciences*, 7(3 & 4), 387–427.
- Graesser, A., & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist*, 45(4), 234–244.
- 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.
- Hayashi, Y., & Miwa, K. (2009). Prior experience and communication media in establishing common ground during collaboration. In *Proceedings of the 31th Annual Conference of the Cognitive Science Society*, 528–531.
- Hayashi Y. (2012 a). On pedagogical effects of learner-support agents in collaborative interaction. In S.A. Cerri & B. Clancey (Eds.): *Proceedings of the 11th International Conference on Intelligent Tutoring Systems (ITS2012)*, Lecture Notes in Computer Science, 7315, 22–32.
- Hayashi Y. (2012 b). Designing affective pedagogical agents: How learners' and agents' gender and age influence emotion in an online tutoring task. In *Proceedings of the 20th International Conference on Computers in Education (ICCE2012)*, 46–50.
- Holmes, J. (2007). Designing agents to support learning by explaining. *Computers & Education*, 48(4), 523–547.
- Kim, Y., Baylor, A. L., & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning*, 23(3), 220–234.
- King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal*, 30, 338–368.
- Lave, J. (1987). *Cognition in practice*. New York: Cambridge University Press.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia learning environments: Role of methods and media. *Journal of Educational Psychology*, 94(3), 598–610.
- Miyake, N. (1986) Constructive interaction and the interactive process of understanding. *Cognitive Science*, 10(2), 151–177.
- Mousavi, S. Y., Low, R. & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87(2), 319–334.
- Okada, T., & Simon, H. (1997). Collaborative discovery in a scientific domain. *Cognitive Science*, 21(2), 109–146.
- Salomon, G. (2001). *Distributed cognition: Psychological and educational considerations*. New York: Cambridge University Press.
- Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science*, 26(4), 469–501.
- Sweller, J., Chandler, P., Tierney, P. & Cooper, M. (1990). Cognitive load as a factor in the structure of technical material. *Journal of Experimental Psychology: General*, 119(2), 176–192.

An Adapted Group Psychotherapy Framework for Teaching and Learning About CSCL

Yotam Hod, Dr. Dani Ben-Zvi, University of Haifa, Israel
Email: yotamhod@edtech.haifa.ac.il, dani.benzvi@edtech.haifa.ac.il

Abstract: This theoretical paper presents a framework adapted from group psychotherapy (GP) for teaching and learning about computer-supported collaborative learning (CSCL). The framework is used both in the design of a learning community and to explain how aspects of learning occur within it. The design, which we enacted in a graduate level course that taught about CSCL, is based on content-process integration, group cohesion, and reflection. The learning is based on transference into a microcosm and developing the intentionality to change. We use narrative illustrations from moderators' and students' discourse to show how this framework impacts teaching and learning about CSCL.

Introduction

Since its inception, the scientific CSCL community has shown substantial interest in implementing its ideas in practical settings, particularly in higher education (Srijbos, Kirschner, & Martens, 2004). A great deal has been published about various educational designs that employ CSCL, with the goal of getting learners to collaborate in ways that are supported by technology (e.g., De Graaf, De Laat, & Scheltinga, 2004). In most of these cases, the content that is collaborated upon is external to the field, such as in mathematics, and the learning process is *through* CSCL. While such research is certainly important, more can be done to build upon existing research in teaching and learning *about* CSCL (e.g., Ronen-Fuhrmann, Kali, & Hoadley, 2008), as opposed to *through* it.

Further researching this particular niche in CSCL involves examining what material should be studied during graduate courses about CSCL as well as the appropriate pedagogical design to foster learning. Regarding the content, the big ideas of CSCL have to do with metaphors or conceptions of learning such as participation (Sfard, 1998), knowledge-building (Scardamalia & Bereiter, 1994) and collaboration (Stahl, Koschmann, & Suthers, 2006). Likewise, developing expertise in CSCL involves studying rich academic literature on a range of topics that include learning, research methods, and technology, just to name a few. While the learning material is not the prime focus of this paper, the epicenter upon which to base it ostensibly should start with academic material published from the CSCL community. From there, it can span out in nearly endless directions to include more peripheral ideas.

Regarding the process of teaching and learning about CSCL, there are good pedagogical reasons to integrate the process with the content. In approaches that aim for students to learn *through* CSCL, the content can be in any discipline so long as the learning process is collaborative and supported by technology. In approaches that aim for learning *about* CSCL, an inverse of this content-process relationship exists. The content, however debatable, must be about CSCL and is therefore fixed; the pedagogical approach is variable. Nevertheless, research into teaching programs show that methods are too often disconnected with what they are trying to teach (Bransford, 2000) and there is growing evidence that they should be integrated (Darling-Hammond, 2000). Therefore, there are good reasons for those seeking to teach about CSCL to integrate content and process. In this way, students simultaneously study CSCL content and develop a rich repertoire of experiences about it.

Relations to Group Psychotherapy

Much like in teaching and learning about CSCL, the field of group psychotherapy (GP) also deals with the integration of content and process. In GP that takes a here-and-now approach, there are two tiers that are critical for the therapy to occur. The first tier has to do with taking an interest in the here-and-now events of a group, or the present process the participants engage in. This process, in effect, becomes part of the content of the second tier, which deals with illuminating this process together with the group. The content is integrated with substantive issues in GP, such as power and dominance or psychic functions, that have to do with interpersonal relationships of the participants (Yalom, 2005). Thus, while the external content of GP and CSCL are different, both have a shared characteristic of integrating their respective content with the process of interpersonal learning.

There are many implications of the shared content-process integration found in GP and teaching and learning about CSCL. These seem to hinge on the idea of reflection. While reflection is a fundamental concept and practice in both, the separation of the fields leads them to use different jargon to explain its role, and research upon which to base findings. Yet, they have a strikingly great deal in common.

In CSCL, reflection is based on literature with a long and well-established history (e.g., Chinn, Yost, Sentner, & Forlenza-Bailey, 2000). Developing reflective students is associated with ideas that have to do

with taking responsibility over one's learning such as meta-cognition (e.g., Brown & Campione, 1994) and intentional learning (Scardamalia & Bereiter, 1994). Indeed, these have been cornerstones in CSCL research (Stahl, Koschmann, & Suthers, 2006). In more contemporary CSCL studies, reflection has been used to develop group members' awareness of their individual and group behavior creating a bridge between social and cognitive processes in collaborative learning (Phielix, Prins, & Kirschner, 2010). In all cases, the primary idea has to do with organizing instruction, through reflective discourse, so that learning can be controlled and intentionally changed.

In GP, reflection is essential to the therapeutic process. The group setting serves as a microcosm, whereby each participant *transfers* (1) their own attitudes, understandings, and behavior in their ordinary life to the group situation. Participants are guided to speak about their feelings and experiences as they relate to the ongoing events during meetings. This is referred to as the *here-and-now* focus. Reflection on the process helps the participants recognize their own attitudes, understandings, and behavior, in relation to others and their lives in general. This allows them to decide if they are satisfied with what they bring to the microcosm, and if not, empowers them to exercise the *will* to change it (Yalom, 2005). Thus, despite the different jargon, the similarity between the two fields is evident. In both cases, deep learning is based upon the intentionality of the learner to make changes, which occurs as a result of a reflective process that they engage in.

In addition to the content-process integration as well as the role of reflection, one final similarity rests with group cohesion, or generally the sense of membership in the group (Yalom, 2005). In GP, this has been identified as an elemental therapeutic factor. In CSCL, it is possible to find some corollaries to this, such as with the study of the design of social infrastructure for effective collaboration (Bielaczyc, 2006). Still, there have been calls in CSCL to further examine "the social (psychological) dimension of social interaction for group forming group structure and group dynamics" (Kreijns, Kirschner, & Jochems, 2003, p. 343). While GP deals with group cohesion more directly as part of its theory and practice, it is clear that it is important to both fields.

While the similarities between GP and teaching and learning about CSCL are striking, there are important differences too. As mentioned already, the content of GP and CSCL differ, each with its own interests. Other key differences have to do with the goals, participants, and settings of each. Even though contemporary GP has moved away from curative goals and today aims for learning and growth like in teaching and learning about CSCL (Yalom, 2005), the goals are therapeutic. In comparison, participants in courses about CSCL are typically in graduate institutions who aim to develop their professional interests and skills. As such, the terms of the agreement and expectations between therapist-client versus teacher-student are different.

Taking these differences into account, we adapted a theoretical framework from GP and applied it to teaching and learning about CSCL. The design based on the adapted framework had three inter-related components: (a) content-process integration, (b) promoting group cohesion, and (c) reflection. We applied this to a blended learning community as part of a graduate-level course. We focus on describing this "grain of sand" seeking to illustrate the learning that emerged about CSCL as a consequence.

Implementation of the Framework

In this paper, we illustrate our preliminary findings in a case study of a blended graduate course as part of the Educational Technologies Graduate Program at the University of Haifa (UH), Israel. The course, titled "Challenges and Approaches to Technology-Enhanced Teaching and Learning" (CATELT) served as an introductory requirement for all students in the multi-year program. CATELT's primary goal was to teach the basics of the learning sciences, with a special focus on CSCL. Originally enacted during the 2006-2007 academic year, five previous annual iterations were progressively refined to form the version (2011-2012) that is presented in this paper.

CATELT 2011/2 included 14 students, aged 23-54, who had no prior relationship and were selected based on their academic records and interest in educational technologies. The primary course moderator, who was accompanied by a teacher's assistant (TA) and a researcher, was a senior lecturer at the UH. In addition to his research interests in innovative educational technology and collaborative learning communities, his prior training in group counseling was relevant to enacting the adapted GP framework.

Data collection and analysis

To draw conclusions, rich micro-level data from ftf class interactions, the online collaborative editing environment (Wiki), personal interviews and observations of moderators/TA meetings were collected throughout the full duration of the course. These were analyzed micro-analytically (e.g., Meira, 1998; Siegler & Crowley, 1991). Findings based on video and audio recordings of the ftf interactions, as well as all written material from the Wiki, were reviewed by the researcher and triangulated by a committee of expert and novice peers. Furthermore, conclusions were reached only after multiple sources of data validated a specific result.

Design

CATELT had a blended course structure that was arranged as a series of 26 alternating and interconnected face-to-face (ftf) and online interactions. These occurred in a small classroom and online collaborative editing environment (Wiki) (Ben-Zvi, 2007). Each pair of interactions occurred over one week, making the course duration a thirteen-week semester. Several activities were run during each interaction, as seen in Figure 1. Overall, reflective activity in the form of groups reflection sessions (GRS) consisted of nearly one-third (758 out of 2410 minutes) of all ftf interaction; reflective diaries and moderator reflective messages were recurring and central components of online interactions. The remaining components were either collaborative activities on the course content, or aimed to build group cohesion between group members.

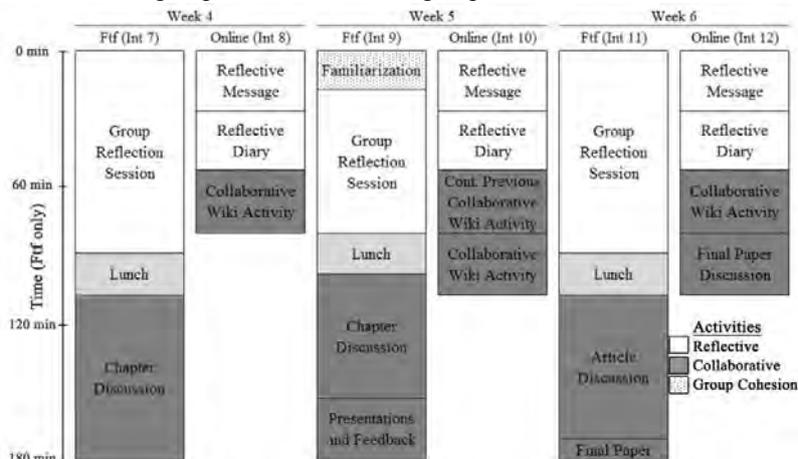


Figure 1. Sample of activities in CATELT's adapted GP framework.

Component 1. Content-Process Integration

Content and process in CATELT were integrated such that students could examine their own collaborative learning process in a technology-enhanced environment. To do this, the learning process that emerged and was practiced became the basis of both ftf and online discussions, primarily in reflective activities. Likewise, CSCL content that was relevant to the students' experience of learning was collaboratively studied. These interacted symbiotically. As students developed more collaborative experience, they could understand CSCL content more deeply; as students understood CSCL content more deeply, they could collaborate more effectively. While this integration in studying about CSCL is similar to that in GP, two key adaptations to the framework were made. First, the actual content studied dealt with CSCL, instead of substantive content of GP. Second, the design was adapted so that students could participate in a process of CSCL instead of GP.

CSCL content was introduced in three broad themes that were addressed sequentially, in relation to the expected learning trajectory of the students. The first theme focused on the individual learner, as the students entered the course primarily with traditional views of learning. Specifically, the first four chapters from "How People Learn" (Bransford 2000) were studied collaboratively to deepen their understandings of learning, mainly from an individualistic perspective, while providing them with collaborative experience. As the students gained this collaborative learning experience, the course moved to the second theme of learning in a community. This consisted of academic articles that had an important influence on CSCL (e.g., Brown & Campione, 1994; Brown, Collins, & Duguid, 1989). The content of the final theme maintained this collaborative focus but in ways that were supported by technology (e.g., Scardamalia & Bereiter, 1994), as the students increasingly engaged in the course Wiki. As such, the content was planned to parallel the broad process of the students' learning.

Having students experience CSCL was also integral to the GP framework. Mediated ftf sessions alone, like in GP, would not have sufficed to make the experience one that entailed CSCL. As such, a number of ftf and online activities were supplemented to give the students a range of CSCL experiences. These were primarily designed on the course Wiki, which the students needed to complete throughout the week. As part of their online assignments, students were typically asked to work in self-created teams and co-author content pages on the Wiki. These were based upon key concepts and ideas found in their assigned readings. Although some Wiki-based activities were completed individually, such as the writing of final course papers, all of these had components, such as peer-feedback, that encouraged the students to collaborate intensely.

Assignments generally had ill-structured designs, leaving the group members to deal with challenges like forming groups on their own and deciding for themselves where to stop. These assignments were designed this way purposely so that students could engage in CSCL experiences that brought about difficulties and challenged them to think deeply about the process they were engaged in. Moreover, the moderators' interventions challenged the students to be in a continual search for improvement and deepening of

understanding which took precedence over praising students for a job well done. This was important because it focused on the process of collaboration over achieving a final product, raising important issues that could be related to CSCL content, such as that learning never ends (lifelong learning) and that learning should be driven by the student (learner-centered design). Designing and intervening for serious challenges gave the group experiences that they could consider, becoming an important source of content in reflective activities. This also explains the reference to *challenges* in the course title.

Component 2: Promoting Group Cohesion

Cohesion building in the group served two primary purposes. The first was to get the students, to the maximum extent possible, to feel a part of and participate in the community. The second purpose was to generate important collaboration-related process issues as content. To fulfill these purposes, several specially designed activities as well as techniques throughout the course were enacted.

Familiarization activities were the most direct method to promote group cohesion. Five different familiarization activities were run towards the start of the semester (weeks one, two, three, five, and nine). Each of these used some artifact (e.g. image, textual prompt) and involved some form of peer-to-peer conversation so that students could more freely share their personal feelings, attitudes, and perspectives with each other in a trusting and safe environment. Additionally, various techniques were used by the moderators to encourage group cohesion, both within and outside of the familiarization activities. For example, at the start of peer discussions students were regularly asked to talk with someone who they didn't know yet to reduce the fear of approaching an unknown person and to deepen their relationship with them. At times, students were encouraged to sit in different places in the room, which the moderators themselves often modeled by switching places themselves. The moderators also discouraged particularly active students from dominating by encouraging them to wait and listen; invited members to join the community by welcoming them warmly on the course Wiki; and built a caring and trusting community with empathy by demonstrating personal knowledge of the students' perspectives. Together with the familiarization activities, these many techniques promoted group cohesion.

In addition to directly building the students' sense of membership to the group a second purpose of promoting group cohesion was to give students experience on what it meant to build a learning community. By being engaged in a process of building group cohesion, they could examine some of the related challenges and issues. For example, one student who reflected upon the first ftf meeting, which was designed largely to promote cohesion, thought of the implications in relation to her own experiences as well as educational design in general.

Patricia (2): The excitement of first grade won't return but this was close... I left the class with a feeling of, "if a student didn't have a teacher that wanted them to feel, this is something they need to go through! So they will have an appetite for more". It was pleasant for me to be in class, many smiles, a lot of listening, a lot of patience, and attempts to calm fears. We the students were in the middle, we told about ourselves and listened to others, who listened to us and cared for us. (3)

As such, cohesion building was not just a mechanism to facilitate the group's functioning but was also part of the content-process integration. In this way, like the other collaborative activities, the process of building group cohesion was a basis for reflection and learning.

Component 3: Reflection

Reflective activities in CATELT had the purpose of creating rich discourse in two tiers. The first tier focused on the here-and-now events of the group, and the second tier on illuminating this process. There were three types of reflective activities in CATELT that enacted these two tiers. These included group reflection sessions (GRS), online reflective diaries, and online moderator reflective messages. While reflection in CATELT was at the core of its resemblance with GP because of these tiers, several adaptations were made.

Among all CATELTs activities, GRS had the most direct resemblance with GP because of the format, which is typically run in small mediated ftf groups (Yalom, 2005). GRS elicited first tier here-and-now discourse based on the group's present and past experiences. To do this, the moderator used a number of techniques all guided by an underlying principle of focusing on the here-and-now. For example, when one student began talking abstractly and analysing his experience, the moderator abruptly stopped and refocused him:

Moderator: You can't make generalizations in a reflection discussion because a large portion of us... act from different paradigms that they bring from their previous lives... which is why I'm asking you to explain your paradigm based on exactly what is happening to you.

As an adaptation to GP, CATELT had many collaborative activities, particularly during the week on the Wiki. These supplemented the GRS discourse, as students were encouraged to talk about the events that occurred collaboratively over the week. While this adaptation broadened the experiences that could be discussed during reflection sessions, it was still operationally consistent with GP, as “past events of the therapy group are a part of the here-and-now” (Yalom, 2005, p. 162).

A second adaptation had to do with the second tier. While both GP and CSCL deal with inter-group relations and therefore have a great deal of content in common CATELT’s focus on CSCL changed the focus of the process commentary. For example, during one reflection session the group discussed their experiences of collaborative editing on the Wiki. One student raised the challenge of editing others’ work aware of the different styles of writing but not wanting to offend a peer by making changes. At this point, Phil offered a solution:

Phil: There is a simple solution: You split the sections between the participants, you upload all of it, and then you decide that one person is the editor of everything.

The moderator used this opportunity to make process commentary based on CSCL content. Instead of focusing on Phil’s motivations for offering advice as may be done in GP the moderator instead focused on his understanding of cooperation and collaboration.

Moderator: This is one of the solutions [to how the group should work together]. You come from a perspective as if you know the answer. But this solution is problematic because when you talk about cooperation versus collaboration, the meaning of collaboration is more or less what Tina and Jane described: They worked nearly completely synchronously, on every word and statement. That is collaboration - a shared product that is a synergy of their work. You offer a different model.

Together, the GRS included two tiers that are found in GP. The first tier elicited the group’s here-and-now experiences as they collaborated, and the second tier illuminated this process in relation to CSCL content.

In addition to the GRS, online reflective diaries and online moderator reflection messages contributed to the reflective discourse. The online reflective diary was a continuation of the GRS, where students were asked to reflect upon their learning from the ftf meeting that passed. They wrote these in the community section of the Wiki, where other members of the class could see. This encouraged a continuation and further deepening of key ideas raised during the ftf meeting. Students were also encouraged to create conversations in each others’ reflective diary discussion pages to deepen dialogues by seeding, migrating, and mutually appropriating ideas (Brown & Campione, 1994). To strengthen the connection between these experiences and GRS, the course moderators would sometimes take meaningful conversations and use them as part of ftf reflections. While the writing of personal diaries after sessions is a common practice in GP, the public online format of these reflections, the discussions around them, and their connection to ftf meetings are adaptations.

Like the reflection diaries, another common practice in GP is for the therapist to write and share their own reflection following each ftf meeting (Yalom, 2005). In CATELT, this was done in the course Wiki. The messages, which were written following reflective meetings by the moderators after each interaction, presented the moderators’ reflection of what was meaningful. They also encouraged exemplary behavior and built on related CSCL content that was brought up during ftf meetings.

Transference and Change

Just like in the GP framework, a critical aspect of CATELT had to do with developing group cohesiveness so that the members could re-experience their attitudes, behaviors, and understandings from their daily lives into the group setting, a process known as transference (Yalom, 2005). The microcosm that developed in CATELT, in comparison to GP, had a more direct focus on collaboration. As such, students had a chance to play out their own past experiences in the collaborative learning microcosm that was formed. These were examined as part of the reflective discourse. As students raised their awareness over who they were as learners along with their role in the group, they were able to intentionally make changes to the way they collaborated.

Transference into a cohesive CSCL microcosm occurred within the context of group cohesiveness. Examples of group cohesiveness could be found by the emergence of group behaviors as the course developed. For example, the students began organizing their own shared lunches, had informal meetings before class over coffee, and were very active online during the week particularly commenting in each others’ diary discussion pages. Evidence came about explicitly as well, such as when one student shared her feelings about the group as part of a GRS. The discussion revolved around the metaphor of a turtle:

Moderator: What is your protection? If you were a turtle, what would protect you?

Beth: ... You are all my family, yes, in the past few weeks I feel this way: Protection.

Given the cohesiveness of the group, transference effects were found prevalently. For example, Patricia was a self-described competitive person, who worked in the hi-tech industry for many years. Showing that her behavior in the course was transference from her “real” world Patricia started the course in the same manner. For example, she favored learning from an authority over her peers. She also made many arguments rejecting the idea of collaborative learning on the grounds of practicality:

Patricia: In a workplace when you come to work, the work is very important at the end of the day. And when I think about these things, I say... what is happening here is an island. ...In real life, ultimately you go to your job and you need to sit with yourself and work.

Because of Patricia’s background her primary preoccupation in the course related to her individualistic, competitive background. The cohesiveness of the microcosm allowed her to transfer her experience to CATELT.

Individual Change

The collaborative microcosm that developed, which formed part of the content of reflective activity, led to discourse where the students examined and tested their own and others’ behavior and identity as learners in the course.

Patricia: From the discussion on the Learning Sciences – [I reached the conclusion] that there is a need to learn more about how we learn and understand. That I need to learn the way I learn myself, what actions I do while reading, when I understand better, how I overcome a text that isn’t clear, and an open question for me - learning in a community, looking at ourselves as a learning community. I am so much of a “lone wolf” in this respect...

Such discourse allowed students to become aware of themselves as collaborative learners, and suggest new ways to participate in the group. For example as part of Patricia’s competitiveness she was very dominant in the group towards the beginning of the semester. By becoming aware of her behavior through reflection, she appeared to show greater sensitivity towards others by taking a less dominant role:

Patricia: In the last few lessons in the reflection, especially in the beginning, I spoke a lot and now I talk less. This is because sometimes I want to hear others. I don’t want to be in a situation – I admit that I talked a lot – and I don’t want to be in a situation where I am always talking. I really want to hear others.

Patricia’s change coincided with an epistemological shift. At first she questioned the logic and effectiveness of learning from peers, arguing that bad ideas could be seeded and migrated. As such, she resisted collaboration with peers on the Wiki. However, upon reflecting on her collaborative experience with her peers, she showed a greater openness and desire to listen and learn from others.

Patricia: I assume that if someone wrote something, they thought about it, and this is what they understood. [I want] to see why they understood it this way because maybe I didn’t understand it.

While Patricia’s increased sensitivity and epistemological understanding was by no means linear nor absolute, it did show that she was engaged in meaningful discourse about herself. Such personal awareness gave her the will and desire to make these intentional changes.

Group Change

The intentional changes also occurred at the group level, which manifested, for example, as a group norms discussion. From the start of the semester, at any time when the group complained about their own collaboration, the moderator encouraged the students to take responsibility over their own norms. Supporting this call, a blank norms page on the Wiki was posted in the main navigation bar. This page included links to previous iterations of the class’ norms pages, suggesting that this stage would be reached later in the course. The moderator did not assign such a task, however, waiting for the group to be ready to take responsibility for it. Indeed, in previous iterations of the course, the group norms discussions manifested in different ways. As members in the group continued to reflect upon themselves and deepen their understanding of collaboration, the idea coalesced that the group norms that emerged were not consistent with their desired behavior as a group.

Patricia: On Wednesday, it became clear to me that I wasn't alone. That a few others thought that there wasn't collaboration and that it is not possible to create collaboration in this way. Something else is needed.

The group, by consensus, chose to engage in a process of discussing their own norms so that they could alter their existing norms. With the support of the group, two student leaders requested time during the subsequent 4th interaction to lead a process where the group renegotiated their norms. They followed this with more discussions on the Wiki. Showing symbolically that they really understood the process orientation of their actions, they chose to work on their agreed upon norms in a page intended for collaborative discussion, instead of a regular content page.

Following the group norms discussion, the group maintained their collaborative focus for the remainder of the semester. For example, when one student in a later assignment suggested that the group divide up a task and complete the separate parts in small groups, the idea was rejected. One student wrote the following, and several others quickly followed in agreement, showing that the group had intentionally changed.

Patricia: Let's try real collaborative editing for once and not cooperative... not like Phil suggested (sorry Phil...). I mean that every person adds his insights to every paragraph and not that we shall break apart [the work] and each group writes something. That we already tried and we all thought this wasn't really collaborative editing...

Discussion and Conclusion

This theoretical paper presents a framework that is adapted from GP that can be used to design, enact, and explain individual and group changes when teaching and learning about CSCL. Content-process integration, group cohesiveness, and reflection are essential components for this to occur. The idea that a learning community can be viewed as a microcosm where transference effects take place, and where reflection empowers students to make intentional changes, is a novel approach that can impact the field. Indeed, our illustrations, while modest, show that there is potential in this framework for the design and promotion of deep learning. Namely, getting individual learners to make intentional changes to their epistemological ideas as well as sensitivity towards others, along with getting groups to move from cooperative to collaborative learning, is a potentially profound result. We posit that this is an important desired result of teaching and learning about CSCL, and a key to getting students to enter the CSCL and Learning Sciences communities.

Certainly, there are many aspects of our design and illustrations that are not considered as part of our adapted GP framework. For example, GP has identified therapeutic factors that we have not considered here, such as the instillation of hope (Yalom, 2005). We have focused on a simplified framework for GP to describe the central mechanisms of individual and group change and what we consider are the most critical elements in the design for this to happen. Our hope is that with more research, we can further develop this framework. Moreover, using design-based research methodology, we can systematically adjust and refine our model to gain a better understanding of how people learn about CSCL as they gain membership in the learning community.

Our next steps include trying to re-teach CATELT with a different moderator. This can help us isolate some of the characteristics of the moderator, such as personality and expertise, and focus instead on the role of the design. Likewise, we are continuing our retrospective analysis of CATELT, micro-analysing the data on related research questions that can provide a more complete description of the learning processes. We hope that through these sustained efforts, we can positively influence programs and courses seeking to teach about CSCL so students can be enculturated into the CSCL and Learning Sciences communities.

Endnotes

- (1) We use the terms *transfer* and *transference* in the context of its use in GP and not in relation to the concept as it is used commonly in the Learning Sciences (e.g., Bransford, 2000).
- (2) All course participants were designated pseudonyms to maintain their confidentiality.
- (3) All quotes were originally said or written in Hebrew, and have been translated for the sake of this paper. As part of our ongoing micro-analytic study, we closely examine the meaning of every word to make sure the translation is as close as possible to the original intention of the contributor.

References

- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, 13(1), 1-14.
- Bielaczyc, K. (2006). Designing social infrastructure: The challenge of building computer-supported learning communities. *Journal of the Learning Sciences*, 15(3), 301-329.

- Ben-Zvi, D. (2007). Using Wiki to promote collaborative learning in statistics education. *Technology Innovations in Statistics Education*, 1(1), 1-18.
- Bransford, J. D. (2000). *How people learn: Brain, mind, experience, and school* (Expanded ed.). Washington, D.C.: National Academy Press.
- Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229-272). Cambridge, U.K.: The MIT Press.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Darling-Hammond, L. (2000). How teacher education matters. *Journal of Teacher Education*, 51(3), 166-173.
- De Graaf, R., De Laat, M., & Scheltinga, H. (2004). CSCL-ware in practice. In J. Strijbos, P.A. Kirschner, and R. Martens (Eds.), *What we know about CSCL and implementing it in higher education* (pp. 201-220). Boston, Mass.: Kluwer Academic Publishers.
- Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior*, 19(3), 335-353.
- Meira, L. (1998). Making sense of instructional devices: The emergence of transparency in mathematical activity. *Journal for Research in Mathematics Education*, 29(2), 121-142.
- Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL-environment: Effects of peer feedback and reflection. *Computers in Human Behavior*, 26(2), 151-161.
- Ronen-Fuhrmann, T., Kali, Y., & Hoadley, C. M. (2008). Helping education students understand learning through designing. *Educational Technology*, 48(2), 26-33.
- Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences*, 3(3), 265-283.
- Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. San Francisco, CA.: Jossey-Bass.
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(2), 4-13.
- Siegler, R. S., & Crowley, K. (1991). The microgenetic method. *American Psychologist*, 46(6), 606-620.
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 409-426). Cambridge, UK: Cambridge University Press.
- Strijbos, J., Kirschner, P. A., & Martens, R. (2004). *What we know about CSCL and implementing it in higher education*. Boston, Mass.: Kluwer Academic Publishers.
- Yalom, I. D. (2005). *The theory and practice of group psychotherapy* (5th Ed.). New York, NY.: Basic Books.
- Yost, D. S., Sentner, S. M., & Forlenza-Bailey, A. (2000). An examination of the construct of critical reflection: Implications for teacher education programming in the 21st century. *Journal of Teacher Education*, 51(1), 39-49.

The Sequential Analysis, Modeling and Visualization of Collaborative Causal Mapping Processes and Effects on Causal Understanding

Allan Jeong, Woon Jee Lee, Florida State University
 Email: ajeong@fsu.edu, woonjee@gmail.com

Abstract: This paper describes a case study that illustrates how particular techniques and two developed software applications can be used to sequentially analyze, model, and visualize the processes and discourse student produce while working collaboratively in pairs to construct a causal diagram. The analysis was conducted at various levels in order to model students' map construction processes, the map construction processes in conjunction with students' discourse, and processes observed within each group versus across all groups. Transitional state diagrams produced from the sequential analysis of each group's behaviors revealed unique sequential patterns in the processes used between the three groups. These observed processes provide potential explanations for the observed differences in the accuracy of the maps between the groups. The implications of the findings and directions for further refinements to the techniques and software tools are identified and discussed in further detail.

One of the essential skills in solving complex problems is the ability to identify, articulate, and understand causal relationships between variables and/or events within a complex system. One way to accomplish this task is to construct a causal map. A causal map is a two-dimensional network of nodes and links that conveys the hierarchical and cause-effect relationships between events or variables within a complex system. Causal maps for example can be used to identify which variables exert direct/indirect effects on an outcome variable, which variables are to be viewed as root causes, and how the effects of variables on a given outcome are mediated by other variables. They can serve as a useful tool for scaffolding learning (Cho & Jonassen, 2002), especially so when students work collaboratively to construct a causal map (Nesbit & Adesope, 2006). Furthermore, causal maps can be used to assess students' causal understanding and systems thinking skills (Jeong & Shin, 2013; McClure & Bell, 1990).

Although various procedures have been developed to provide guidance on how to construct causal maps (Jonassen & Ionas, 2008; Bryson et al., 2004; Scavarda et al., 2006; Clarkson & Hodgkinson, 2005; Decision Systems Inc., 2012), there is little empirical research that have modeled the processes students use and that have identified the specific processes that create more accurate maps (Jeong & Lee, 2012). Based on a review of the empirical research and frequent reports on the high amount of variance in quality often observed across students causal maps, Ruiz-Primo & Shavelson (1996) concluded that maps should not be used in the classroom for large-scale assessments until students' facility, prior knowledge/skills with using maps, and associated training techniques are thoroughly examined. The point is that we do not yet know at this time to what extent students' causal maps (when used as an assessment tool) are measures of their causal understanding or measures of their causal mapping skills/processes. Furthermore, few studies have examined how peer interaction are integral to the processes students use when working collaboratively to construct causal maps, and how these processes affect map accuracy and/or causal understanding.

The purpose of this case study was develop, test, and illustrate a set of techniques and software tools that were developed and used to sequentially analyze, model, and precisely visualize students' map construction processes across multiple levels – map construction behaviors only, the interplay between map construction behaviors and student discourse, and the map construction processes within group versus across all groups. To illustrate our approach, this case study examined the following questions:

1. What sequential patterns exist in students' map drawing processes?
2. What sequential patterns exist between map drawing behaviors and students' dialog when working collaboratively on a shared causal map?
3. Which patterns help to explain observed differences in the accuracy of students' maps?

Method

Procedures. The participants were six graduate students (all female with ages ranging from 22 to 38) enrolled in a computer multimedia development course at a large university in the southeast region of U.S. in fall 2011. Students completed a 15-minute practice session on how to use the jMAP software (Jeong, 2012) to create causal maps and received from the instructor an introduction to causal maps with example applications. The students were paired up with another student, presented with a copy of the jMAP software with 15 pre-specified

events (Figure 1), and collaboratively constructed a causal map to identify chains of events and critical events that explains how and why a given multimedia tutorial (presented in a hypothetical scenario) was producing inferior learning. To correctly link the 15 given events into the correct hierarchical sequence/structure, students had to apply their knowledge of multimedia e-learning principles studied in the course. Each group was recorded on video camera.

The accuracy of each group's map was assessed in comparison to the instructor's map. In figure 2 is the instructor's map in jMAP viewed in relation to each group's map (dark green links = correct link with correct causal direction; light green = correct link with incorrect causal direction; gray = missed link; green halo = correctly identified root cause). Map accuracy was based on the total number of dark green links, nodes with green halos, and number of 1, 2nd, and 3rd order links that stem directly from a correctly identified root cause.

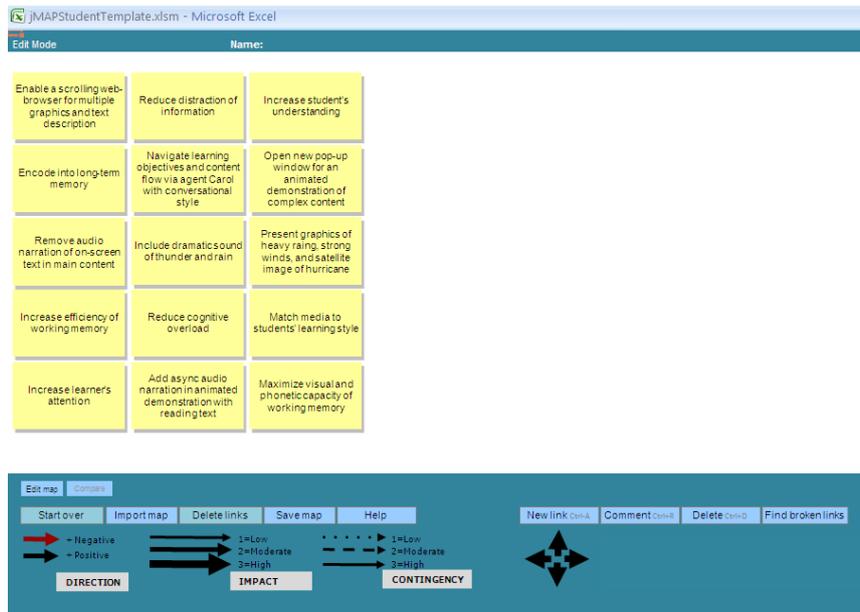
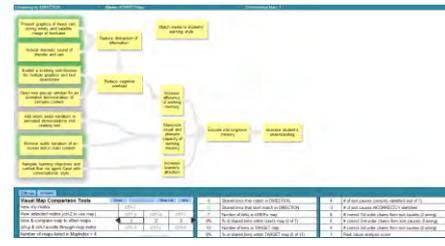
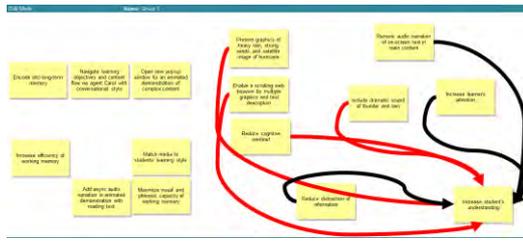
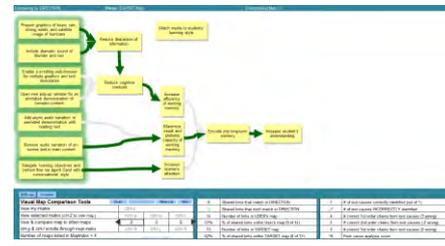
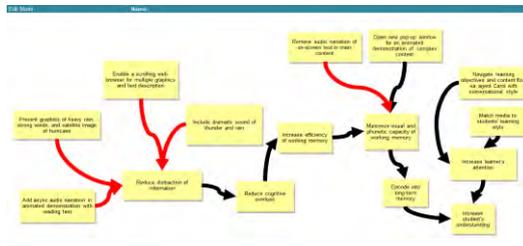


Figure 1. A screen capture of the initial causal map presented to students in jMAP.

Group 1
Low
accuracy



Group 2
High
accuracy



Group 3
Medium
accuracy

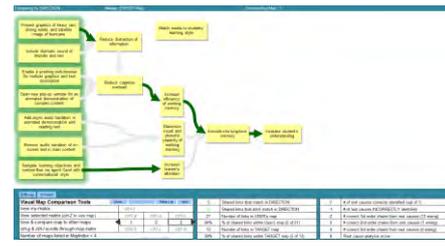
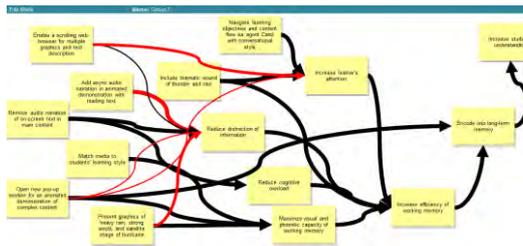


Figure 2. The maps of each group and the instructor's map viewed in relation to the group's map.

Data Analysis. The map drawing behaviors observed in the video recordings were coded into the six categories: PN - Position a node for the first time, RN - Reposition a node, AL- Add a causal link between two nodes, CA - Change the attributes of a causal link, DL- Delete a causal link, CL- Insert a comment to explain a causal link. The inter-reliability between two coders showed very good agreement (Cohen's Kappa coefficients $k = .926, .893$ and $.889$ for codes assigned to behaviors observed in Group 1, 2, and 3, respectively). A total of seven dialog moves were identified from a close review and analysis of the discussion transcripts (Table 1). The inter-reliability between two coders showed good agreement ($k = .712$ in Group1, $.734$ in Group2, and $.748$ in Group3). Using these seven dialog moves, student conversations were manually classified into six forms of discourse based on specific dialog move sequences (Table 2) to better examine the relationship between specific map construction behaviors and the discourse that took place prior to, after, and/or concurrently with each observed behavior. In the future, multidimensional scaling can be used to classify dialog move sequences with more scientific rigor.

Table 1: Seven codes used to classify dialog moves

Codes	Code Descriptions
EXP	Explain one's ideas in reference to specific nodes and their causal relationships.
DIR	Provide directions on how to construct the map (e.g., move the node, make the link, etc.) and how to use the mapping tool (e.g., "click the arrow button")
QUE	Ask questions. The question may require a specific answer, an opinion, or simply a confirmation (e.g., "right?").
AGR	Agree with partner's explanation or directions; positive responses to the confirmative questions.
DIS	Disagree with partner's explanation or direction, plus showing negative responses to the confirmative questions.
STA	Self-talk aloud to verbalize current actions performed on the causal map.
EVI	Provide evidence based on the reading materials. (e.g., it is saying that ...)

Table 2: Six forms of discourse based on observed dialog move sequences

Discourse	Description	Indicators in terms of dialog move sequences
Collaborative	Information and idea sharing	EXP→AGR→AGR
		EXP→AGR→EXP→AGR
		EXP→EXP→EXP→... →AGR
Interrogative	Exchange of questions & answers	QUE→EXP
		QUE→AGR
Argumentative	Conflicting viewpoints	EXP→DIS→EXP
		QUE→DIS
Confirmative	"Right?" and "Yes"	QUE ("Right?") →AGR ("Yes")
Explanation	Single explanation	EXP
Agreement	Single expression of agreement	AGR

For each group, the codes for the map behaviors were entered sequentially into a spreadsheet column with each code assigned a sequence number in the adjacent column based on their natural chronology (e.g., first action performed by the group is assigned the sequence number 1). This data was used to model the map drawing processes independent of the discourse that took place during the map construction process. To examine the interplay between the map construction process and discourse, the codes for both the map drawing behaviors and discourse were entered sequentially into another spreadsheet with each code assigned to a sequence number reflecting their chronological order. For behaviors and forms of discourse that occurred simultaneously and/or concurrently, these events were assigned the same sequence number.

Each of the resulting data sets were sequentially analyzed using the Discussion Analysis Tool (DAT) (Jeong, 2012b). DAT produced frequency matrices to compute transitional probabilities for each event pair. To determine which two-event sequences could be deemed to be a behavioral "pattern", the DAT software computed z -scores for each reported probability. Probabilities that were significantly higher/lower than the expected probability were identified by z -scores that were greater than the critical z -score of 1.64 at $\alpha = .10$ with observed frequencies of no less than five (Bakeman & Gottman, 1997). The DAT software was then used

to generate transitional state diagrams (Figure 3) to visually convey the observed probabilities so that similarities and differences in behavioral patterns between groups can be easily identified. In the state diagrams, the thickness of the arrows in the diagrams is in direct proportion to the observed transitional probabilities. The black and gray arrows identify probabilities that are and are not significantly greater than expected. The first and second numerical value displayed in each node identifies the number of times the given action was performed and the number of events that followed the given action. The size of the glow emanating from each node conveys the number of times the given action was performed. Note that the state diagrams in figure 3 and 4 do *not* reveal the transitional probabilities found to be significantly *lower* than expected (showing action sequences that a particular group may have the tendency to avoid). If necessary, these probabilities can be identified with gray arrows drawn with sparsely dotted lines (e.g., - - - >).

What is important to note is that the relative position of each node in a state diagram is identical in all state diagrams for other groups. As a result, the similarities and differences in sequential patterns between the groups are immediately apparent, making the patterns easier to identify and easier to interpret. If the location of any given node were to vary from one state diagram to another, the differences and similarities become very difficult to discern (the same kind of difficulty one experiences when trying to compare the causal maps produced by different groups as illustrated at the top half of figure 2). However, holding the node positions is a fixed location in the state diagrams also presents some disadvantages. The nodes must be positioned in a circular alignment so that all possible transitions between nodes can be clearly displayed in the state diagram. When the number of nodes exceeds six (as in Figure 4), there is insufficient space to clearly display all observed transitions because the arrows cross over and partially obstruct one another. As a result, a state diagram with more than 6 nodes shows only those transitions found to be significantly higher than the expected probability (diagrams with black arrows only with gray arrows omitted).

Results

Mapping processes used by students across all groups. To identify sequential patterns in the map construction process exhibited by students across all three groups, the sequential data from each group was appended into one spreadsheet column and entered into the DAT software to produce the left state diagram in Figure 3. This state diagram revealed six sequential patterns to suggest two general procedures. In one procedure, students moved a node into position, then attached a link to either point to or from the node, changed the link attribute (positive or negative in causal relationship), then inserted a comment into the link to explain the causal relationship, and either revised the existing comment or inserted a new comment to another previously added causal link. In the second procedure, students deleted a link from the causal map (which was performed relatively infrequently), then repositioned a node just prior to inserting a link to point either from or to that node. Once the link was added, they progressed through the same sequence of actions by changing the link attribute, and inserting one or more comments to explain one or more causal relationships. In the future, data will be collected (using the same procedures presented in this case study) to generate a state diagram that identifies the processes used by the instructor and other experts. This can then be compared to the findings revealed in Figure 3 to reveal potential differences (if any exists) and possible deficiencies in student's mapping processes and the processes to scaffold and facilitate in future causal mapping software systems.

Mapping processes associated with map accuracy. The state diagram in Figure 4 revealed some unique patterns between groups (with only 1 of the 10 total patterns observed in all three groups). Despite the absence of an expert process model in this case study, the unique patterns revealed in Figure 4 provide plausible explanations as to why differences in map accuracy were found between the groups. For example, the low performing group, which produced the least accurate map among the three groups, exhibited a four-step linear process of positioning node, inserting causal link, changing link attributes, and inserting comments on link. The medium performing group, which produced the second most accurate map, used more of a stepwise process in which multiple nodes were placed into position first before links were added to the nodes. In contrast, the high performing group, which produced the most accurate map among the three groups, exhibited a three-step linear but iterative process of positioning nodes, adding links, and specifying link attributes (spending no time adding comments into the links to explain the causal relationship). A close review of the video showed in fact that this group work progressively backwards from nodes that had the most direct to the least direct effect on the outcome variable. Although the sample size was not sufficient to make these findings conclusive, the findings serve to illustrate how the techniques and tools described in this study can be used in the future to analyze larger data sets. This is now possible given that the latest version of the jMAP tool automatically logs up to 26 unique and more precise actions that can be performed on a causal map at any time (Jeong, in press).

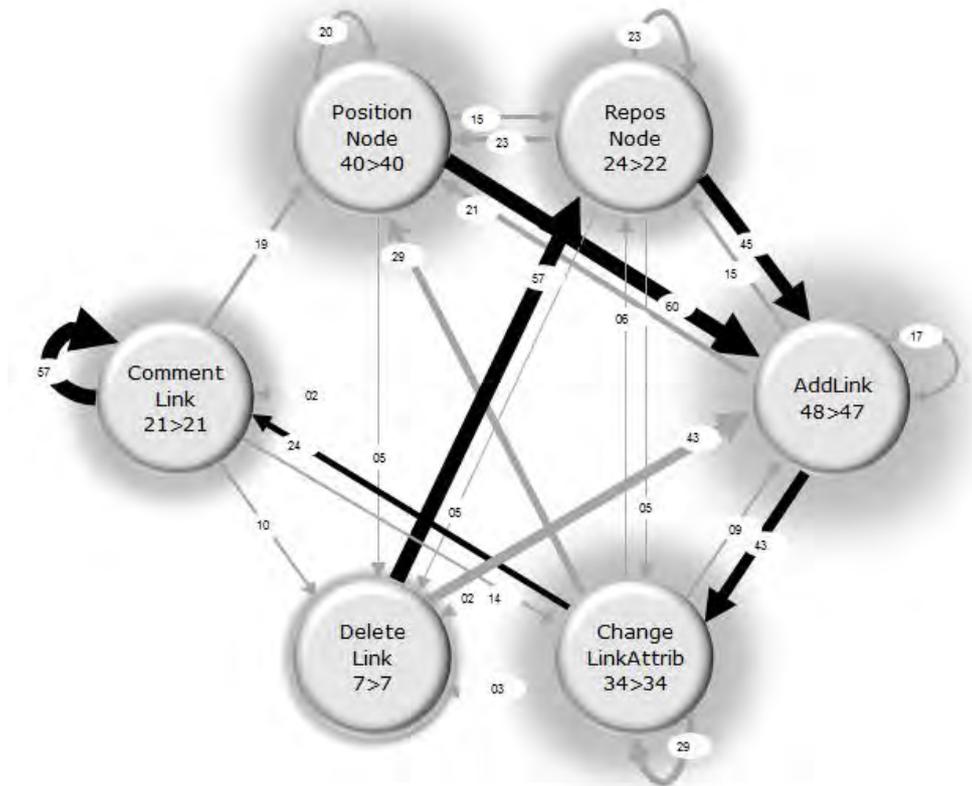


Figure 3. Transitional state diagrams of the map construction process across *all* groups

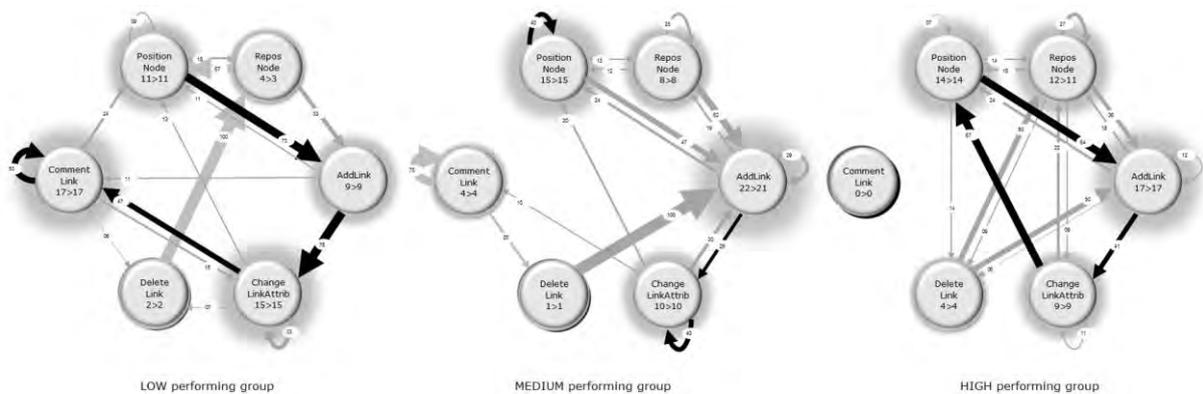


Figure 4. Transitional state diagrams of the map construction processes by group

Map and discourse processed performed by student across all groups. The state diagram in Figure 5 revealed a total of six action sequence patterns between mapping actions (move and/or reposition a node, add and/or delete a node) and forms of discourse. The main patterns or processes revealed in Figure 5 are that: a) the movement of nodes was generally preceded by either confirmative → collaborative or interrogative discourse; and b) the adding and deletion of links were generally preceded by argumentation → agreement → explanation. Once again, future plans are in the works to generate an expert process model to determine to what extent the processes observed in Figure 5 are different to those processes performed by experts (if any exists). Comparing the findings in Figure 5 (or similar findings produced with a larger data set) to the expert model can help to determine how to scaffold the communications between group members to trigger the mapping actions most likely to produce more accurate causal diagrams. The ultimate goal is to help students produce more accurate causal diagrams in order to increase students’ ability to explain the causal mechanisms underlying direct and indirect cause effect relationships between variables in complex systems.

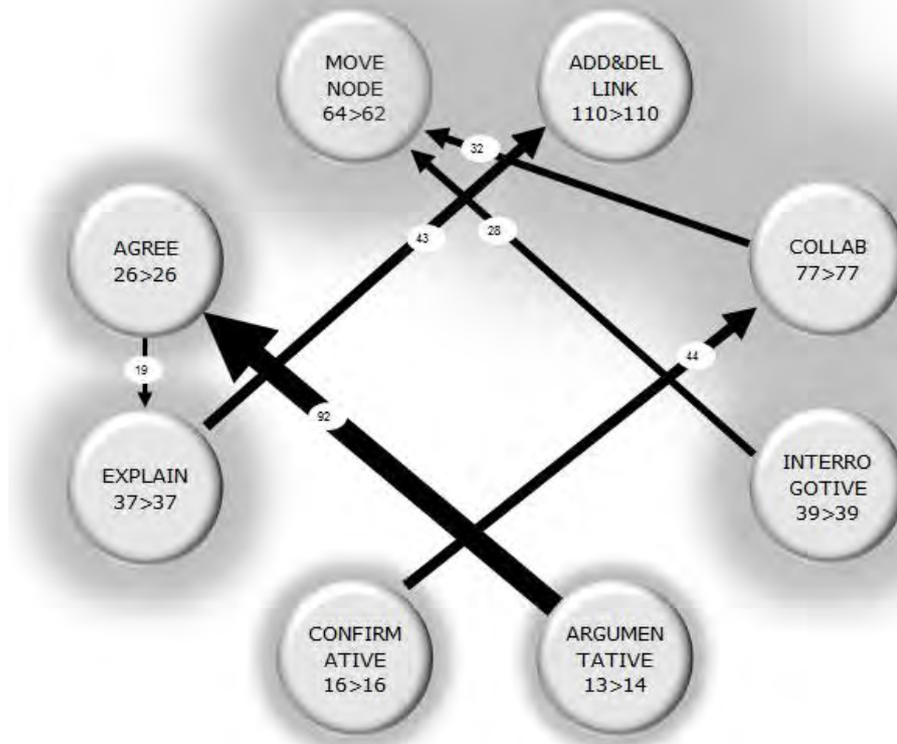


Figure 5. Transitional state diagram of map and discourse processes performed across all groups

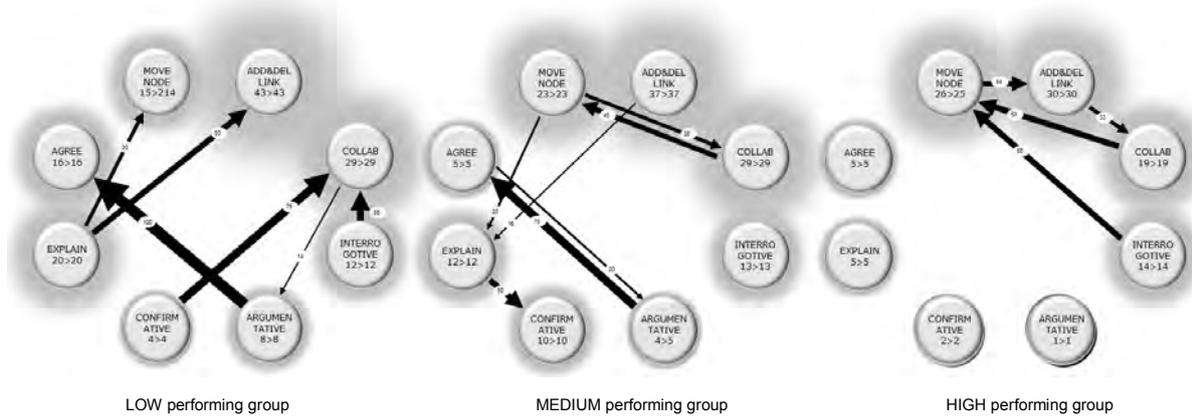


Figure 6. Transitional state diagrams of the map construction *and* discourse processes by group

Mapping and discourse processes associated with map accuracy. The state diagrams in figure 6 reveal a total of 15 unique patterns between mapping actions and discourse. Of these 15 unique patterns, not one of these patterns was observed in all three groups. Among the patterns that were observed in at least two of the groups was the collaborative discourse that took place just before moving a node, and the argumentation that took place just prior to expressing agreement. With respect to map accuracy, here are just some of the findings revealed via a comparative analysis of the three state diagrams in Figure 6. In both the high and medium performing groups (but not in the low performing group), the movement of nodes was preceded by collaborative discourse. These differences suggest that collaborative dialog \rightarrow move node action sequence may have contributed to increases in map accuracy relative to the low performing group. To help explain how the high performing group produced a more accurate causal map than the medium performing group, the state diagrams show that the high performing group moved nodes and linked the nodes just prior to engaging in collaborative discourse. These particular changes to the causal map (moved node with added link) may have been tentative changes used to create a shared visual artifact used to scaffold and facilitate the collaborative discussion and assessment of the proposed actions - similar to the processes described in Cho & Jonassen's (2002) study on the use of argument diagrams for scaffolding group debates. Once the actions were assessed through collaborative discourse, the state diagram shows that the collaborative discourse was then followed by further actions on the placement of the nodes. A close review of the discussion transcripts will be necessary to identify excerpts that illustrate and validate this particular process.

In a similar manner, the process of making tentative changes to the causal map to scaffold further discussion was also observed in the medium performing group, which helps to explain why this group produced a better causal map than the low performing group. The state diagram for the medium group shows that explanative discussion immediately followed the movement of a node and the addition/deletion of a link. The explanative discussion was then immediately followed with confirmative exchanges. In contrast, the low performing group to an extent performed this process in reverse order – explaining (or simply announcing) the act of moving a node or adding/deleting a link, then executing the proposed and/or explained action. A close review of the video revealed that in the low performing group, one of the students controlled and dominated the task, with the dominant student often explaining or simply announcing (as opposed to discussing) her next course of action.

Discussion

Although the sample size in this case study was not sufficient to determine conclusively if the observed patterns are stable or completely unique to the groups across different levels of performance, this study's primary purpose was to illustrate how the described techniques and software tools can be further refined and used in the future to better understand the complex interplay between students' actions and supporting discourse when working collaboratively to construct causal diagrams. While the findings suggest that the processes students perform while constructing causal diagrams can vary widely (which is consistent with Ruiz-Primo & Shavelson's 1997 findings), this study illustrates how a comparative analysis of processes between groups can provide plausible explanations as to why, when, and how some groups produce more accurate maps than others (and/or achieve and exhibit better causal understanding and causal reasoning skills). This approach may be most appropriate when studying complex ill-defined domains where it may be the case that there is no one particular process (or expert model) that produces the best results.

With the ability to automatically log and capture more precise actions performed on a causal diagram within the jMAP software, larger data sets can now be acquired and processed more quickly. But more importantly, the type of data captured in the jMAP logs will enable future studies to examine and model students' causal reasoning processes in far greater detail. For example, the captured log data can now be used to determine when a student is deleting the link $A \rightarrow C$ (when the causal map is currently showing $A \rightarrow C$ and $B \rightarrow$) and inserting a link $A \rightarrow B$ in order to produce the causal chain of events $A \rightarrow B \rightarrow C$ (having realized that B alone is sufficient to cause C and that the effects of A on C is mediated by B). This type of data can help determine to what extent a student is using a backward/deductive vs. forward/inductive approach and depth vs. breadth-first approach (work in progress).

Overall, this paper describes an approach that can be used to conduct further and larger scale studies to identify and validate the key processes that produce higher quality causal diagrams. Once these key processes are identified and tested, the next generation of causal mapping software can be developed to actively monitor, scaffold and standardize the processes (using either a fixed or a dynamic model) to help ensure that the causal maps students produce are accurate measures of their causal understanding rather than a measure of their causal mapping and causal reasoning skills. Such an application can then be used not only as an instructional tool, but also as an alternative and potentially powerful tool for assessing causal understanding in science education on a large-scale basis.

References

- Bakeman, R., & Gottman, J. (1997). *Observing interaction: An introduction to sequential analysis*. Cambridge, UK: Cambridge University Press.
- Bryson, J. M., Ackerman, F., Eden, C., & Finn, C. B. (2004). *Visible thinking: Unlocking causal mapping for practical business results*. Chichester: John Wiley and Sons.
- Cho, K. L., & Jonassen, D. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research & Development, 50*(3), 5-22.
- Clarkson, G. P., & Hodgkinson, G. P. (2005). Introducing Cognizer™: A comprehensive computer package for the elicitation and analysis of cause maps. *Organizational Research Methods, 18*(3), 317-344.
- Decision Systems, Inc. (2003). Root cause analysis using REASON. Last retrieved July 2012 at <http://www.rootcause.com>.
- Jeong, A. (in press). Sequentially analyzing and modeling causal mapping processes that support causal understanding and systems thinking. To appear in Ria Hanewald & Dirk Ifenthaler's (Eds.) *Digital Knowledge Maps in Education: Technology Enhanced Support for Teachers and Learners*, Springer.
- Jeong, A. & Shin, H. (2013). The sequential analysis of causal mapping processes that support causal understanding and systems thinking. Paper to be presented at the 2013 American Educational Research Association conference in San Francisco, CA.
- Jeong, A. & Lee, W.J. (2011). Developing causal understanding with causal maps: The impact of total links, temporal flow, and lateral position of outcome nodes. *Educational Technology Research & Development, 60*(2), 325-340.
- Jeong, A. (2012). *jMAP*. Last retrieved on November, 2012 at <http://myweb.fsu.edu/ajeong/jmap>.
- Jeong, A. (2012b). *Discussion Analysis Tool*. Last retrieved on November, 2012 at <http://myweb.fsu.edu/ajeong/dat>.
- Jonassen, D. H., & Ionas I. G. (2008). Designing effective supports for causal reasoning. *Educational Technology Research and Development, 56*, 287-308.
- McClure, J., & Bell, P. (1990). Effects of an environmental education-related STS approach instruction on cognitive structures of pre-service science teachers. University Park, PA: Pennsylvania State University. (ERIC Document Reproduction Service No. ED 341 582).
- Nesbit, J. C., & Adesope, O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413-448.
- Ruiz-Primo, M., Shavelson, R., & Schulz, S. (1997). On the validity of concept map-base assessment interpretations: An experiment testing the assumption of hierarchical concept maps in science (CSE Technical Report 455). Los Angeles: University of California.
- Scavarda, A.J., Bouzdine-Chameeva, T., Goldstein, S.M., Hays, J.M., & Hill, A. V. (2006). A methodology for constructing collective causal maps. *Decision Sciences Journal, 37*(2), 263-284.

When Instruction Supports Collaboration, but Does not Lead to Learning – the Case of Classroom and Small Group Scripts in the CSCL Classroom

Ingo Kollar, Christof Wecker, Sybille Langer & Frank Fischer, Ludwig-Maximilian University of Munich,
Email: ingo.kollar@psy.lmu.de, christof.wecker@psy.lmu.de, sybille-langer@web.de,
frank.fischer@psy.lmu.de

Abstract: Typically, instructional guidance for CSCL has two aims, namely to help students (a) collaborate on a higher level and (b) become more proficient in the practices that are facilitated during collaboration. This paper presents an empirical study in which high school students' collaborative online search behavior (as a learning process measure) as well as their online search competence (as individual learning outcome) were targeted by small group collaboration scripts and classroom scripts. Both the small group collaboration script and a plenary-plus-group level classroom script yielded positive effects on online search activities during collaboration. When no or just one scaffold was given, correlations between the quality of collaborative online search activities and individual online search competence (as learning outcome) were positive. When both interventions were combined, however, the correlation disappeared, indicating that although this combination helped students act on a higher level during collaboration, it did not contribute to individual learning.

Introduction and aims of the study

Numerous studies have demonstrated that to make CSCL effective, instructional guidance, e.g. through the provision of argument maps (Suthers & Hundhausen, 2003), group awareness tools (Buder & Bodemer, 2008) or collaboration scripts (Kollar, Fischer & Slotta, 2007), is necessary. The aim of such guidance typically is two-fold: first, it is meant to help learners collaborate on a higher level (e.g., by increasing the level of knowledge building activities; see Schellens, De Wever, van Keer & Valcke, 2007) than without guidance. Second, guidance also aims at helping students learn more, i.e. reach higher individual knowledge and skill levels measured after collaboration (see, for example, Weinberger, Ertl, Fischer & Mandl, 2005).

Sometimes, multiple sources of guidance are combined to reach these objectives. For example, Kollar et al. (2012) combined small group collaboration scripts with heuristic worked examples in a CSCL environment designed to acquire mathematical argumentation competence. Likewise, Kopp and Mandl (2011) used both a small group collaboration script and a content scheme to provide guidance in a case-based CSCL environment. Providing learners with several sources of instructional guidance typically has the aim to produce synergistic scaffolding effects (Tabak, 2004), i.e. that the potentially positive effects of the used scaffolds do not simply add up, but rather interact positively with each other, with the result that the effects of each scaffold are amplified by the simultaneous provision of the other. However, as research shows, combining scaffolds in a way that synergistic scaffolding occurs is difficult. For instance, Kollar, Wecker, Langer and Fischer (2011) showed that combining a small-group collaboration script guiding dyadic online search processes and a plenary-plus-group level classroom script (i.e., an instructional intervention that alternated the modeling of online search activities as plenary activities and dyadic online search phases as group level activities in the classroom) did not yield a synergistic scaffolding effect on the acquisition of online search competence. Although both interventions were effective when the other one was not provided, their combination did not affect the effectiveness of the small group collaboration script and even slightly reduced the effectiveness of the plenary-plus-group level classroom script. However, in the Kollar et al. (2011) study, no process-based explanation for this result was provided. Therefore, this paper aims at providing an analysis of collaborative online search activities that occurred in the Kollar et al. (2011) study and investigates how these collaborative online search activities relate to online search competence students displayed in a subsequent individual posttest.

Fostering online search competence as a grand challenge for education

To participate in societal debates about science-related issues (e.g., whether nuclear power plants should be shut down), members of the information society need to have well-developed strategies to find and use relevant and credible information. Much of this information is available on the Internet. Since anyone can publish on the Internet, however, credibility, lopsidedness, timeliness and relevance of information are critical issues. Thus, supporting learners in their development of online search competence is a grand challenge for education.

Gerjets, Kammerer and Werner (2011) proposed a five-step model of successful online search: (1) Users face an information need and define a search goal. (2) They select a search engine, choose search terms and send their query off. (3) They scan the resulting search results page and evaluate it based on a set of criteria

such as relevance or credibility. (4) Once they have selected a website for closer inspection, users scan it and extract the required information (again, based on quality criteria such as relevance and credibility). Finally, (5) users need to compare and integrate the information they found on the selected website(s) into a coherent solution for the information problem. As a wealth of empirical research has shown, this ideal online search strategy can however hardly be observed in students across different age groups and educational contexts. For example, Tomaiuolo and Packer (1996) demonstrated that many university students have problems employing appropriate search terms for solving simple retrieval tasks such as “locate the full text of the Magna Charta”. Likewise, Brand-Gruwel, Wopereis and Walraven (2009) showed that students often use inadequate criteria to assess the quality of websites, such as the language in which they are written or text length. Thus, there clearly is a need to design instructional interventions that help learners gain online search competence.

Scripting as a way to foster online search competence

A promising way to foster online search competence is having students collaborate (Lazonder, 2005) during their online search and by structuring their collaboration through scripting (e.g., Fischer, Kollar, Mandl & Haake, 2007). Scripts provide learners with direct guidance on how to structure their collaboration by assigning activities and roles to different learners within a social learning setting. One way to differentiate different types of scripts is to distinguish between classroom scripts and small group collaboration scripts (see Kollar et al., 2011). *Classroom scripts* provide coarse-grained activity structures that distribute learning activities over the social levels of the classroom (see Dillenbourg & Jermann, 2007). For example, group level classroom scripts would have all learning activities within a classroom carried out solely by small groups, while a plenary-plus-group-level classroom script might alternate between modeling (as a plenary activity) and dyadic learning activities (as group level activities). Of course, further classroom scripts are conceivable. *Small group collaboration scripts*, in turn, provide more fine-grained guidance with respect to the specific activities that are to be shown within small group collaboration. For example, a small group collaboration script may have one learner of a dyad suggest which link to click on a search results page, while the other learner is prompted to critically reflect upon his/her learning partner’s choice based on credibility considerations. As the study by Kollar et al. (2011) demonstrated, both classroom scripts and small group collaboration scripts can be designed in a way that online search competence (as an individual learning outcome) can effectively be facilitated. Further research has produced a wealth of evidence for the potentials of classroom scripts (e.g., Dillenbourg & Hong, 2008; Hmelo-Silver, 2004; Kolodner, 2007) and small group collaboration scripts (e.g., Kollar, Fischer & Slotta, 2007; Rummel & Spada, 2005; Schellens et al., 2007; Tsovaltzi et al., 2010; Wecker & Fischer, 2011) also for the acquisition of skills and competences beyond the online search field.

The present study

As already mentioned, this study provides an in-depth analysis of collaborative online search activities from the study by Kollar et al. (2011). For the purposes of that study, a 4.5 weeks curriculum unit for 9th grade biology classrooms was designed. Over the course of the unit, students had repeated opportunities to use the Internet to develop a well-warranted position on whether Genetic Engineering should be allowed or not. In a 2x2 factorial design, we systematically varied two independent factors: “type of classroom script” (group level classroom script vs. plenary-plus-group-level classroom script) and “small group collaboration script” (with vs. without). In Kollar et al. (2011), the main dependent variable was the students’ level of online search competence after completion of the curriculum unit (prior online search competence was controlled for). The results demonstrated that both the small group collaboration script and the plenary-plus-group level classroom script had a positive effect, as long as students were only provided with one of the two. Concerning the combination of the small-group collaboration script and the plenary-plus-group level classroom script, we expected to find a synergistic scaffolding effect, i.e. that receiving modeling of good online search would especially pay off if the subsequent dyadic online search would be structured by appropriate prompting. However, no synergistic scaffolding effect (Tabak, 2004) was found. As these results have already been published, they are not further reiterated in the analyses of this paper. Instead, this paper tries to answer two research questions aiming at a better understanding of the results with respect to the acquisition of online search competence as reported in Kollar et al. (2011):

1. What are the effects of providing learners with a small group collaboration script (vs. unscripted small group collaboration), a plenary-plus-group level classroom script (vs. a group-level classroom script), as well as their interaction on the quality of online search activities exhibited during collaboration?
2. How does the quality of the collaborative online search activities relate to the online search competence individual students demonstrate in a subsequent posttest?

Based on the learning outcome analyses reported in Kollar et al. (2011), we expected an analogous result pattern for the quality of the collaborative online search activities that were shown during collaboration. More specifically, we expected positive effects of both the small group collaboration script and the plenary-plus-group level classroom script, as long as they were provided individually, on the quality of the exhibited

collaborative online search activities. In the combined condition, we expected a significantly lower quality of collaborative online search activities when compared to the condition “plenary-plus-group level classroom script/without small group collaboration script”, but a comparable level to students from the condition “group level classroom script/with small group collaboration script”. Learners who neither received the small group collaboration script nor the plenary-plus-group level classroom script were expected to show the lowest levels of collaborative online search activities. With respect to the relation between the quality of collaborative online search activities in the process and individual online search competence displayed in the individual posttests, we expected significant and positive correlations in all conditions, i.e. the higher the quality of the collaborative online search activities in the dyads, the higher the individuals’ online search competence after collaboration.

Method

Participants and design

Overall, 174 students from eight classes of four high schools from Southern Germany participated in the study. However, only for 151 9th graders data from online search phases as well as the individual posttest were available. Therefore, only these 151 students were included in this study. As described, we established a quasi-experimental 2x2-factorial pre-post test design with the independent factors “type of classroom script” (group level classroom script vs. plenary-plus-group-level classroom script) and “small group collaboration script” (with vs. without). Eight classes were randomly assigned to the four experimental conditions, i.e. each condition was implemented in two classes (see table 1).

Table 1: Design of the empirical study.

		Small-group collaboration script	
		Without	With
Type of classroom script	Group-level	N = 36 students (2 classes)	N = 48 students (2 classes)
	Plenary-plus-group level	N = 22 students (2 classes)	N = 45 students (2 classes)

Instructional setting and independent variables

The experimental design was integrated in a curriculum unit during which students received the task to use the Internet to develop a well-warranted position on the question whether Genetic Engineering should be allowed or not. For this purpose, each single student was equipped with a laptop computer on which a LAN connection was established to allow for Internet browsing. Since the experiment took place in the regular Biology lessons of the participating classes and followed the regular timetable of each class, only two lessons per week in each class took place. Overall, the curriculum unit spanned seven lessons. One additional lesson right before the start of the intervention was used for the administration of pretests, and one lesson right after the end of the intervention was used for posttests. During the seven learning sessions, after a general introduction by the teacher to the topic and to successful online search behavior, three content-specific learning cycles were created. Cycle 1 dealt with economic issues, cycle 2 with ecological issues, and cycle 3 with health-related issues of Genetic Engineering. Each of these cycles consisted of three steps. In step 1, students had the opportunity to browse an online environment that held relevant Biological content knowledge on Genetics and Genetic Engineering. The online environment was created in WISE (Slotta & Linn, 2000), and its content design was based on regular 9th grade Biology textbooks. In step 2, the actual online search phase took place, during which the experimental variation was implemented (see below). During this step, learning mainly took place in dyads. In step 3, classes in all conditions engaged in a plenary discussion that invited the students to exchange and critically discuss the arguments and pieces of evidence they had found or developed during their Internet search.

Independent variables

The two independent factors “type of classroom script” and “small-group collaboration script” were systematically varied during step 2 in each of the three learning cycles, i.e. in the phases in which students were supposed to search the Internet for arguments and evidence that would seem helpful to develop a position in the Genetic Engineering debate. In all four experimental conditions, the Internet browsers of two learning partners each were connected to each other, i.e. whenever one learning partner went to a new website, his/her partner’s browser would go there too. This was realized by a browser plug-in called S-COL (Wecker et al., 2010).

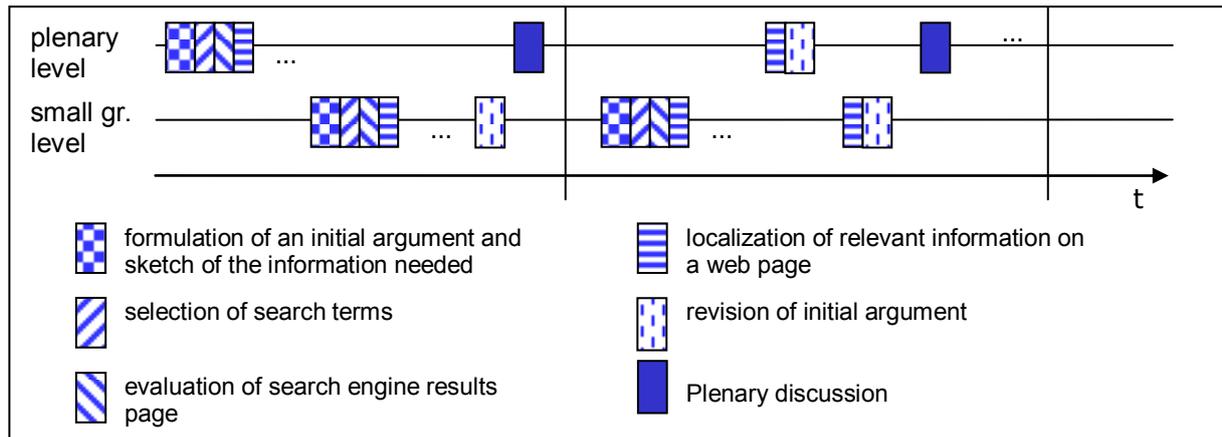


Figure 1. Graphical representation (derived from Dillenbourg & Jermann, 2007) of the plenary-plus-group level classroom script (lines represent the two social levels “plenary” and “small group”, boxes represent the different online search activities exhibited on the respective social levels.

In the *plenary-plus-group level classroom script* condition (see figure 1), dyadic online search during step 2 of each cycle was interrupted from time to time by modeling phase (as plenary activities), during which the teacher and a student or two students demonstrated what a successful online search looks like, before student dyads went on with their own online search. The design of the plenary-plus-group level classroom script as well as the content of the modeling phases was based on an adaptation of the five-step online search competence model by Gerjets et al. (2011). Thus, students were instructed to follow a five-step online search strategy: (1) *formulating an initial argument and a sketch of the information needed*, (2) *selecting search terms*, (3) *evaluating the search engine results page (SERP)*, (4) *localizing relevant information on a web page*, and (5) *formulating the final elaborated argument*. For all steps, certain quality criteria were also provided in the modeling phases. For example, during the evaluation of the hit list, students were instructed to discuss the credibility and relevance of the single links. In contrast, in the *group level classroom script* condition, all search activities in step 2 of each content-specific learning cycle were to be conducted in dyads, i.e. dyadic search activities were not interrupted by modeling phases.

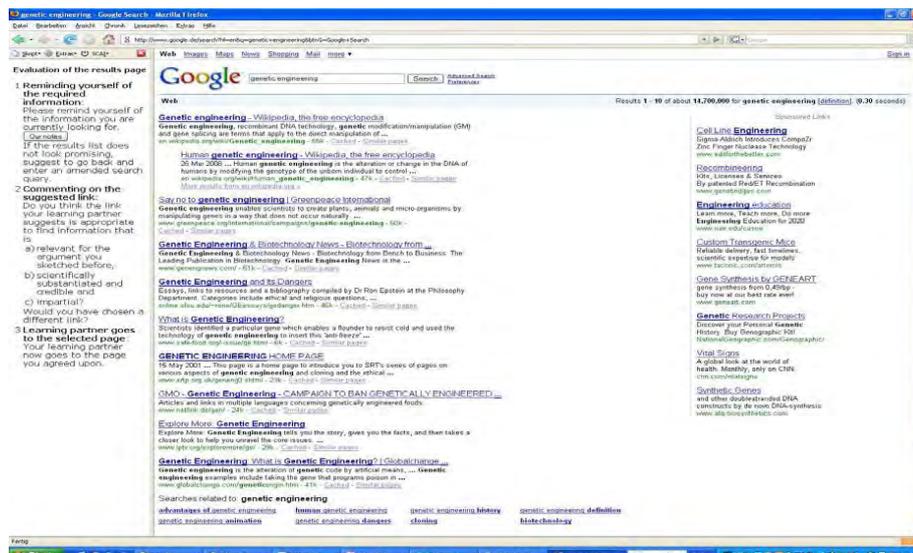


Figure 2. Screenshot of the small group collaboration script (view of the screen of one of the two learners during the step “scan search engine results page”; the right side of the screen displays a regular Google results list; the left side of the screen provides prompts related to this online search step for one of the two learners).

Just like the plenary-plus-group level classroom script, the *small group collaboration script* (see figure 2) was designed on the basis of the adapted five-step online search competence model by Gerjets et al. (2011). For each of the five online search steps, the two learners of each dyad received complementary prompts that told them what to do. Typically, learner A was supposed to suggest what to do next (e.g., suggest a link to click on while the browser displayed a hit list), and learner B was supposed to critically reflect upon the suggestions of

learner A (e.g., by a prompt “Do you think the link your learning partner suggests is appropriate to find information that is (a) relevant for the argument you sketched before, (b) scientifically substantiated and credible and (c) impartial? Would you have chosen a different link?”). With each new search, these two roles were switched among the learning partners. When the small group collaboration script was combined with the plenary-plus-group level classroom script, modelling was alternated with structured dyadic search activities based on the prompts specified in the small group collaboration script. When the small group collaboration script was combined with the group-level classroom script, all search activities were to be conducted in dyads, and all dyads received the prompts that made up the small group collaboration script; however, dyads in this condition did not receive modelling of successful online search. In the condition without small group collaboration script, no prompts were displayed during dyadic online search; however, the Internet browsers of the two learning partners were connected in the way that was described earlier.

Dependent variables

The *quality of the collaborative online search activities* the dyads exhibited during treatment was analyzed based on screen-and-audio-recordings. The first ten minutes of each dyad in each search phase were video-coded based on a coding scheme that was designed on the adapted competence model of successful online search by Gerjets et al. (2011). E.g., we coded whether the students currently performed activities belonging to the step of sketching an argument, or whether they performed activities belonging to the step of evaluating a search engine results page, as well as what quality criteria they applied. The unit of analysis for this analysis were segments of ten seconds, and the predominant activity in these 10 seconds was coded. The mutually exclusive codes were: (1) *formulation of an initial argument and a sketch of the information needed*, (2) *selection of search terms*, (3) *evaluation of the hit list*, (4) *localization of relevant information* on a web page, (5) *formulation of the final elaborated argument*, and (6) *other*. Because all time samples were drawn from the beginning of the online search phases, activities belonging to early steps of the underlying online search model by Gerjets et al. (2011) were more appropriate than later steps. Therefore, a composite indicator of the quality of collaborative search activities was computed that reflects the appropriateness of the activities belonging to each of the five steps of the strategy during the first ten minutes of each search phase. This composite indicator was calculated as the sum of the proportions of time spent on the activities belonging to the more appropriate first two steps “*formulation of an initial argument and sketch of the information needed*” (with fourfold weight) and “*selection of search terms*” (with double weight), minus the sum of the proportions of time spent on the activities belonging to the less appropriate final three steps “*evaluation of the hit list*” (with unit weight), “*localization of relevant information*” (with double weight), and “*written formulation of the final elaborated argument*” (with threefold weight). This indicator was calculated separately for the learners and their learning partners, resulting in two composite indicator variables for strategy performance. A value of higher than 0 would mean that a person displayed more reasonable (step 1 and 2) than unreasonable (steps 3, 4 and 5) activities during the first ten minutes of each search phase, while a value of lower than 0 would mean that s/he displayed more unreasonable (steps 3, 4 and 5) than reasonable (steps 1 and 2) activities. To determine objectivity, a subsample of 11% of the data from this and a further study (Wecker, Kollar & Fischer, 2011) that used the same task and coding scheme (but different experimental variations) was coded by three independent raters, with ICCs for each of the five online search steps of about .90 (*formulation of an initial argument and sketch of the information needed*: ICC = .96, *selection of search terms*: ICC = .90, *evaluation of the hit list*: ICC = .95, *localization of relevant information*: ICC = .97, *written formulation of the final elaborated argument*: ICC = .88, composite indicator of *strategy performance*: ICC = .97).

Online search competence as an individual learning outcome was measured in an individual posttest (see also Kollar et al., 2011) that asked students to describe in as much detail as possible how they would use the Internet to arrive at a reasoned position in a science-related debate different from Genetic Engineering (whether nuclear power plants should be shut down or not). For the pretest, an analogous test on a different science topic was used. Since the results analyzing the effects of the two treatments and their combination on this measure have already been published elsewhere, interested readers are referred to the corresponding paper (Kollar et al., 2011). To understand the analyses of this paper, the result pattern on the acquisition of online search competence however needs to be kept in mind: As reported above, Kollar et al. (2011) found that although both interventions were effective when the other one was not provided, their combination did not yield synergistic scaffolding effects, i.e. it did not affect the effectiveness of the small group collaboration script and even slightly reduced the effectiveness of the plenary-plus-group level classroom script.

Statistical analyses

To determine the effects of the two independent variables on the quality of the collaborative online search activities, an ANCOVA with “type of classroom script” and “small group collaboration script” as fixed factor, classes as further fixed factor nested within the experimental conditions (to account for the hierarchical data structure), the composite indicator for quality of collaborative online search activities as the dependent variable

and prior online search competence as a covariate was conducted. To answer the question on the relation between learning activities and outcomes, bivariate correlations were computed between “quality of collaborative online search activities” and “online search competence (outcome)”. For all analyses, the significance level was set to 5%.

Results

With respect to RQ 1 on the effects of the two treatments and their different combinations on the quality of online search activities during collaboration, the descriptive data (see table 2) showed that learners who received the small group collaboration script together with the plenary-plus-group level classroom script exhibited the highest quality levels. The lowest levels were observed when learners did not receive the small group script and followed the group level classroom script. Students from the other two conditions (“with small group collaboration script and group level classroom script” and “without small group collaboration script and plenary-plus-group level classroom script”) were in between and reached comparable levels.

Table 2: Means and standard deviations of the (composite indicator of) quality of collaborative online search activities for the four experimental conditions.

	Without small group collaboration script				With small group collaboration script			
	Group level classroom script		Plenary plus group level classroom script		Group level classroom script		Plenary plus group level classroom script	
	M	SD	M	SD	M	SD	M	SD
Quality of online search activities	-0.68	0.33	-0.49	0.28	-0.33	0.47	-0.01	0.50

An ANCOVA with “type of classroom script” and “small group collaboration script” as fixed factors, classes as a further fixed factor nested within the experimental conditions (to account both for the hierarchical data structure and to control for possible teacher effects), the composite indicator for quality of collaborative online search activities as the dependent variable and prior online search competence as a covariate revealed a significant main effect for the small group collaboration script, $F(1; 142) = 29.23, p < .01$, partial $\eta^2 = .17$, with students who had learned with the small group collaboration script outperforming students who did not receive a small group collaboration script. Also, we found a significant main effect for type of classroom script, $F(1; 142) = 14.72, p < .001$, partial $\eta^2 = .09$, favoring the conditions with plenary-plus-group level classroom script over learners who had learned with the group level classroom script. There was no significant interaction effect, $F(1; 142) = 0.02, p = .88$, partial $\eta^2 < .01$.

With respect to RQ 2 concerning the relation between the quality of collaborative online search activities and the levels of online search competence students displayed in the individual posttest, bivariate correlations were calculated. When all four conditions were taken together, this correlation turned out to be insignificant ($r = .05, p = .24$). However, we also calculated separate correlation analyses for the combined condition (small group collaboration script and plenary-plus-group level classroom script) and the remaining three conditions. This revealed an interesting pattern: When both scaffolds were combined, there was a slightly significant negative correlation between the quality of the search strategy during the process and online search competence that students exhibited in the posttest ($r = -.27; p = .08$; two-tailed), while for the other three experimental conditions taken together, the correlation was positive and approached statistical significance ($r = .18; p = .08$; two-tailed).

Discussion

This paper provides an in-depth process analysis of a study presented in Kollar et al. (2011) which showed that both a small group collaboration script and a plenary-plus-group level classroom script that were employed in a curriculum unit on Genetic Engineering in 9th grade Biology classrooms were effective ways of fostering online search competence. Once these two scaffolds were combined, however, neither an addition of effects nor a positive interaction (meaning that both treatments would have amplified each other’s effects) appeared. To the contrary, adding the small group collaboration script while a plenary-plus-group level classroom script was implemented tended to reduce the effectiveness of the latter, while the effectiveness of the small group collaboration script compared to unsupported collaboration was not affected by the type of classroom script that was employed. Thus, with respect to the acquisition of individual online search competence, combining a small group collaboration script and a plenary-plus-group level classroom script may have produced what Dillenbourg (2002) termed “overscripting” (for the exact analyses that underlie these results, please see Kollar et al., 2011).

The results of the process analyses presented in the current paper only partially reflect these results, since with respect to the quality of the collaborative online search activities, the two treatments did not work best in isolation, but instead when they were combined, as was indicated by the additive main effects that were found with respect to research question 1. On the one hand, given the incongruence with respect to the results of the learning outcome analyses reported in Kollar et al. (2011), this result is surprising. On the other hand, theoretically it may have been expected, since both the small group collaboration script and the plenary-plus-group level classroom script were specifically designed with the aim to support students during their collaborative online search activities. That the two scripts were effective with respect to this aim, can be regarded as support for previous research that has demonstrated (a) the effectiveness of small group collaboration scripts to facilitate collaborative learning activities (e.g., Schellens et al., 2007) and (b) the effectiveness of instructional classroom approaches that implement specific distributions of learning activities over the different social levels of the classroom, such as Problem-based Learning (Hmelo-Silver, 2004), Learning-by-Design (Kolodner, 2007) or Reciprocal Teaching (Palincsar & Brown, 1984), even though it has to be noted that few empirical studies exist that systematically compared the effects of different distributions of activities over social levels on individual learning outcomes. However, even though we found that the combination of learning with a small group collaboration script and a plenary-plus-group level classroom script was most successful, we still did not find a true synergistic scaffolding effect (Tabak, 2004). Such an effect would mean that the two scaffolds would mutually amplify their effects (i.e. cause a positive interaction effect). It is possible that true synergistic scaffolding was not produced because both scaffolds were redundant, since their design followed the same theoretical assumptions concerning good online search strategies. It might be that once students have seen the teacher or other students model high level online search strategies, they may not have seen the necessity of paying attention to the small group collaboration script prompts anymore. If this is true, this may indeed be labelled an over-scripting effect (Dillenbourg, 2002). Further research is necessary to test the validity of this interpretation.

Still, that the additive effects pattern we found with respect to the quality of the collaborative online search activities did not appear on the learning outcome level (using students' performance in the online search competence posttest) requires further consideration, and the correlational analyses we ran with respect to research question 2 are helpful in this respect. As these analyses revealed, when the three conditions that either employed the small group collaboration script or the plenary-plus-group level classroom script alone, or none of the two, were taken together, there was a weak, but (marginally) significant positive correlation between the quality of collaborative search activities and online search competence measured in the posttest. However, in the condition that combined the small group collaboration script and the plenary-plus-group level classroom script, a marginally significant negative correlation of moderate size was observed. In other words, although the combination of these two scaffolds helped students act on a higher level during collaboration, it did not help them to actually acquire the competence; it even tended to hinder their competence acquisition. A tentative explanation for this result could be that students may exhibit a high degree of collaborative online search strategy use without actually internalizing the strategy due to over-reliance on the rich scaffolds with the combination of the two scaffolds. Another interpretation might be that the combination of the two scaffolds has led to an advanced automatization of the online search strategy that was proposed in the two scripts, which may have made it hard for students to have the strategy later available declaratively (since the online search competence test had students describe – and not perform – an ideal online search). Future research is necessary to test these assumptions.

In summary, our results imply that the combination of small group collaboration scripts and plenary-plus-group level classroom scripts seems helpful to help dyads perform higher-level search activities. If the goal is to produce positive effects on individual online search competence, yet, the two scaffolds should rather be given without presenting the other as well.

References

- Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of Information Problem Solving while using Internet. *Computers & Education*, *53*, 1207-1217
- Buder, J. & Bodemer, D. (2008). Supporting controversial CSCL discussions with augmented group awareness tools. *International Journal of Computer-Supported Collaborative Learning*, *3*, 123-139.
- Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL* (pp. 61-91). Heerlen: Open Universiteit Nederland.
- Dillenbourg, P., & Hong, F. (2008). The mechanics of CSCL macro-scripts. *International Journal of Computer-Supported Collaborative Learning*, *3*, 5-23.
- Dillenbourg, P. & Jermann, P. (2007). Designing integrative scripts. In F. Fischer, I. Kollar, H. Mandl & J. Haake (Eds.), *Scripting Computer-Supported Collaborative Learning: Cognitive, computational, and educational perspectives* (pp. 259-288). New York: Springer.

- Fischer, F., Kollar, I., Mandl, H. & Haake, J. (Eds.) (2007). *Scripting Computer-Supported Collaborative Learning: Cognitive, computational, and educational perspectives*. New York: Springer.
- Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction, 21*, 220-231.
- Gillies, R. M. (2004). The effects of cooperative learning on junior high school students during small group learning. *Learning and Instruction, 14*(2), 197-213.
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*, 235-266.
- Kollar, I., Fischer, F., & Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learning and Instruction, 17*(6), 708-721.
- Kollar, I., Ufer, S., Lorenz, E., Vogel, F., Reiss, K. & Fischer, F. (2012). Using heuristic worked examples and collaboration scripts to help learners acquire mathematical argumentation skills. In J. van Aalst, K. Thompson, M. J. Jacobson & P. Reimann (Eds.), *The future of learning – ICLS 2012 Conference Proceedings (Volume 1 – full papers)* (pp. 331-338). International Society of the Learning Sciences.
- Kollar, I., Wecker, C., Langer, S., & Fischer, F. (2011). Orchestrating Web-Based Collaborative Inquiry Learning with Small Group and Classroom Scripts. In H. Spada, G. Stahl, N. Miyake & N. Law (Eds.), *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Proceedings, Vol. 1 - Long Papers* (pp. 422-429). International Society of the Learning Sciences.
- Kolodner, J. L. (2007). The roles of scripts in promoting collaborative discourse in learning by design. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting Computer-Supported Collaborative Learning—Cognitive, Computational, and Educational Perspectives* (pp. 237-262). New York: Springer.
- Kopp, B., & Mandl, H. (2011). Fostering argument justification using collaboration scripts and content schemes. *Learning and Instruction, 21* (5), 636–649.
- Lazonder, A.W. (2005). Do two heads search better than one? Effects of student collaboration on Web search behavior and search outcomes. *British Journal of Educational Technology, 36*, 465-475.
- Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences, 14*(2), 201-241.
- Schellens, T., Van Keer, H., De Wever, B., & Valcke, M. (2007). Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups? *International Journal of Computer-Supported Collaborative Learning, 2*(2-3), 225-246.
- Slotta, J. D. & Linn, M. C. (2000). How do students make sense of Internet resources in the science classroom? In M. J. Jacobson & R. Kozma (Eds.), *Learning the sciences of the 21st century*. Mahwah, NJ: LEA.
- Suthers, D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences, 12*(2), 183-218.
- Tabak, I. (2004). Synergy: A complement to emerging patterns of distributed scaffolding. *The Journal of the Learning Sciences, 13*(3), 305-335.
- Tomaiuolo, N. G. & Packer, J. G. (1996). Web search engines: Key to locating information for all users or only the cognoscenti. In Online Information 96 (Eds.), *Proceedings of the International Online Information Meeting* (20th, Olympia 2, London, England, United Kingdom, Dec. 3-5, 1996), Retrieved 8/11/2008, http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/f9/a2.pdf.
- Tsovaltzi, D., Rummel, N., McLaren, B. M., Pinkwart, N., Scheuer, O., Harrer, A. & Braun, I. (2010). Extending a virtual chemistry laboratory with a collaboration script to promote conceptual learning. *International Journal of Technology Enhanced Learning, 2*(1/2), 91-110.
- Wecker, C. & Fischer, F. (2011). From guided to self-regulated performance of domain-general skills: The role of peer monitoring during the fading of instructional scripts. *Learning and Instruction, 21* (6), 746-756.
- Wecker, C., Kollar, I., & Fischer, F. (2011). Explaining the Effects of Continuous and Faded Scripts on Online Search Skills: The Role of Collaborative Strategy Practice. In H. Spada, G. Stahl, N. Miyake & N. Law (Eds.), *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Proceedings, Volume 1 - Long Papers* (pp. 390-397). International Society of the Learning Sciences.
- Wecker, C., Stegmann, K., Bernstein, F., Huber, M. J., Kalus, G., Kollar, I., Rathmeyer, S. & Fischer, F. (2010). S-COL: A Copernican turn for the development of flexibly reusable collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 5*(3), 321-343.
- Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science, 33*(1), 1-30.

Acknowledgments

The research reported here was funded by the Deutsche Forschungsgemeinschaft (DFG). The authors would like to thank the students and teachers who have participated in the study.

Interface Tangibility and Gesture in Mediating Individual Agency Within Group Spatial Problem Solving With an Ecosystem Simulation

Helen Kwah, New York University, 82 Washington Square East, 6th Floor, New York, NY 10003

Leilah Lyons, University of Illinois at Chicago, 851 S. Morgan (M/C 152), Chicago, IL 60607
Dixie Ching, New York University, 82 Washington Square East, 6th Floor, New York, NY 10003

Email: helen.k@nyu.edu, llyons@uic.edu, dixie@nyu.edu

Abstract: This paper examines how a tangible interface facilitates gesture-mediated spatial reasoning during collaborative problem solving, as evidenced by sensitivity to emergent spatial patterns within a complex system simulation of a watershed. We tested the interface against two non-tangible input comparison conditions (single- and multi-mouse) to control for access differences. To determine if groups' solutions displayed sensitivity to emergent spatial patterns, we constructed a quantitative "dynamism" measure, and found that solutions produced in the tangible condition were significantly more spatially sensitive. To better understand why, we conducted a case study by selecting two representative participant groups and performing qualitative multimodal analyses of participant speech and gesture. Our findings indicate that the tangible interface's greater affordance of gesture allowed participants to express concepts containing both spatial and temporal properties, with the added benefit of increasing the agency of less-verbally participatory group members to explore and contribute their ideas more equitably.

Introduction

When a complex system simulation is spatial, meaning that the location of simulation elements has an impact on the emergent outcomes of the simulation, special supports may be needed to assist learners as they come to reason spatially about the represented domain. By "reasoning spatially," we do not refer to classic visuospatial cognition literature (e.g., wayfinding or mental rotation tasks), but rather to the class of problems identified by the NRC Committee on Geography's (2006) "Learning to Think Spatially" report, wherein reasoning spatially entails the ability to "perceive, remember, and analyze the static and, via transformations, the dynamic properties of objects and the relationships between objects." In our problem space (the integration of green infrastructure into urban landscapes) spatiality is highly important: a given green infrastructure element may have a large impact or none at all depending on its placement. Because traditional desktop simulations ask the user to transmute spatial manipulations through (single-user) input devices like mice, we designed a tangible simulation interface to support more direct spatial manipulations by multiple users and conjectured that the tangibility would improve users' abilities to construct solutions sensitive to the simulation's emergent spatial phenomena. We tested the interface against two non-tangible mouse input conditions, and gauged sensitivity of solutions with a "dynamism measure"- finding that solutions produced in the tangible condition were significantly more spatially sensitive. A case study was then conducted to examine why, and our findings, which are presented in this paper, indicated that the tangible interface afforded gesture-mediated spatial reasoning, which enabled less verbally participatory group members to better consider spatial and temporal information and more equitably express their reasoning, resulting in improved group problem solving outcomes.

Prior Work

In human-computer interaction research, tangible user interfaces (TUIs) in the form of multi-touch tabletop displays have been shown to result in more equitable participation by group members working on tasks with spatial or other physical constraints, like arranging office seating (Marshall et al., 2008) or planning itineraries (Rogers & Lindley, 2004). In a study that compared a multi-touch tabletop to input using TUIs, the use of tangibles encouraged greater participation from people who normally found it difficult to contribute verbally in group settings (Rogers et al., 2009). Other research has found that the use of TUIs facilitated individual spatial problem solving as users leveraged the physical affordances of tangible objects (Antle et al., 2009). This research adds to this prior literature by examining the affordance of TUIs for gesture-mediated spatial reasoning in the context of group spatial problem solving with a dynamic, computer-based simulation.

There is a growing body of research on the role gesture plays in facilitating spatial cognition (Alibali, 2005; Goldin-Meadow, 2003). There are various dimensions or types of gesture that have been identified, including deictic (pointing) and iconic gestures (McNeill, 1992). Iconic gestures are considered 'representational' as they present kinesthetic images of objects or actions; for example, waving a cupped hand

up and down while talking about playing basketball. Iconic gestures can thus facilitate spatial cognition by helping the gesturer to focus upon and represent dynamic, temporal and spatial information—information that would be difficult to convey in speech but that is necessary for understanding a topic (Poizzer-Ardenghi & Roth, 2007). In fact, gesture appears to provide a visuospatial modality for learners to explore ideas that cannot yet be articulated in speech because the domain is new or the conceptual vocabulary has not been acquired (Alibali & Goldin-Meadow, 1993; Roth & Welzel, 2001). In research on computer-supported collaborative learning, gestures have been acknowledged as an important resource for knowledge building although the focus has primarily been on deictic gestures for establishing joint attention and shared reference (Stahl, 2003). Studies of small-group science inquiry have examined both deictic and iconic gestures, and found that iconic gestures in particular can provide support for spatial reasoning and representation of information that could be shared and negotiated by the group (e.g., Radinsky et al., 2008; Roth & Lawless, 2002). The studies by Radinsky, Goldman, and Singer (2008; Singer et al., 2008) also examined tracing gestures, although traces were grouped with deictic gestures, thus highlighting their indexical function. Tracing gestures can have both deictic and iconic elements (Goodwin, 2003), and in this study we use the term "iconic-tracing" to highlight the dual function of such gestures to publically index an object visualized in the environment, and to represent the visuospatial or dynamic characteristics of the object.

Study Design and Methods

29 triads (mostly undergrads and high school students) were given the task of creating optimal rainwater infiltration solutions by placing gardens, called swales, on an urban map. Groups were given twelve minutes to create multiple swale configurations, which they tested by viewing a simulation based upon their placements. The groups had to balance ground rainwater infiltration against swale cost. To achieve a high combined score, the participants had to become more effective at placing swales in locations that would more efficiently capture more of the rainfall. The rain fell evenly across the map in the simulation, but the combination of ground elevation and man-made features (like paved roads and water-diverting sewers) produced emergent flows and pools of surface water. Participants could see these flow paths and pools by watching the simulation's visualization as it ran. A simulation run terminated automatically when there was no longer any surface water (having either drained into sewers, infiltrated into the groundwater supply via swales, or flowed off of the map through a "sink" - a low-elevation point at the edge of the simulated map).

Groups performed this task in three input conditions (tangible interface, single-mouse, and multi-mouse). A repeated-measures-with-rotation design was used to alternate the order of input conditions as well as to alternate the specific urban maps. Participation was incentivized by cash paid to each member based on the distance of the group's best combined infiltration and swale cost scores from the ideal solution. In the tangible interface condition participants were given physical tokens representing swales that they could place on a paper map representing the landscape. An overhead camera recorded the configurations and sent the input to a computer which ran the simulation for the participants to view on a display placed behind the map. In the mouse input conditions, participants used one or three mouse controllers to place swales on the landscape represented on the screen of a shared laptop. The results of all configurations were exported at the click of a button to the simulation computer, so in all conditions the simulation run was witnessed at the same viewing angles and screen size. The three maps given to participants were determined to be equally difficult but different in surface details to mitigate a practice effect.

We wanted to know whether or not the participants became any better at placing the swales in response to the observable emergent patterns in surface water flow, so we constructed a "dynamism" measure for each grid square on each map. We defined "dynamism" to be the amount of water inflow each grid square received from its neighbors over the length of the simulation run, in gallons, which we then normalized by dividing by the highest inflow value obtainable on that particular map (these maximum values were roughly equivalent across the three maps). To assess participants' placements in a particular trial, we computed an average normalized dynamism value, summing the normalized dynamism values of all of their chosen swale locations and dividing by the number of swales placed. A higher average normalized dynamism value indicated that participants were placing swales in locations that were more effective at trapping water for infiltration, whereas lower dynamism values indicated that participants were less successful in placing their swales in locations that would intercept surface water. We noticed that participants produced solutions that had higher average normalized dynamism in the tangible condition than the other conditions (see results section) which prompted us to select two representative cases to examine more deeply.

All sessions were videotaped in order to quantify the number of different configurations and qualitatively analyze the conversations and actions. For the case study (Yin, 2003), we selected two groups that were representative of the higher average normalized dynamism observed in the tangible condition, but which were different in all other ways. They were of different age groups (undergraduate vs. high school) and genders (males vs. females), and experienced the tangible condition in opposite order (tangible, single mouse, multi-mouse vs. single-mouse, multi-mouse, tangible). They also showed different relative dynamism values for

single-mouse and multi-mouse (higher for single-mouse in Group 1, higher for multi-mouse in Group 2) despite their mouse conditions being ordered identically (single-mouse followed by multi-mouse). We transcribed and coded the videotaped sessions for the two groups using a multimodal format adapted from Goodwin (2003) in order to examine both speech and gesture. We also segmented participants' speech into *utterances* following a procedure similar to Kintsch (1998) where each utterance is defined as a meaningful unit that expresses a proposition or sentiment. Therefore, even when a participant expressed agreement through a single word (e.g., "okay") in a single conversational turn, this was counted as an utterance. Utterances were then coded for evidence of contributions to collaborative problem solving for the following categories: 1) asserting an idea, 2) expressing agreement, and 3) expressing disagreement.

Gestures were annotated following a procedure modified from McNeill (1992) for their timing with utterances, type, and description. Gesture types included deictic, metaphoric, iconic, and iconic-tracing gestures. As explained earlier, we added an "iconic-tracing" category to emphasize *both* the emergent deictic aspect for actively pointing out references, and the iconic aspect to describe the visuospatial or dynamic characteristics of the references. Iconic gestures were distinguished by their holistic representation of an object, whereas iconic-tracing gestures were distinguished by their schematic use to trace out the shape or flow of an object. Purely iconic and metaphoric gestures did not appear significantly in the data and are not included in the analysis.

Findings

In this section, we first present the overall quantitative dynamism results, followed by the qualitative findings for Group 1 and Group 2. These cases are presented separately with an overview of the three sessions' group dynamics and gesture use, followed by a more detailed presentation for the tangible interface session. In the final section, Groups 1 and 2 are discussed together. Note that group members are referred to by the Red, Green, or Yellow color wrist bands worn during the sessions (e.g., Group 1 members are Yellow1, Green1, and Red1).

Quantitative Dynamism Results

The average normalized inflow values seen across all 29 groups of three participants was 0.074 (SD = 0.087) for paper, 0.065 (SD = 0.088) for multi-mouse, and 0.046 (SD = 0.045) for single-mouse, which showed a significant effect of interface style on the ability to target high-dynamism locations for swale placement [$F(2,581) = 7.15, p < 0.001$]. Post hoc comparisons using the Tukey HSD test indicated that the paper condition differed significantly from the single mouse condition ($p < 0.01$), although there were no significant differences between the paper and multimouse or between the multimouse and single mouse conditions.

Group 1 Overview of Three Sessions

Group 1 consisted of three male college students. The order of conditions that Group 1 received was: 1) tangible interface, 2) single mouse, and 3) multi-mouse. Table 1 below presents numeric summaries of the communicative output across the three conditions. From this data, it is evident that Yellow1 dominated for utterances in every session, and for gestures in the second and third sessions—although mostly with deictic gestures. Red1 produced the most gestures in the first session with many iconic-tracing gestures, but both his utterances and gestures dropped precipitously over the following two conditions. Green1 consistently produced a moderate number of utterances and smaller number of mostly deictic gestures, and appeared to grow more assertive over the three sessions.

Table 1. Group 1: Communicative output per condition

Data	Tangible Interface			Single Mouse			Multi-Mouse		
	Yellow1	Green1	Red1	Yellow1	Green1	Red1	Yellow1	Green1	Red1
Total # of Utterances	74	47	55	54	32	23	73	65	56
Asserting Idea	10	8	5	11	9	3	9	8	4
Express Agreement	7	4	16	8	5	10	7	5	15
Express Disagreement	4	4	3	4	4	1	3	4	3
Total # of Gestures	13	7	16	11	7	3	7	2	2
Deictic	10	6	9	8	7	3	7	2	1
Iconic-tracing	3	1	7	3					1

Fortunately the video data provides rich detail about how Red1's iconic-tracing gestures helped him to consider dynamic spatial information, and assert his ideas verbally to the group. As will be described further below, Red1's assertion of his main idea in the tangible interface session also effected an interesting moment of convergence and agreement in the group.

Group 1, Session 1: Tangible Interface

In Group1’s first trial, Yellow1 led by stating his understanding of the pattern of rainwater absorption he observed in the first run of the simulation, and by correcting Green1’s orientation, who then realized he had incorrectly transposed the landscape visualized on the simulation screen to the landscape represented on the paper map. Yellow1 also initiated placing the first swales and suggesting that they pursue a strategy of either lining a street with swales or alternating swales in a checkerboard pattern (see Table 2, line 1). Green1 immediately joined Yellow1 in placing swales, but Red1 hesitated and made his first assertion (line 2) which was not picked up by the others to spread out the swales more. This early group interaction (Table 2) shows Yellow1 and Green1’s agency in action and Red1’s attempt to make an assertion which was ignored. Also, Red1 provides an early verbal expression of an idea about ‘spreading’ that he only reasserts successfully later after first inscribing it in gesture.

Table 2. Group1: Articulating first ideas and actions.

1	[02:18.20] Yellow1: (<i>Yellow1 and Green1 are placing swales on map as Yellow1 is talking</i>) So we should create this and see if we just need lines? Or do we want to alternate? [02:23.04]
2	[02:23.05] Red1: Or do you wanna spread em out maybe?... [02:25.00]
3	[02:25.13] Green1: Actually [02:26.03]
4	[02:26.10] Yellow1: Or to go, go opposite?[02:26.19]
5	[02:33.00] Green1: We also got quite a bit over here [02:33.08]
6	[02:36.04] Red1: (<i>Still has not made any placement yet</i>) Yeah (<i>pause</i>) should we...[02:40.23]

Over the next trial, Red1 produced gestures as if considering the flow of water over the landscape but accompanied by minimal or no speech. Such gestures indicate that they were made *for thinking* (McNeill, 2005) rather than for speaking. In Table 3, frame 3a (below), Yellow1 had just finished pointing out how the water ends up in the lower left quadrant of the map (not pictured) when Red1 moved his left hand in a wave-like motion to start an iconic-tracing gesture at [06:41.01] before he actually spoke at [06:42.20] saying only, “and um...” Through this gesture, Red1 appeared to be considering the emergent flow of water as he held his left hand palm down and fingers splayed out, moving over the line of swale placements down towards the left quadrant Yellow1 had just pointed out.

Table 3. Red1 Gestures for thinking examples.

 <p>3a.</p>	 <p>3b.</p>
<p>[06:41.01] Red1: {... and um...} (<i>Red1 iconic-tracing gesture, starts speaking at [06:42.20]</i>)</p>	<p>[08:48.20] Red1: {um I don’t feel like...} (<i>Red1 iconic-tracing gesture</i>)</p>

Two minutes later in Table 3, frame 3b (see above), Red1 moved his left hand out along the left line of swale placements again and said, “um I don’t feel like...” With this iconic-tracing gesture, Red1’s palm is vertical to the map and loosely cupped, as if both considering the impact of the water flow down the street and how much could be contained (cupped image) in this area. Table 4 below shows how Green1 then asserted a disagreement with the current configuration (line 1), which gave Red1 an entry for asserting his idea leading in with a pointing gesture (line 2) before speaking (line 4). As soon as Red1 pointed out the area of concern (lines 4, 5), there was an immediate convergence of overlapping speech, agreement, and pointing gestures by all three group members. Red1 then more fully articulated his idea for spreading out the swales (line 9), and making it “a gradual change” (lines 15, 16), which echoed his initial assertion that was ignored at [02:23.05] (Table 2) to “spread them out.” However, Red1’s idea and reasoning is more clear here, and presented an effective spatial strategy for reducing the emergent water flow to the zone that Yellow1 and Green1 had lined with a checkerboard of swales.

Table 4. Red1 articulates idea of “spreading” and “gradual change.”

<p>1. [08:48.24] Green1: There's a little bit too much. 2. [08:49.20] Red1: {Yeah} (<i>moving right hand in, pointing gesture</i>); 3. [08:50.21] Yellow1: Okay. 4. [08:54.04] Red1: ... {So let's, um, right here like, right here} (<i>long pause</i>)</p>		
		
<p>5. [08:56.04] Red1: [{seems like right here}]. 6. [08:56.04] Yellow1: [{These two are neighboring}]. 7. [08:56.10] Green1: [{Yeah}]. 8. [08:56.20] Red1: [{They're so close] and that's neighboring}</p>	<p>9. [08:58.10] Red1: that we could easily just take it away (<i>pause</i>) 10. [09:00.17] Red1: [like that]. 11. [09:00.17] Green1: [also, it didn't seem] to have much of an effect in that area.</p>	<p>12. [09:04.14] Yellow1: Let's spread this out maybe, we can get rid of this one? 13. [09:07.16] Green1: Yeah seems like that would be better (<i>Red1 and Green1 look up and check the simulation</i>).</p>
		
<p>14. [09:10.22] Red1: Yeah {because its concentration is centered here}</p>	<p>15. [09:12.00] Red1: at least make {it like a gradual (<i>pause</i>) change}.</p>	<p>16. [09:13.26] Yellow1: Mm-hm (<i>agrees</i>).</p>

Group 2 Overview

Group 2 consisted of three female college students. The order of conditions that Group 2 received was: 1) single mouse, 2) multi-mouse and 3) tangible interface. Table 5 presents a summary of the communicative output for all three conditions. From this data it appears that Green2 dominated in every session for utterances, but gestured only in the first two mouse input sessions with mostly deictic pointing gestures. Yellow2 was consistent over the sessions but produced more utterances and gestures in the mouse conditions than Red2, who was inconsistent by producing many utterances in the first session (although few gestures because she was in control of the single mouse), but then dropping in utterances and gestures in the second session. In the third session, Red2 dramatically increased her number of utterances and produced a large number of iconic-tracing gestures, unlike her partners.

From the video data for Group 2, it is evident that Red2 began to use iconic-tracing gestures in the third (tangible) session to better articulate her thinking about the emergent spatial flows of rainwater. Over the three sessions, Red2 appeared to have difficulty in verbally articulating her spatial reasoning, but by the tangible interface session, Red2 became more insistent and better able present her ideas using iconic-tracing gestures as a support.

Table 5. Group 2: Communicative output per condition.

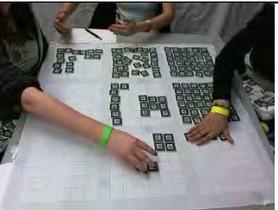
Data	Single Mouse			Multi-Mouse			Tangible Interface		
	Yellow2	Green2	Red2	Yellow2	Green2	Red2	Yellow2	Green2	Red2
Total # of Utterances	60	103	89	69	126	38	40	93	78
Asserting Idea	11	18	6	5	16	5	6	18	13
Express Agreement	9	5	8	8	5	5	9	5	8
Express Disagreement	1	6	2	0	7	2	1	14	8
Total # of Gestures	23	42	4	10	14	5	5	13	26
Deictic	21	35	2	9	12	3	5	13	12
Iconic-tracing	2	7	2	1	2	2			14

Group 2, Session 3: Tangible Interface

In this session, Green2 made the first move by asserting her idea of placing swales on the “dark spots” (i.e., low elevation locations) which led to a strategy pursued by all three group members to fill up an entire quadrant on the map with swales, creating a catchment effect (where water would pool over a cluster of swales). However, after viewing the simulation results from the initial swale configuration, Red2 referred to the edge of the map closest to her and tried to assert a different idea. At [07:12.03] she said, “Let’s just try...” but because Yellow2 and Green2 were busy talking, she went ahead and started placing swales near her edge. Red2 was more specific at [08:29.26] with, “I don’t know... I really want it along the edge though right here (*pause*) really, really, really want that!” However, when Red2 added iconic-tracing gestures, she started to elucidate her reasoning.

Table 6 below provides examples of these gestures and the added spatial information that supported her reasoning. At [08:48.03] (Table 6, frame 6a), Red2 first argued, “Because I saw the water coming down like this,” while moving her right hand in a waving motion over the middle section of the paper map towards her edge. About 20 seconds later (not pictured in Table 6), Red2 then repeated this gesture for herself without speaking. By [12:13.21] (Table 6, frame 6b), Red2 had picked up more emergent information about the water flows and used two hands to trace the flows down farther along the right side of the map, and then with two cupped hands to move along her edge. Red2 was more sensitive to the location of the flows now and why there was so much pooling towards her edge. While Green2 had been consistently arguing against Red2’s concern for the edge, by [14:03.04] Green2 gave in and agreed, although out of friendly exasperation, “Alright! Just try things!” In Group2, although there wasn’t a clear moment of group convergence, Red2’s use of iconic-tracing gesture enabled her to sensitize to emergent spatial flows in the system and better articulate her spatial reasoning, which perhaps gave her the confidence to keep asserting her ideas until they became incorporated by the group.

Table 6. Red2 adds spatial information to her gestures.

 <p>6a.</p>	 <p>6b.</p>	 <p>6c.</p>
<p>1. [08:48.03] Red2: Because I saw the water coming down like this.</p>	<p>1. [12:13.21] Red2: The water at the end was towards the edges...</p>	<p>2. [12:14.11] Red2: ...and onto the streets. 3. [00:12:19.05] Red2: I don't think it was up there though (<i>to Yellow2</i>).</p>

Discussion

As mentioned earlier, we chose Groups 1 and 2 to compare on the basis of having similar dynamic outcomes despite different factors like age or gender or receiving the tangible interface condition in opposite order, which potentially rules out the effect that familiarity with the task would have on participant performance in a given input condition. Nevertheless, we found similar interaction patterns in the tangible interface condition for the two groups in that the least dominant group member of each group, Red 1 and Red2, experienced failure at first in articulating and asserting their ideas verbally, but then succeeded after using iconic-tracing gestures to pick up and describe emergent spatial information. What were the particular representational affordances of the tangible interface that enabled this iconic-tracing gesture based spatial sensitivity and reasoning? In both cases, it appeared that Red1 and Red2 visualized patterns of water flows *over* the paper map, indicating that the paper map provided a material support for their visualization. From a distributed cognition framework, Hutchins (2005) describes how material forms serve as *material anchors* for stabilizing conceptions, and the paper map appeared to serve this function. Similarly, the physical swale tokens also provided material anchors for visualizing water infiltration strategies. The physicality of the tokens may even have provided a representation of depth to the landscape. In fact at one point, Red2 began to stack the physical swale tokens on top of each other because she wanted to address the greater need an area where water was pooling. But even given the ways that the tangible interface functioned as a material anchor, it was nevertheless the use of iconic-tracing gestures that coordinated between what would be a mental visualization and the material anchor for the visualization. This coordinating function of gesture in bridging between the imagination and the material world has been noted in cognitive linguistics studies (e.g., Williams, 2008), and deserves further attention in educational technology design research.

Another similarity between Red1 and Red2 was their focus upon the pattern of emergent water flows over the landscape, although Red1 identified a different spatial strategy than Red2. Red1 expressed a “spread out” idea that reflected a sponge-like solution, anticipating the flow of rainwater and absorbing it as early as

possible through a gradual distribution of swales. Red2 expressed a barrier-like solution of adding a cluster of swales to a point right before her catchment area to absorb and redirect the influx of rainwater. In fact, there was actually no single best solution for the rainwater infiltration problems presented to the groups because solutions depend upon both fixed landscape features like the gradients, streets, and sewer locations as well as the dynamic swale placements (e.g., the swale “barrier” Red 2 created in Table 6, frame 6b served to both absorb and divert the surface flow into her catchment area). Therefore the better solutions could only come about through sensitivity to emergent patterns of water flow observed in the simulation, which entails a sensitivity to how patterns unfold over time.

The positive impact of both Red1 and Red2’s contributions to group spatial problem solving in the tangible interface condition was also evident in the quantitative dynamism measures (Table 7). The dynamism values are highest for both groups in the tangible interface condition, mirroring what was seen across the 584 trials generated by the other 81 participants in the study. For Group 1, the effect of the interface is significant according to a one-way ANOVA [$F(2,14) = 138.76, p < 0.0001$], with a *post hoc* Tukey HSD test confirming significance for Tangible vs. Single-mouse ($p < 0.01$), Tangible vs. Multi-mouse ($p < 0.01$), and Single-mouse vs. Multi-mouse ($p < 0.01$). For Group 2, although the interface condition did have a significant effect [$F(2,23) = 11.26, p < 0.001$], this held only for the Tangible vs. Single-mouse ($p < 0.01$) and Multi-mouse vs. Single-mouse ($p < 0.01$).

Table 7. Dynamism measures: Groups 1 and 2.

Group 1			Group 2		
Session 1: Tangible	Session 2: Single mouse	Session 3: Multi-mouse	Session1: Single mouse	Session 2: Multi-mouse	Session 3: Tangible
0.0464	0.0246	0.0181	0.0077	0.0270	0.1108

Implications and Conclusion

This study’s findings indicate that a tangible interface provides a beneficial affordance for gesture-mediated spatial reasoning, especially for less verbally participatory or assertive group members in the context of group solving of a complex spatial problem modeled on a computer based simulation. As a modality for spatial reasoning and communication, gesture (and iconic-tracing gesture in particular) increased the communicative agency of less dominant group members and resulted in equalizing group participation. In addition, the greater sensitivity to emergent spatial information that these less dominant group members articulated through gesture first and then speech had the effect of improving group spatial problem solving performance. While the role of gesture in spatial reasoning and group communication has been studied before, we have focused on the particular benefit of iconic-tracing gestures for working with emergent patterns, and how a tangible interface supported such gesture use.

We speculate that the materiality of the elements of the tangible interface, including the paper map and physical tokens, provided a *material anchor* (Hutchins, 2005) for stabilizing what would otherwise be only mental visualizations of emergent water flows and varying spatial features of urban landscapes, and that iconic-tracing gestures provided the bridge to link participants’ visualizations to these material anchors. Understanding tangible interfaces in terms of material anchors and the use of gesture brings useful elements from distributed cognition and cognitive linguistics frameworks into the design of interfaces for computer based group learning.

Recently, several studies have demonstrated that students who were required to imitate the gestures that a teacher performed for solving algebraic math problems did significantly better on subsequent tests than controls, and researchers speculated that the visuospatial reasoning strategies inscribed in gesture provided the benefit (Goldin-Meadow et al., 2009). To extend our current findings, we propose to borrow this strategy of intentional gesture by testing the current simulation in science classrooms with teachers intentionally modeling specific iconic-tracing gestures when introducing the simulation to students. In addition, teachers could simultaneously add conceptual vocabulary to the gestural communication (e.g., words such as ‘surface flow’ and ‘catchments’), which would ground the terminology in relevant multimodal imagery (Alibali & Nathan, 2011). Similarly, another possibility for future research is to design a tangible interface that necessitates the performance of iconic-tracing type gestures for manipulating the spatial parameters of the ecosystem simulation. The one implication that is important to communicate from this study is that both individual group members and rest of the group can be helped by more opportunities to draw upon the dual function benefits of iconic-tracing gesture for both actively reasoning about emergent spatial phenomena and communicating and coordinating the reasoning with others. The emergent properties of complex systems have proven to be challenging for many learners (Sweeney & Serman, 2007) and if gesture can provide an integral modality for reasoning about emergent phenomena, then the design of educational interventions—whether in the form of teacher communication strategies or tangible interfaces for computer based simulations—should consider how iconic-tracing and other types of gestures can play a role.

References

- Alibali, M., & 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(4), 468–523.
- Alibali, M. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition & Computation*, 5(4), 307–331.
- Alibali, M., & Nathan, M. J. (2011). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences*, 1–40.
- Antle, A. N., Droumeva, M., & Ha, D. (2009). Hands on what?: comparing children's mouse-based and tangible-based interaction. In *Proc. IDC '09*, ACM, 80–88.
- Goldin-Meadow, S. (2003). *Hearing gesture: how our hands help us think*. Cambridge, MA: Harvard Univ. Press.
- Goldin-Meadow, S., & Beilock, S. L. (2010). Action's Influence on Thought: The Case of Gesture. *Perspectives on Psychological Science*, 5(6), 664–674.
- Goodwin, C. (2003). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 217–242). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics*, 37(10), 1555–1577.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. NY: Cambridge Univ. Press.
- Maher, M., & Kim, M. (2005). Do Tangible User Interfaces Impact Spatial Cognition in Collaborative Design? In Y. Luo (Ed.), *Cooperative Design, Visualization, and Engineering*, (3675) 30–41. Springer Berlin/Heidelberg.
- Marshall, P., Hornecker, E., Morris, R., Sheep Dalton, N., & Rogers, Y. (2008). When the fingers do the talking: A study of group participation with varying constraints to a tabletop interface. In *Proc. of Tabletop 08*, 37–44.
- McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. Chicago: University of Chicago Press.
- McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- National Research Council (2006). *Learning to Think Spatially: GIS as a Support System in the K-12 Curriculum*. Washington, D.C.: National Academies Press.
- Pozzer-Ardenghi, L., & Roth, W. (2007). Performing concepts during science lectures. *Science Ed.*, 91(1), 96–114.
- Radinsky, J., Goldman, S., & Singer, M. (2008). Students' sense-making with visual data in small-group argumentation. In *Proc. ICLS'08*, 237–245.
- Rogers, Y., Lim, Y., Hazlewood, W., & Marshall, P. (2009). Equal Opportunities: Do Shareable Interfaces Promote More Group Participation Than Single User Displays? *Human-Computer Interaction*, 24(1), 79–116.
- Rogers, Y., & Lindley, S. (2004). Collaborating around vertical and horizontal large interactive displays: which way is best? *Interacting with Computers*, 16(6), 1133–1152.
- Roth, W.-M., & Lawless, D. (2002). Scientific investigations, metaphorical gestures, and the emergence of abstract scientific concepts. *Learning and Instruction*, 12, 285–304.
- Roth, W.-M., & Welzel, M. (2001). From activity to gestures and scientific language. *Journal of Research in Science Teaching*, 38(1), 103–136.
- Singer, M., Radinsky, J., & Goldman, S. R. (2008). The Role of Gesture in Meaning Construction. *Discourse Processes*, 45(4-5), 365–386.
- Stahl, G. (2002). Contributions to a theoretical framework for CSCL. In *Proc. CSCL'02*, 62–71.
- Sweeney, L. B., & Serman, J. D. (2007). Thinking about systems: student and teacher conceptions of natural and social systems. *System Dynamics Review*, 23(2-3), 285–311.
- Williams, R. F. (2008). Guided conceptualization: Mental spaces in instructional discourse. In T. Oakley & A. Hougaard (Eds.), *Mental Spaces in Discourse and Interaction* (pp. 209–234). Amsterdam: John Benjamins.
- Yin, R. (2003). *Applications of case study research* (2nd ed.). Thousand Oaks: Sage Publications.

Acknowledgements

This work is funded by National Science Foundation REESE grant 1020065.
Design and implementation of simulation and experiment by Tia Shelley, Chandan Dasgupta, and Brian Slattery.

Teacher framing, classroom collaboration scripts, and help-seeking and help-giving behaviors

Eleni A. Kyza, Yiannis Georgiou, Cyprus University of Technology
Eleni.Kyza@cut.ac.cy, Ioannis.Georgiou@cut.ac.cy
Demetra Hadjichambi, Andreas Hadjichambis, Cyprus Ministry of Education and Culture
demhad@ucy.ac.cy, a.chadjihambi@cytanet.com.cy

Abstract: This case study investigated students' collaborative help-seeking and their teachers' help-giving behaviors in inquiry-based learning. Data from two pairs of middle-school students, using two different scaffolding scripts, and from their biology teacher, were collected and analyzed. The following research questions were pursued: How does each collaboration script influence students' help-seeking and teachers' help-giving activity? Data included videotapes of each pair's interactions, the discussions between the pairs and the teacher, whole-class discussions, learning assessments and a teacher interview. Findings indicated that the pair in the *ImplicitScaffolding* script sought help less frequently than the *ExplicitScaffolding* pair while the nature of the help sought was different. Findings also showed that the different scaffolding scripts impacted student motivation and framed the teacher expectations differently, regardless of the type of help sought by the students. These findings highlight the connection between collaboration scripts, teacher cognition and scaffolding, and bear implications about students and teachers.

Introduction

While inquiry is widely accepted as a means to facilitate learning in science, inquiry-based tasks can still be highly challenging for students, especially when students are not appropriately and sufficiently supported (Kollar, Fischer, & Slotta, 2007; Linn, 2006). Research has provided evidence that help-seeking is often related to better learning (Nelson-Le Gall, 1981) while at the same time there is evidence that students do not often engage in such activities (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003). Effective help seeking strategies could promote student learning in such inquiry-based learning environments (e.g. Aleven et al., 2003). Considering that help seeking is an important meta-cognitive skill (Nelson-LeGall, 1981; Newman, 1994) several researchers have focused on designing learning environments promoting help-seeking behaviors (Aleven & Koedinger, 2001), tried to explicitly teach this skill to students (e.g. Ames, 1983; Nelson-LeGall, 1981; Newman, 1994) and tried to motivate students to seek help more often. While students' help-seeking behavior has received some attention in the literature, it appears that teachers' help-giving behaviors have been largely ignored in such discussions (Butler, 2006).

The present study is a multiple case-study into students' collaborative help-seeking behavior. It examines students' collaborative help-seeking and teacher's help-giving in two comparable contexts, each using a different classroom script to support student collaboration. The first context involved a guided inquiry environment, in which students collaborated to address a biological problem. This environment (*ExplicitScaffolding*) relied on the teacher, peer collaboration, and worksheets to support the learning process. The second context involved a less-structured inquiry environment, in which students collaborated to address the same biological problem on the computer. This environment (*ImplicitScaffolding*) relied on the web-based inquiry investigation, peer collaboration and the teacher for guidance. We investigate the following research questions: How does each classroom script influence the framing of the teaching and learning activity? Specifically, a) how does the group help-seeking behavior differ in explicit vs. implicit scripted learning environments and b) how does the teacher help-giving behavior differ in such context?

Theoretical Framework

Our work draws on literature conducted on small group collaboration scripts, students' self-regulated learning as it relates to help seeking, scaffolding and teacher cognition. In this study, we specifically seek to explore whether the collaboration script in inquiry-based learning, a topic discussed in Makitalo-Siegl, Kohnle and Fischer (2011) influenced the help-seeking behavior of the students and the help-giving behavior of the teacher. Collaboration scripts seek to structure collaborative activity in meaningful ways and support intersubjective learning (Kobbe et al., 2007; Kollar, Fischer, & Friedrich, 2006; Kollar, Fischer, & Slotta, 2007). Kobbe et al. (2007) identified five important components of collaboration scripts: learning objectives, type of activities, sequencing, role distribution, and representation type. Whereas the Makitalo-Siegl et al. (2011) study explored

two different versions of collaboration scripts both provided by the Web-based Inquiry Science Environment (WISE) learning platform we wanted to explore students' help-seeking behavior in more traditional (learning environment+worksheets) setting versus in more open-ended, yet scaffolded versions (web-based learning environment + web-based scaffolded articulation spaces).

Makitalo-Siegl, Kohnle and Fischer (2011) investigated whether students, working in collaborative inquiry settings employing high and low-structured classroom scripts, appropriately used available help. Confirming prior studies, Makitalo-Siegl et al. found that students did not engage in much help-seeking in either context. Furthermore, exploring the effect of low-structure and high-structure collaboration scripts provided through the WISE platform, Makitalo-Siegl et al. found that students in the low-structure situation engaged in statistically significant more help seeking than students in the high-structure situation, and that students in the low-structure situation exhibited smaller learning gains than the students in the high-structure condition.

While teacher scaffolding (e.g. Tabak, 2004) and the use of technology to scaffold student learning (e.g. Quintana et al., 2004; Lu, Lajoie, & Wiseman, 2010) are topics that have received significant attention to date, teachers' help-giving behavior, in response to student help-seeking is a topic that has not been discussed much. We believe that coupling the investigation of the two phenomena of interest (students not seeking help much even though help-seeking is positively associated with learning and teacher cognition in providing help to support collaborative inquiry learning) merits further examination.

Methodology

Participants

Teacher

One biology teacher, with five years of teaching experience and a master's in science education, and her 7th grade students at an urban middle-school were involved in this study. The teacher (Mrs. Tonia, a pseudonym) was a member of a participatory design group who collaborated to design the learning materials with the goal of supporting inquiry learning; she did not participate in the planning of this research study nor was she aware of the specific research questions. Mrs. Tonia enacted the learning module using the *ExplicitScaffolding* collaboration materials with one of her classes and used the web-based, *ImplicitScaffolding* materials with her second class. The teacher was familiar with both approaches to inquiry-based learning.

Students

Two intact classes, taught by the same teacher (Mrs. Tonia) participated in this study. Conceptual understanding assessments and motivation surveys were administered to all students following a pre-post design. The analysis of the conceptual understanding pre-test indicated no statistically significant difference in terms of the two classes' content knowledge before the teaching intervention ($Z=-.052$, $p=.958$), indicating that students in each collaboration script were equivalent in terms of their prior knowledge regarding the human reproductive system. Similarly, the comparison of students' motivation in the two collaboration scripts did not indicate any statistically significant differences in terms of students' views about a traditional biology lesson ($Z=-1.191$, $p=.234$) or about an ideal biology lesson ($Z=-1.772$, $p=.076$). Hence, these results, confirmed that both classrooms were equivalent in terms of students' motivation.

Two pairs of students (one from each of the classes taught by the same teacher) formed our case study groups in this study. Students in each pair were 12-year-old girls of comparable academic performance and motivation to engage in science learning, as assessed by their teacher and as confirmed by their grades in language arts and science education. The pairs were selected by the teacher due to their comparability and to their willingness to be videotaped. Data from two other pairs were also collected but were not analyzed for this paper.

Implementation context

Intervention

The teaching intervention consisted of seven, 80-minute sessions. Students, assuming the role of a doctor, investigated a scenario-driven problem that dealt with human reproduction; according to this scenario students were involved with a socio-scientific investigation, gathering, interpreting and synthesizing secondary information in order to take an evidence-based decision and to offer advice to a young couple seeking medical advice in order to have a baby.

Collaboration Scripts

Two collaboration scripts, one in each classroom, were employed. In the *ExplicitScaffolding* script each pair worked with learning materials and worksheets and relied on the teacher, on peer collaboration and on the worksheets to progress with solving the problem posed to them. In the *ImplicitScaffolding* script each pair had access to the same learning materials organized using a web-based educational application and used online articulation spaces with general type prompts to organize their inquiry. Students relied on prompts provided online, on peer collaboration and on the teacher for guidance. In both conditions, the teacher was available during the sessions and walked around the classroom in a non-systematic manner while students were working in pairs. At the end of the unit, a plenary discussion was led by the teacher in both conditions. One researcher was present during the lessons to collect enactment data from each classroom script.

Data Coding and Analysis

All classroom interactions were videotaped, including discourse at the level of the pair and whole class discussions with the teacher. The data were transcribed verbatim and were coded using qualitative analysis software. A forty-minute interview with the teacher after the implementation investigated how she considered her role in each learning environment in terms of scaffolding the students. Conceptual understanding achievement tests and motivation surveys were also administered prior to and after the learning intervention to students in both classrooms.

We coded all instances of each pair's help-seeking and teacher's help-giving behavior using a coding scheme informed by prior studies (Makitalo-Siegl, Kohnle & Fischer, 2011). The coding was performed at episode level and focused on six areas: context of help-seeking, type of help sought, content of help sought, type of help provided, recipient of help, and usage of help. All coding categories and their definitions are shown in Table 1. An inter-rater reliability check performed by two of the authors yielded a score of 90% agreement. All disagreements were first discussed and resolved between the two coders; subsequently, all data were coded accordingly by one of the coders.

Table 1: Help-seeking and help-giving coding scheme

Coding category	Definition
1. Context of help-seeking	This category describes the context in which students seek for helping
1.1 Student-centered	The students ask for help while collaborating for their investigation
1.2 Teacher-centered	The students ask for help while participating into a whole-class discussion led by the teacher
2. Type of help sought	This category characterizes the type of help sought by the students
2.1. Executive	Students request direct answers such as "is this right"
2.2. Instrumental	Students request help to support them figuring out the problem on their own
3. Content of help sought	This category characterizes the content of help sought by the students
3.1 Domain knowledge	Students ask for help regarding a domain related issue (e.g. reproductive system)
3.2 Inquiry skills	Students ask for help regarding an inquiry skill (e.g. how to formulate hypotheses)
3.3 Procedural issues	Students seek help regarding procedural or writing issues (e.g. asking for the repeat of a statement so that they can copy it correctly, discussing grammar, etc.)
4. Type of help provided	This category characterizes the type of help given to the students
4.1 Executive	Direct answers are provided as help
4.2 Instrumental	Students are scaffolded in finding the answer on their own
4.3 None	No help is provided to the students
5. Recipient of help	This category describes how help giving is provided
5.1 Help provided to the group	Help is provided to the students who sought it
5.2 Help provided to the whole class	Help is provided to the whole class, in response to the students' help seeking

6. Usage of help	This category codes whether students have followed up on the advice provided by the teacher or other peers.
6.1 Help used	The students utilize the help given
6.1 Help not used	The students do not utilize the help given
6.2 Insufficient data	The available data do not allow determining whether the help provided was utilized by the students

After the initial identification and coding of episodes (n=42), episodes were analyzed for emerging patterns, examining the data first at the episode level and then contrasting episodes with each other. Episodes presenting similar patterns were then grouped to determine dominant trends in each of the two scripted environments. This analysis represents the crux of what we are interested in, as we seek to explore the relationship between scripts, students' collaborative behavior and the teacher's response.

Findings

The analysis of the data informs our case study approach (Yin, 2003), with the multiple sources of data and the different data analyses serving to triangulate our findings. In this section we first present findings on the pairs' help-seeking behavior and then present the results of the pattern analysis which examines students' help-seeking behavior as coupled with the teacher's help-giving behavior.

Help-seeking

Of the 42 episodes identified, 27 belonged to the explicit and 15 to the implicit scaffolding script. Whereas all help-seeking in the implicit scaffolding script (*ImplicitScaffolding*) took place in the context of the collaboration of the pair, 12 of the 27 episodes (44%) of the explicit scaffolding script (*ExplicitScaffolding*) occurred in the context of whole-class discussions with the teacher. The *ExplicitScaffolding* pair sought executive help more often, seeking directly the "right answer" whereas the *ImplicitScaffolding* pair sought more instrumental and procedural help. Table 2 presents these findings.

Table 2: Help-seeking and help-giving findings

	No of episodes and percentages	
	<i>ImplicitScaffolding</i> Script (N=15)	<i>ExplicitScaffolding</i> Script (N=27)
Context of help seeking		
<i>Student-centered</i>	15 (100%)	15 (56%)
<i>Teacher-centered</i>	0 (0%)	12 (44%)
Type of help sought		
<i>Executive</i>	5 (33%)	18 (67%)
<i>Instrumental</i>	10 (67%)	9 (33%)
Content of help sought		
<i>Domain knowledge</i>	5 (34%)	21 (78%)
<i>Inquiry skills</i>	0 (0%)	2 (7%)
<i>Procedural issues</i>	8 (53%)	4 (15%)
Type of help provided		
<i>Executive</i>	11 (73%)	17 (63%)
<i>Instrumental</i>	3 (20%)	3 (11%)
<i>None</i>	1 (7%)	7 (26%)
Recipient of help		
<i>Help provided to the group</i>	14 (100%)	12 (60%)
<i>Help provided to whole class</i>	0 (0%)	8 (40%)
Usage of help		
<i>Help used</i>	14 (100%)	12 (60%)
<i>Insufficient data</i>	0 (0%)	8 (40%)

Help-giving

In both collaboration scripts, the teacher provided students mainly with executive (more direct answers and explicit guidance) rather than instrumental help (73% of the episodes for the *ImplicitScaffolding* pair and 63% of the time for the *ExplicitScaffolding* pair). In 7 out of 27 episodes (26%) the pair was not provided with help

in the explicit guidance script, whereas also in 8 of the 20 episodes when help was given it was provided in the context of a whole-class discussion and was not related to the specifics of the collaborative work conducted by the pair.

Analysis of help-giving and help-seeking patterns

Each pair's help-giving and help-seeking episodes were analyzed separately to identify collaboration patterns. The dominant patterns (showing up in at least 5 episodes each) are shown in Figure 1.

ImplicitScaffolding SCRIPT	ExplicitScaffolding SCRIPT	
1.1 Student-centered context	1.1 Student-centered context	1.2 Teacher-centered context
2.2 Instrumental help sought	2.1 Executive help sought	2.1 Executive help sought
3.3 Help for procedural issues	3.1 Help on domain knowledge	3.1 Help on domain knowledge
4.1 Executive help provided	4.1 Executive help provided	4.1 Executive help provided
5.1 Help provided to the group	5.1 Help provided to the group	5.2 Help provided to whole class
6.1 Help used	6.1 Help used	6.2 Insufficient data
Pattern I (53.3 %)	Pattern I (18.5 %)	Pattern 2 (22.2 %)

Figure 1. Help-giving and help-seeking dominant patterns identified

Six different patterns were identified in the *ImplicitScaffolding* script, with one of them being a dominant pattern since it was identified in eight of the episodes. According to this pattern the two girls at several points sought instrumental help since they were asking for supportive guidance to continue with their process of inquiry. The teacher responded to the pair's help-seeking by providing executive help, helping define next steps; this help was taken up by the pair. The following brief excerpt illustrates this pattern:

[Episode 9, *ImplicitScaffolding* script]

Maria: Mrs...

Teacher: Yes?

Maria: There's an instruction here asking us to complete the template "From fertilization to birth" but there is also another template...

Teacher: Have you completed the "Menstrual Cycle" [template]?

Maria: Yes.

Teacher: Ok... Then this is the right template.

Maria: Just a moment.

Teacher: This is the right template.

Anna: Let's read [the sources] first.

Maria: No.

Teacher: Name your template first.

Maria: What should we name it?

Anna: "Birth" because it refers on the natural way of childbearing...

Maria: Fetus...

Teacher: So now you should answer all of these questions [on the template] regarding the natural way of childbearing.

The pair continues with their work.

On the other hand, the pattern analysis in the 27 help-seeking/help-giving episodes of the *ExplicitScaffolding* script revealed fourteen different patterns, two of which were identified as dominant, being detected in five and six different episodes respectively. For example, according to the first dominant pattern the pair sought executive help on several occasions since they were asking directly for the "right" answer regarding domain knowledge necessary to finish their task. The teacher provided the pair with executive help, giving "right answers"; this help was taken up by the pair, who mechanically wrote the right answers. The following brief excerpt illustrates this pattern:

[Episode 8, *ExplicitScaffolding* script]

Fiona: Mrs Tonia, should we write that [the ovaries] stop releasing ovules after a certain age?

Teacher: Yes... Because this is the way they function...

The girls record the right answer on their worksheet.

As the teacher explained in her post-enactment interview her instructional objectives remained the same in both situations: “*The aims were the same... What had changed was the approach in order to see how the development of inquiry skills, content knowledge and collaboration could be promoted. It shouldn’t be right to have different aims... I had the same aims but within a different approach in order to see the result...*”

However, the teacher felt that in spite of identical educational objectives her role should be differentiated. Mrs. Tonia felt that in the *ImplicitScaffolding* script there was no need to guide her students except with some procedural issues that concerned the details of using the templates provided. In contrast, in the *ExplicitScaffolding* script she felt that she should comment on students’ work more frequently, correcting their answers to their worksheets and providing more explicit guidance through whole class discussions.

[In the ImplicitScaffolding script] I was talking as less as I could compared to the other class. The only thing that [my students] needed was that they wanted to know what exactly they should do in some of the templates ... I didn’t want to guide them. I wanted each pair to move at their own pace...

[In the ExplicitScaffolding script] things were more guided. The whole class should stop at the same point in order to make some comments, to move forward and to check the answers to the questions they worked on by the end of the lesson. And after all, the lesson was guided by me... I mean that they followed their worksheets but we should stop at the same point, as indicated in lesson plan, in order to watch a video or to talk about how to go on... It was something very different...

The [ImplicitScaffolding] provided by the web-based tool could function more autonomously. Students could work without my own help... In contrast, in the other class, if I were not to give instructions, the students could not work so easily I guess...

Conceptual understanding achievement tests

The statistical analysis of the students’ pre- and the post-test on conceptual understanding, employing the Wilcoxon sign rank test, yielded statistically significant results for the students in the *ImplicitScaffolding* script ($Z=-2.785$, $p=.000$) as well as for the students in the *ExplicitScaffolding* script ($Z=-4.012$, $p=.000$). Therefore, it can be concluded that students in both classroom scripts presented a statistically significant difference in their learning gains, as an outcome of the learning intervention.

Motivational assessment test

The comparison of students’ responses to a survey about the motivational potential of a traditional biology lesson and the motivational potential of an ideal lesson, indicated that a traditional biology lesson was much less motivating than an ideal biology lesson; the difference was statistically significant for both the students in the *ImplicitScaffolding* script ($Z=-2.925$, $p=.003$) as well as for the students in the *ExplicitScaffolding* script ($Z=-3.142$, $p=.002$). Even though there was no statistically significant difference for the students of the *ExplicitScaffolding* script ($Z=-.735$, $p=.462$) between the inquiry lessons they attended and their traditional biology lessons, the analysis of the responses of the students in the *ImplicitScaffolding* script indicated a statistically significant difference between the inquiry lessons and their traditional biology lessons ($Z=-2.591$, $p=.010$), in favor of the inquiry lessons.

Discussion and Implications

Help seeking processes can be affected by patterns of classroom interaction and facilitated by instruction (Aleven et al., 2003; Karabenick & Newman, 2009; Oortwijn, Boekaerts, Vedder, & Strijbos, 2008). However, there has been little research on the question concerning how different patterns of classroom interaction may differentiate the process of help-seeking and help-giving to date. Aiming to shed some light to this underexplored area, the present study focused on help-seeking / help-giving behaviors using two different classroom scripts to teach an inquiry-based biology learning module.

According to our findings, the *ImplicitScaffolding* scripts provided our case study students with the necessary structure and guidance and supported their learning. In contrast, the *ExplicitScaffolding* classroom

script constrained the pair's autonomous inquiry, as the students' help-seeking and help-giving behavior showed. The differentiation observed in the process of students' help seeking could be partially attributed to each classroom script, which afforded different teacher and students' interactions that differentially shaped the process of help-seeking and help-giving. The dominant patterns of student help-seeking and teacher help-giving behaviors were different in each scaffolding script. In the case of the *ImplicitScaffolding* script, the pair working with the web-based learning environment framed their inquiry as a multi-step process. In this context, students asked mainly for instrumental help when they felt uncertain about how to carry forward with their investigation. In contrast, in the *ExplicitScaffolding* script, students sought help more frequently, sought more executive help and requested direct answers relating to domain knowledge. While students in the *ImplicitScaffolding* script sought less help when compared to the help-seeking behavior of the students in the *ExplicitScaffolding* script, no statistically significant differences were found between the classes and both classes improved on the conceptual understanding test. This suggests that, perhaps, students in the *ImplicitScaffolding* script voiced fewer questions because the scaffolds may have afforded a process that helped them answer questions in the context of collaboration and thus reducing the need for teacher guidance. Such results are supported by discussions about the reflective affordances of articulation tools and processes (e.g. Chi, de Leeuw, Chiu, & Lavancher, 1994; Quintana et al., 2004).

The results of all students' motivation surveys, comparing the students' views between the current (inquiry) and past (traditional) biology lessons indicated a statistically significant difference between the two collaboration scripts. The results suggest that students found the *ImplicitScaffolding* script more motivating than past instructional settings; this is important as motivation has been reported as having the capacity to result to enhanced learning (Blumenfeld, Kempner & Krajcik, 2006). As Makitalo-Siegl et al. (2011) discussed, different classroom scripts impact help-seeking behavior; Makitalo-Siegl et al. found that students who participated in collaborative learning environments providing appropriate structure and scaffolding, tended to seek help much less and to achieve better learning results. In our case, data support the first but not the second conclusion. However, it maybe that the learning assessments only capture part of the learning occurring in such contexts; indeed, if self-regulated inquiry learning is what we are pursuing the use of other assessments tapping into inquiry-related and metacognitive skills should be explored.

The teacher also responded differently in each of the classroom scripts, as suggested by the analysis of the video and teacher interview data. The teacher's behavior was framed by the context, as the teacher assumed that students in the *ImplicitScaffolding* script should be able to work more autonomously, while she felt that the students in the *ExplicitScaffolding* script needed more guidance and structure. Hence, it is no surprise that in this context, students' help-seeking/help-giving episodes in the *ExplicitScaffolding* script often took place into a more teacher-centered environment, while the opposite was true for the *ImplicitScaffolding* script. Levin, Hammer and Coffey (2009) discussed the concept of teacher framing as a way to understand teacher attention. According to Levin and his colleagues "whether and how teachers attend and respond to student thinking largely reflects how they frame what is taking place in their classes" (p.143). While Levin et al. suggested that institutional requirements frame teacher activity to selectively attend to issues of learning and teaching, curricular materials, especially new methods of teaching may also have a similar impact. This topic merits further exploration as it relates to teacher professional development through curricula that have educative properties, even when these properties are not explicitly communicated to teachers (Davis & Krajcik, 2005).

A, rather, unexpected finding is that the teacher provided executive help to most of the pairs' requests, regardless of the collaboration script. This may be explained by the finding that students in one situation (*ExplicitScaffolding*) requested such type of response whereas most of the questions students posed in the other situation (*ImplicitScaffolding*) were instrumental for their process but yet could be often answered via executive support. Future qualitative exploration of our dataset, as well as the analysis of the videotaped interactions of two additional pairs not included in this analysis (one in each scaffolding condition) can help elucidate our findings and provide better insights into the conditions under which each type of support provided may be most helpful.

These findings contribute to our understanding of students' help-seeking behavior and how to better support it, and can inform teacher professional development. In terms of student help-seeking, and as suggested by the conceptual understanding achievement tests administered in this study, it seems that environments providing *ImplicitScaffolding* may be providing the support that students need to achieve positive learning results, while in environments where this type of support is missing students necessarily turn to teachers to find that additional support necessary to complete the task. Where the participating teacher was concerned, our findings suggested that the framing of the task may have heavily influenced the teacher's perceptions of what she was expected to do and how she should guide student inquiry; both tools and teachers' reflection-on-action (Schon, 1983) are necessary to help teachers move towards improved support of students' self-regulated behavior. These findings highlight the connections between collaboration scripts, teacher cognition and scaffolding and bear implications about students and teachers. Future studies should examine the topic

employing different methodologies and large sample sizes, connecting analyses to specific tasks, and investigating the issues in other contexts.

References

- Aleven, V., & Koedinger, K. R. (2001). Investigations into help seeking and learning with a Cognitive Tutor. In R. Luckin (Ed.), *Papers of the AIED-2001 Workshop on Help Provision and Help Seeking in Interactive Learning Environments*. Available: <http://www.hcrc.ed.ac.uk/aied2001/workshops.html>
- Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research*, 73(3), 277-320.
- Blumenfeld, P., Kempler, T. M., & Krajcik, J. S. (2006). Motivation and cognitive engagement in learning environments. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 475-488). New York: Cambridge University Press.
- Butler, R. (2006). An achievement goal perspective on student help seeking and teacher help giving in the classroom: theory, research, and educational implications. In S. A. Karabenick & R. S. Newman (Eds.), *Help Seeking in Academic Settings: Goals, Groups, And Contexts*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Chi, M.T., De Leeuw, N., Chiu, M., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18(3), 439-477. doi: 10.1016/0364-0213(94)90016-7
- Davis, E. A., & Krajcik, J. S. (2005). Designing Educative Curriculum Materials to Promote Teacher Learning. *Educational Researcher*, 34(3), 3-14. doi: 10.3102/0013189x034003003
- Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., & Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, 2(2), 211-224. doi: 10.1007/s11412-007-9014-4
- Kollar, I., Fischer, F., & Hesse, F. (2006). Collaboration Scripts – A Conceptual Analysis. *Educational Psychology Review*, 18(2), 159-185. doi: 10.1007/s10648-006-9007-2
- Kollar, I., Fischer, F., & Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learning and Instruction*, 17(6), 708-721.
- Levin, D. M., Hammer, D. & Coffey, J. E. (2009). Novice Teachers' Attention to Student Thinking. *Journal of Teacher Education* 60(2): 142-154.
- Linn, M. C. (2006). The knowledge integration perspective on learning and instruction. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 243-264). New York: Cambridge University Press.
- Lu, J., Lajoie, S., & Wiseman, J. (2010). Scaffolding problem-based learning with CSCL tools. *International Journal of Computer-Supported Collaborative Learning*, 5(3), 283-298. doi: 10.1007/s11412-010-9092-6
- Mäkitalo-Siegl, K., Kohnle, C., & Fischer, F. (2011). Computer-supported collaborative inquiry learning and classroom scripts: Effects on help-seeking processes and learning outcomes. *Learning and Instruction*, 17(6), 257-266.
- Nelson-Le Gall, S. (1981). Help-seeking: an understudied problem-solving skill in children. *Developmental Review*, 1, 224-246.
- Newman, R. S. (1994). Adaptive help seeking: A strategy of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 283-301). Hillsdale, NJ: Erlbaum.
- Oortwijn, M. B., Boekaerts, M., Vedder, P., & Strijbos, J. W. (2008). Helping behaviour during cooperative learning and learning gains: the role of the teacher and of pupils' prior knowledge and ethnic background. *Learning and Instruction*, 18(2), 146-159.
- Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D.C., Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, 13(3), 337-386.
- Schön, D. (1983). *The Reflective Practitioner. How professionals think in action*. London: Temple Smith.
- Tabak, I. (2004). Synergy: A Complement to Emerging Patterns of Distributed Scaffolding. *Journal of the Learning Sciences*, 13(3), 305-355.
- Yin, R. K. (2003). *Case study research: design and methods* (3rd ed.) Thousand Oaks, CA: Sage Publications Inc.

Acknowledgments

This study was funded by the “Science in Society” Initiative of the Seventh Framework Research Programme (FP7) of the European Community, under the PROFILES grant (266589). Opinions, findings, and conclusions are those of the authors and do not necessarily reflect the views of the funding agency. We would like to acknowledge our appreciation of the teacher and students who participated in this study.

Using Gartner’s Hype Cycle as a basis to analyze research on the educational use of ubiquitous computing

Jari Laru and Sanna Järvelä

Faculty of Education, University of Oulun yliopisto, PO Box 2000, 90014, Finland

Jari.laru@oulu.fi, sanna.jarvela@oulu.fi

Abstract: In this paper, Gartner Group’s Hype Cycle is used as the basis for categorizing and analyzing research on the educational use of ubiquitous computing. There are five stages of the Hype Cycle: technology trigger, peak of inflated expectations, trough of disillusionment, slope of enlightenment, and plateau of productivity. The first decade of research on the educational use of mobile technology is divided in this paper into four stages: (i) a period of mobility and personal digital assistants; (ii) the era of wireless Internet learning devices; (iii) the introduction of social mobile media; and (iv) a ubiquitous future. In addition, three empirical case studies are used as examples of these developmental stages. These case studies demonstrate the diversity of contexts, methods, and technologies used, ranging from workplace to nature trail, from inquiry learning to collaborative knowledge building, and from PocketPCs to smartphones.

Introduction

The evolving role of smartphones, Internet tablets, and other mobile devices in our everyday life is an example of *ubiquitous computing*, a term coined by Weiser (1991), who wrote that “the most profound technologies are those that disappear [because t]hey weave themselves into the fabric of everyday life until they are indistinguishable from it” (p. 94). Weiser is widely considered to be the father of ubiquitous computing, an environment in which the computer is integral to but embedded in the background of daily life.

In this paper, Gartner Group’s Hype Cycle is used as the basis for categorizing and analyzing research on the educational use of ubiquitous computing because the Hype Cycle characterizes the typical progression of an emerging technology. As depicted (Figure 1), there are five stages of the Hype Cycle: technology trigger, peak of inflated expectations, trough of disillusionment, slope of enlightenment, and plateau of productivity. Because the technology is at different levels of development during each of the five portions of the cycle, research into the educational use of the technologies can be made in steps. These steps are not linear in the strictest sense; rather, they follow the steps of the development of research in the field. In this paper, the Hype Cycle is also used to structure an examination of the development of the general idea of mobile computer-supported learning. This is achieved by adding a layer of several megatrends in the technology-enhanced learning field on top of the Hype Cycle (shown at the top of Figure 1) (O’leary, 2008).

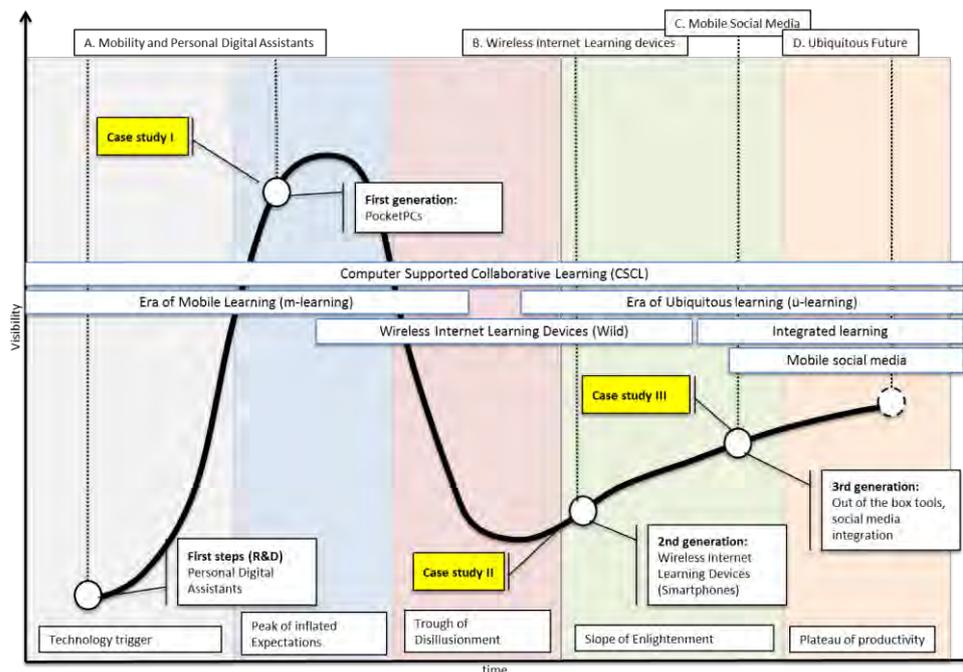


Figure 1. Gartner's Hype Cycle and educational use of ubiquitous computing

In addition to analysis of the general idea of mobile devices in education, three empirical case studies are included in this paper as examples of the developmental stages. These case studies demonstrate the diversity of contexts, methods, and technologies used, ranging from workplace to nature trail, from inquiry learning to collaborative knowledge building, and from PocketPCs to smartphones. The methodological and technological choices made in these case studies have enabled examination of the nature of social interaction in small-group settings supported with ubiquitous technologies in three different contexts with different needs for data collection and analysis.

First years of research on the educational use of ubiquitous computing: Mobility and PDA(s)

The history of the educational use of ubiquitous computing begins with the era of technology triggers (stage 1), including the product launches by Apple Newton, Palm Pilot, and Nokia Communicator in the late 1990s, followed by Microsoft's PocketPC in the early 2000s. Later devices are considered first-generation gadgets in this cycle. These early developments in ubiquitous communication led to a peak of inflated expectations (stage 2) when some scholars thought that mobile devices would revolutionize education (Trifonova, 2003). It was typical during this period to refer to the educational use of mobile devices under the terms "mobile learning" and "m-learning" (Keegan, 2005; Park, 2011; Quinn, 2000).

The idea of *mobile learning* was presented by Sharples (2000), who said that new technological affordances enabled a "new genre of educational technology—personal (handheld or wearable) computer systems that support learning from any location throughout a lifetime." The various educational affordances of wireless technologies suggested by researchers thus far (Roschelle & Pea, 2002) have paved the way for the emergence of so-called mobile learning or ubiquitous learning initiatives, such as G1:1 learning (Chan et al., 2006). While some researchers elaborate terms deeply in scientific practices, many still understand mobile devices and wireless networking technologies in education as "an extension of e-learning" (Quinn, 2000) or the mainstream, pervasive learning delivery medium. However, these simplistic views ignore the fact that modern education and pedagogy puts a high value on active, productive, creative, and collaborative learning methods that go far beyond the absorption of codified information (Hoppe, Joiner, Milrad, & Sharples, 2003).

Case study 1: Designing a new virtual master's program in the context of a distance education network

This study was conducted in realistic settings with the University Learning Center, which offers distance education on information processing sciences through several retraining programs in seven independent regional learning centers. The voluntary participants (N = 10) were split into three teams at two different locations in a northern area of Finland. The participants (nine men and one woman) comprised four project managers, a lecturer, a computer specialist, an educational designer, and three new media designers. All participants had previous experience in working together in the same distributed organization.

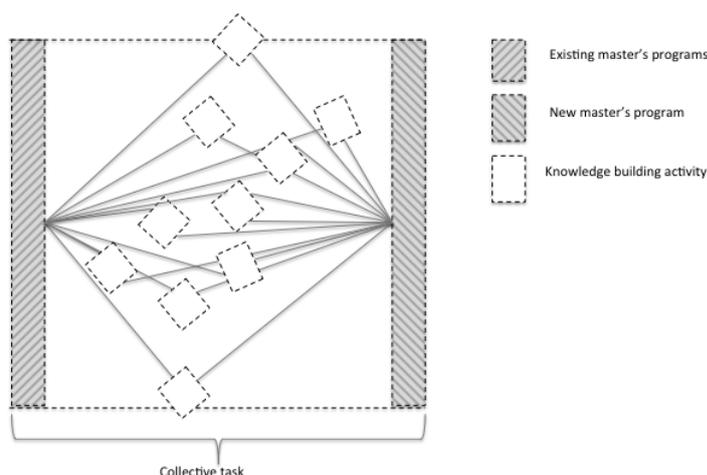


Figure 2. Instructional design of the first case study

In this case study (Laru & Järvelä, 2008), the participants shared a major problem, which was to design a new distance education master's program in a new domain (see Figure 2). The instructional design in this first case study was simplified: a knowledge-building tool was just embedded into existing practices. To design the program, participants were offered a mobilized version of collaborative technology (FLE3mobile) with a

dialogue model of knowledge building at their disposal. Ideas of knowledge building and progressive inquiry learning were operationalized in sentence openers. At the pedagogical level, participants were free to collaborate as they desired while designing the program.

Appearance of first wireless Internet learning devices together with pedagogically ambitious learning goals

In the third stage, that of disillusionment, critical accounts toward technology determinism started to appear. A considerable amount of research effort in this decade was driven by technological challenges, and few studies dealt with questions of how meaningful and productive mobile technology-supported (collaborative) learning actually is (Järvelä, Näykki, Laru, & Luokkanen, 2007; Park, 2011). These concerns are explicitly enumerated in an extensive review of mobile learning projects by Frohberget et al. (2009), where the authors argue that “tool support of most projects is not pedagogically ambitious, [and] a strong minority provide tools that aim at realizing higher pedagogical goals” (p. 317).

In order to ensure engaged learners, a proper pedagogical or lesson design is needed for when enthusiasm for using the new technologies begins to wear out (Looi et al., 2009). Yet, although many scholars, most notably Roschelle and Pea (2002), have predicted tensions between traditional learning models, which are highly centralized, and emerging pedagogical ideas amplified with mobile technologies, which are naturally situated, collaborative, and distributed, educational technologists tend to create applications that are designed to work within “inherited educational ideas rather than transform them” (Squire & Dikkers, 2012). Roschelle and Pea (2002) also predicted how mobile technology might revolutionize the role of teachers by breaking contrastive teaching paradigms of “sage on the stage” (teacher-centered instruction) and “guide by side” (teacher-guided discovery). Instead, they offered the idea of “conductor of performances,” which has been further developed by other scholars (Dillenbourg & Jermann, 2010) using the term “orchestration” to describe run-time adjustments in complex socio-technical designs that include multiple social planes in different contexts mediated by multiple devices.

Case Study II: Field trip to a nature park in a wilderness forest setting in the context of informal K-12 education

The participants in the second case study were primary school students (N = 22, all 12 years of age) who participated in a one-day learning project during a field trip to a nature park in a wilderness forest setting in northern Finland. The field trip activities in this case study were designed and developed by the research team in collaboration with the nature park’s local expert, a biologist. The students were randomly assigned to eight groups (six triads and two dyads), and each group was provided with a mobile phone. Before the experiment, the principles and procedures of collaborative inquiry learning and argumentation were presented, and practical training for the fieldtrip was given in the classroom by the researchers and the biologist.

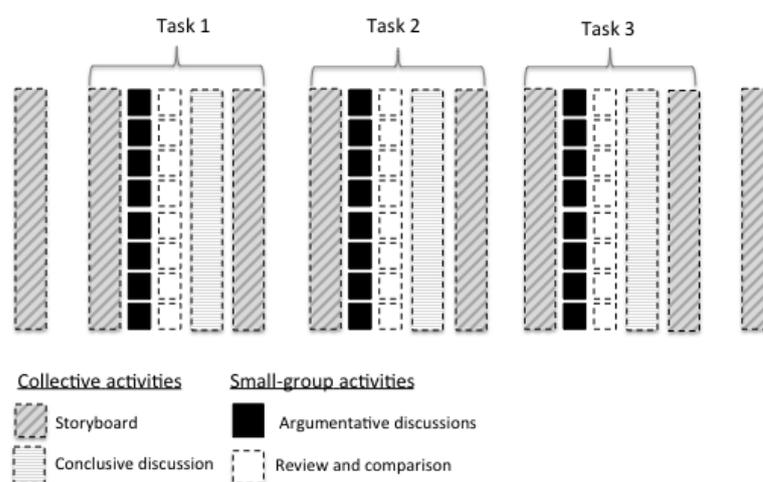


Figure 3. Instructional design of the second case study

The major problem in this study was to explore inanimate and animate traces of nature in small groups in order to create argumentative knowledge claim messages (Laru, Järvelä, & Clariana, 2012). This study is an example of a teacher-led outdoor learning activity in which students learn in groups within confined time periods, which is a subtype of “formal learning in informal settings” (Wong & Looi, 2011).

From the perspective of instructional design (see Figure 3), a collaborative core activity was aimed at scaffolding co-construction of argumentative discussions in small groups during inquiry learning. It consisted of “soft” scaffolding, provided by tutors and the nature guide, and “hard” argumentation scaffolds, provided by the messaging tool (sentence openers). In addition, the instructional design included pre-structuring activities that provided procedural scaffolding in the form of storyboard messages (Laru et al., 2012) as well as post-structuring activities that included debriefing activities such as a review and comparison phase in the collaborative and conclusive synthesis at the end of each task at the collective level.

Era of social mobile learning: Combining affordances of social software and mobile learning

The developments described for the previous phase, together with new affordances of mobile technologies, led to the Hype Cycle stage of enlightenment (Figure 1). The affordances provided by the combination of mobile devices and social software tools led us into a new phase in the evolution of technology enhanced learning, one that forges new learning spaces and continuity between pedagogical phases of the instructional design (Laru, Näykki, & Järvelä, 2012; Multisilta & Milrad, 2009; Wong & Looi, 2011). In practice, the increasing use of mobile social media in education is stitching learners’ formal and informal learning contexts together and bridging individual and social learning, which leads to seamless learning.

However, most papers considered in the extensive literature review made by Wong and Looi (2011) tend to discuss or analyze personalized and social learning in their studies separately or to only focus on one of these aspects. Further, very few papers discuss the mechanisms of bridging individual and collaborative activities. The third case study in the current paper is focused on bridging individual and collaborative activities as well as face-to-face and mobile social media activities. It includes a full activity design, as suggested by Wong and Looi (2011), with multiple phases; the mobile-mediated conceptualization activity was just one phase of the instructional design. Products created in that phase can be characterized as artifacts that were used as a mediating tool for reflections, elaborations, reviews, and knowledge building (Wong & Looi, 2011).

Case Study III: Future scenarios and technologies in learning: A course in the context of higher education

For the third case study, the research participants were 21 undergraduate students in a five-year teacher education program at the Faculty of Education in the University of Finland. All students were enrolled in a required course entitled *Future Scenarios and Technologies in Learning* during the spring semester of 2009. The 21 participants comprised 16 women (76%) and 5 men (24%). The prevalence of women reflected the gender ratio of education majors at the university. The mobile phone-mediated activities in this course are an example of course-related activities outside of the normal class hours, such as artifact creation in daily life (largely incidental encounters or improvisations), which is another subtype of formal learning in informal settings (Wong & Looi, 2011).

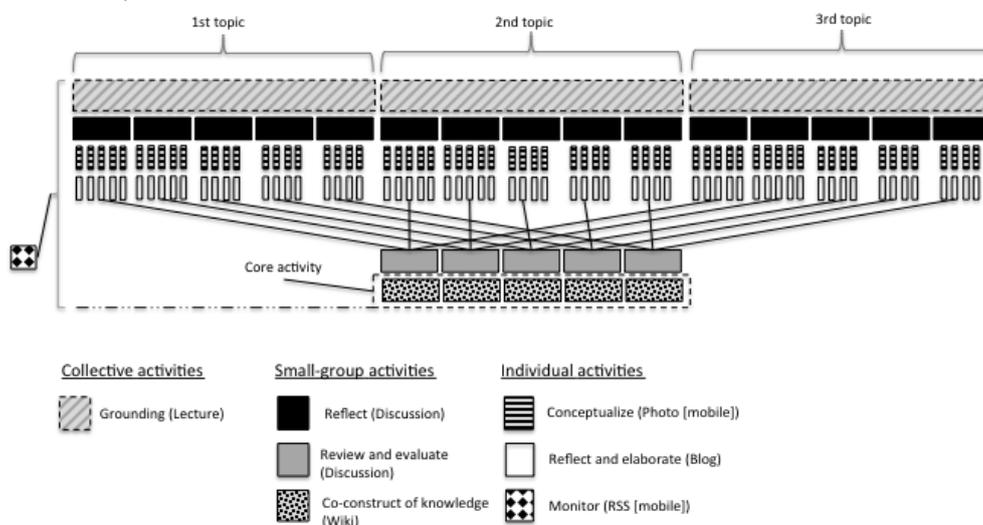


Figure 4. Instructional design of the third case study

In this case study, the same content was elaborated multiple times when students encountered multiple representations of each of the content topics (six altogether) using different analogues, examples, and metaphors. In other words, the instructional design required students to revisit “the same material, at different times, in rearranged contexts, for different purposes and from different conceptual perspectives” (Spiro,

Feltovich, Jacobson, & Coulson, 1991, p. 28). From the perspective of ill-structured problems and tasks, the students split one problem into multiple smaller problem-solving tasks as phases in the instructional design proceeded.

In this experiment, the learners' core task was to integrate selected individual blog reflections and visual representations into coherent and a comprehensive wiki (see Figure 4). Although this wiki was also the main outcome of the activity (the end goal for their activities), it was not specified as such. There were also multiple individual and collective phases before the wiki activity, and the goals for these were not specified either.

The students needed to make choices in three phases concerning their learning objectives aimed at solving ill-structured problems, thus:

1. *Reflection* (collaborative): After a grounding lecture in which students discussed the lecture topic in groups and formulated a problem to be solved during the following individual learning phases;
2. *Conceptualization* (individual): Following the reflection phase, which included an activity in which students were required to conceptualize their group members' shared interests (i.e., shared problem); this task can be considered as a standalone ill-structured task that led students to qualitative modeling in order to reformulate group-level problems;
3. *Knowledge co-construction* (collaborative): An assigned task focused on integrating each group's selected blog entries and photos into a cohesive and comprehensive group wiki; this activity could not be conducted without qualitative modeling to reformulate shared learning objectives and problems, because individual activities affected the shared objectives and problems.

The instructional design of the third experiment enabled students to make comparisons between the cases. This was done both in face-to-face activities and with the help of technological tools. The activities involving comparison comprised two phases:

1. *Reflection and elaboration* (individual): After individual conceptualization, students were required to analyze photos taken using mobile phones in order to discard ideas that were not relevant to their groups' shared learning objectives; they were also required to write blog entries on selected photos, in which they further elaborated associations between the photos, group-level objectives, and students' everyday situated practices (note: students were able to see photos taken and blog entries written by other students and in other groups by monitoring their activities using an RSS reader);
2. *Review and evaluation* (collaborative): After individual reflection and elaboration, students were tasked with reviewing group members' blogs and evaluating the usefulness of blog entries in the context of their shared learning objectives.

Ubiquitous tomorrow: Learning environment consisting of an amalgam of tools around the corner

From the present perspective, this field of research is currently in the phase of the plateau of productivity. The world is entering the Age of Mobilism (Norris & Soloway, 2011). Ubiquitous computing has evolved from Weiser's initial ideas about the interplay between the human world and communication technologies with the widespread adoption of mobile devices that require proactive involvement rather than the calm computing originally suggested by Weiser. Mobile phones have grown beyond a tool for conversations, to become connected computing devices that offer a multitude of services and which currently are perceived as much more than just phones, having also developed into movie players, gaming platforms, cameras, etc. (Pea & Maldonado, 2006). Current trends are increasingly focusing on effective personal learning environments as being characterized by an amalgam of technology devices, software, and services; access to a variety of digital tools simultaneously for everyone, anywhere, anytime; and choices about which technology is most appropriate in a given situation (van't Hooft & Swan, 2007). In many techno-centric papers on context-aware technology, previous killer features—contemporary human/computer interaction paradigms (RFID tags, QR-Codes, GPS, etc.)—are fast becoming regarded as mainstream in current mobile devices. Timely, contextualized information afforded by these can serve as evidence to support partially formed ideas and misunderstandings and to trigger comparison with previously stored data on the device, as well as to support an inquiry process or dialogue in situ. Actually, these affordances are enabling the preparation of instructional designs based on the ideas suggested a decade ago (Roschelle & Pea, 2002).

Western students today may have "one or more devices per student" if needed, but the number of devices in the ubiquitous environment is quite variable. Indeed, device-to-user ratios range from the use of

multiple computing devices (like sensors) by one student (10:1) to a class of students with one interactive whiteboard (1:all), and encompass the in-between usage scenarios of 1:1 (as G1:1 initiative members originally suggested), 1:2 (as in pair work sharing a device), and 1:4 (as in small-group work discussions mediated by a shared device) (Dillenbourg, 2010 in Wong & Looi, 2011). These device-user ratios set new challenges for instructional designers, because each ratio provides different dynamics of interaction and collaboration (Wong & Looi, 2011).

In other words, different device-student ratios are an example of converged cognitive tools that we unconsciously and effortlessly use for achieving the benefits of distributed intelligence (Pea & Maldonado, 2006). From an educational perspective, this is a part of an environment in which “all students have access to a variety of digital devices and services, including computers connected to the Internet and mobile computing devices, whenever and wherever they need them” (van’t Hooft, Swan, & Cook, 2007, p. 6). It is also line with the tenets of constructivism insofar as it involves a learning environment in which both teachers and students are active participants in the learning processes (critically analyzing information, creating new knowledge in a variety of ways, communicating what they have learnt) mediated by tools they have chosen and that are appropriate for particular tasks (Dabbagh & Kitsantas, 2011).

Discussion

Overall, decades of research in the field of educational use of ubiquitous computing and rapid technological evolution (both described in Figure 1) illustrates the rich field of business and research opportunities. Van Lente, Spitters, and Peine (2013) have argued that hypes thrive in rich environments, where research, business, and wider social activities contribute to the creation, sharing, and refinement of expectations. This paper follows studies conducted by Järvenpää and Mäkinen (2008) and van Lente et al. (2013), which have bridged empirical measures to the Hype Cycle. Our paper represents an exploratory and empirically driven study seeking indicators in the three case study designs for the Hype Cycle in relation to the evolution of educational use of ubiquitous computing.

The Hype Cycle and case studies described here emphasize that pedagogically grounded instructional design is needed in order to put emergent technologies into effective use. The employment of mobile devices, including mobile phones and tablets, is a growing trend in education. The practice has been widely technology-driven and often justified simply by the importance of using new technology in classroom. Since we are currently living between the stages of mobile social learning and ubiquitous future, the role of mobile technologies in different learning contexts is still a challenge for researchers and practitioners. Our claim is that seamless learning can be one productive way for schools and other educational institutions to promote learning skills, namely, self-regulated learning and collaboration, and to prepare people for the 21st century learning society. To advance research on self-regulated seamless learning, we propose few design guidelines for self-regulated seamless learning.

We share the constructivist belief that students should learn in environments that deal with “fuzzy,” *ill-structured problems*. Designing challenging collaborative learning tasks provides students with an opportunity for multiple strategic activities and for self-regulation and shared regulation of learning. There should not be one right way to reach a conclusion, and each solution should bring a new set of problems. These complex problems and challenging learning tasks should be embedded in authentic tasks and activities, the kinds of situations that students would face as they apply what they are learning to the real world (Needles & Knapp, 1994). Challenging learning tasks require scaffolds and support. For example, Belland (2011) has suggested the following guidelines for the creation of appropriate scaffolds: (a) Support problem reformulation through qualitative problem modeling; (b) do not give specific end goals; (c) enable students to make comparisons between cases; and (d) enable students to work collaboratively.

As suggested by Spiro et al. (1991), the same content can be *elaborated multiple times*. In practice, this means that students encounter multiple representations of content using different analogues, examples, and metaphors, for example, by using mobile tools or social software. The instructional design required then is for students to revisit the same material, at different times, in rearranged contexts, for different purposes and from different conceptual perspectives. The same content can be also elaborated with multiple individual and collaborative phases before the collective activity allowing students opportunities for self-, co-, and shared regulatory processes (Järvelä & Hadwin, 2013).

References

- Belland, B. (2011). Distributed cognition as a lens to understand the effects of scaffolds: The role of transfer of responsibility. *Educational Psychology Review*, 23(4), 577–600. Retrieved from <http://dx.doi.org/10.1007/s10648-011-9176-5>
- Chan, T.-W., Roschelle, J., Hsi, S., Kinshuk, K., Brown, T., Brown, T., Patton, C., et al. (2006). One-to-one technology-enhanced learning: an opportunity for global research collaboration. *Research and Practice in Technology Enhanced Learning*, 1(1), 3–29. doi:10.1142/S1793206806000032

- Dabbagh, N., & Kitsantas, A. (2011). Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *The Internet and Higher Education*, 15(1), 3–8. doi:10.1016/j.iheduc.2011.06.002
- Dillenbourg, P., & Jermann, P. (2010). Technology for classroom orchestration. In M. S. Khine & I. M. Saleh (Eds.), *New Science of Learning Cognition Computers and Collaboration in Education* (pp. 1–20). Springer New York. doi:10.1007/978-1-4419-5716-0
- Frohberg, D., Göth, C., & Schwabe, G. (2009). Mobile Learning projects - a critical analysis of the state of the art. *Journal of Computer Assisted Learning*, 25(4), 307–331. doi:10.1111/j.1365-2729.2009.00315.x
- Futurelab, N., Naismith, L., Lonsdale, P., Vavoula, G., & Sharples, M. (2005). *Literature Review in Mobile Technologies and Learning Literature Review in Mobile Technologies and Learning*. *Educational Technology*.
- Hoppe, U., Joiner, R., Milrad, M., & Sharples, M. (2003). Guest editorial: Wireless and mobile technologies in education. *Journal of Computer Assisted Learning*, 19, 255–259.
- Järvelä, S., & Hadwin, A. F. (2013). New frontiers: regulating learning in CSSL. *Educational Psychologist*, 48(1), 25–39. doi:10.1080/00461520.2012.748006
- Järvelä, S., Näykki, P., Laru, J., & Luokkanen, T. (2007). Structuring and Regulating Collaborative Learning in Higher Education with Wireless Networks and Mobile Tools. *International Journal of Engineering*, 10(4), 71–79. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Structuring+and+Regulating+Collaborative+Learning+in+Higher+Education+with+Wireless+Networks+and+Mobile+Tools#0>
- Järvenpää, H., & Mäkinen, S. J. (2008). Empirically detecting the Hype Cycle with the life cycle indicators: An exploratory analysis of three technologies. *2008 IEEE International Conference on Industrial Engineering and Engineering Management* (pp. 12–16). IEEE. doi:10.1109/IEEM.2008.4737823
- Keegan, D. (2005). The incorporation of mobile learning into mainstream education and training. (H. Van Der Merwe & T. Brown, Eds.) *World Conference on Mobile Learning Cape Town, 2008* (September 17), 1–17. Retrieved from <http://mlearning.noekaleidoscope.org/public/mlearn2005/www.mlearn.org.za/CD/papers/keegan1.pdf>
- Laru, J., & Järvelä, S. (2008). Social patterns in mobile technology mediated collaboration among members of the professional distance education community. *Educational Media International*, 45(1), 17–32. doi:10.1080/09523980701847131
- Laru, J., Järvelä, S., & Clariana, R. B. (2012). Supporting collaborative inquiry during a biology field trip with mobile peer-to-peer tools for learning: a case study with K-12 learners. *Interactive Learning Environments*, 20(2), 103–117. doi:10.1080/10494821003771350
- Laru, J., Näykki, P., & Järvelä, S. (2012). Supporting small-group learning using multiple Web 2.0 tools: A case study in the higher education context. *The Internet and Higher Education*, 15(1), 29–38. doi:10.1016/j.iheduc.2011.08.004
- Looi, C.-K., Wong, L.-H., So, H.-J., Seow, P., Toh, Y., Chen, W., Zhang, B., et al. (2009). Anatomy of a mobilized lesson: Learning my way. *Computers & Education*, 53(4), 1120–1132. doi:10.1016/j.compedu.2009.05.021
- Multisilta, J., & Milrad, M. (2009). Sharing Experiences with Social Mobile Media. *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services MobileHCI 09*, (10096213363491501752related:uHZ244_0HlwJ), 1. doi:10.1145/1613858.1613977
- Needles, M. C., & Knapp, M. S. (1994). Teaching writing to children who are underserved. *Journal of Educational Psychology*, 86(3), 339–349. doi:10.1037//0022-0663.86.3.339
- Norris, C. A., & Soloway, E. (2011). Learning and Schooling in the Age of Mobilism. *Educational Technology*, 51(6), 3–10. Retrieved from <http://ezproxy.usq.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=68945513&site=ehost-live>
- O'leary, D. (2008). Gartner's hype cycle and information system research issues. *International Journal of Accounting Information Systems*, 9(4), 240–252. doi:10.1016/j.accinf.2008.09.001
- Park, Y. (2011). A pedagogical framework for mobile learning: Categorizing educational applications of mobile technologies into four types. *International Review of Research in Open and Distance Learning*, 12(2), 78–102. doi:10.3394/0380-1330(2006)32
- Pea, R., & Maldonado, H. (2006). WILD for Learning: Interacting through new computing devices anytime, anywhere. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 427–441). Cambridge University Press.
- Quinn, C. (2000). mLearning: Mobile, Wireless, In-Your-Pocket Learning. *LiNE Zine*. Retrieved from <http://www.linezine.com/2.1/features/cqmmwiyp.htm>

- Roschelle, J., & Pea, R. (2002). A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. *International Journal of Cognition and Technology*, 1(1), 145–168. doi:10.1075/ijct.1.1.09ros
- Sharples, M. (2000). The design of personal mobile technologies for lifelong learning. *Computers & Education*, 34(3-4), 177–193. doi:10.1016/S0360-1315(99)00044-5
- Spiro, R., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Technology*, 31(5), 24–33. Retrieved from http://74.125.155.132/scholar?q=cache:At4-p7m5PiwJ:scholar.google.com/&hl=en&as_sdt=0,23
- Squire, K., & Dikkers, S. M. (2012). Amplifications of learning: Use of mobile media devices among youth. *Convergence: The International Journal of Research into New Media Technologies*, 1354856511429646–. doi:10.1177/1354856511429646
- Trifonova, A. (2003). Mobile learning-review of the literature. *Technology Report No DIT03009 University of Trento Department of Information and Communication Technology*. Retrieved from <http://eprints.biblio.unitn.it/archive/00000359/01/009.pdf>
- Van't Hooft, M., & Swan, K. (2007). *Ubiquitous Computing in Education - invisible Technology, Visible Impact*. London: Lawrence Erlbaum Associates, Inc.
- Van't Hooft, M., Swan, K., & Cook, D. (2007). What is Ubiquitous Computing. In M. Van T Hooft & K. Swan (Eds.), *Ubiquitous Computing in Education* (pp. 3–17). Lawrence Erlbaum Associates, Inc.
- Van Lente, H., Spitters, C., & Peine, A. (2013). Comparing technological hype cycles: Towards a theory. *Technological Forecasting and Social Change*, null(null). doi:10.1016/j.techfore.2012.12.004
- Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 3(3), 3–11. doi:10.1145/329124.329126
- Wong, L.-H., & Looi, C.-K. (2011). What seems do we remove in mobile assisted seamless learning? A critical review of the literature. *Computers & Education*, 57(4), 2364–2381. doi:10.1016/j.compedu.2011.06.007

Repurposing everyday technologies for math and science inquiry

Sarah Lewis, Wendy Ju, Stanford University, 353 Serra Mall, Stanford CA 94305
sarahl@stanford.edu, wendyju@stanford.edu

Abstract: Students often have far more sophisticated scientific instruments in their pockets than in their physics classrooms. Today's cell phones and game controllers offer sensors, cameras and communication technologies that can be used for in-depth exploration of physical phenomena. Because everyday toys and tools offer connections to children's social worlds, they are particularly useful for integrating classroom science with everyday intuitions and experiences. Drawing on data from a multi-year research project to help children hack gaming technologies for science inquiry, we examine both technological and social advantages that repurposing everyday technologies for offers for learning abstract STEM concepts. In light of trends towards increased decentralization of education, we extend these findings into a general discussion of the potential for embedding CSCL into the design of everyday things.

Introduction

In many communities kids live and grow in social worlds embedded with interactive technologies of a potency rarely found in classrooms. Sensors in toys, computers, and phones capture changes in position, direction, acceleration, location, and proximity; data streams seamlessly between devices; screens, lights, and audio represent this data locally or remotely. Together, these technologies enable more of us to spend increasing amounts of time learning outside of schools. However, they also offer deep, yet often hidden, affordances useful for classrooms. Even moderately-priced cell phones, game controllers, stuffed toys, watches, handheld computers and workout equipment contain sensing technologies—such as accelerometers and infrared cameras—that could be tapped to reveal the science and mathematics that underlie and describe the workings of the physical world. Because these objects are in everyday use, they also hold the potential to bring difficult concepts in math and science into children's day-to-day social contexts.

A handheld game controller is one such an inexpensive everyday device with components that lend themselves readily to math and science inquiry. The Wii Playstation Remote, which features a three-axis accelerometer, a gyroscope, an infrared camera, seven buttons, a speaker, a haptic motor and two-way wireless Bluetooth communication, cost less than \$20 USD at the time of writing. Affordable, powerful and hackable, they are also pervasive in many children's social worlds. (Nintendo's survey data indicates 46% of Americans aged 6 to 74 played a Wii or Nintendo DS in 2010 (Iwata, 2010).) Components in these devices measure physical phenomena related to motion, such as distance, rotation, velocity and acceleration, common topics of study in science classrooms. Given its ubiquity, the Wii Remote and devices like it could be as familiar and as ready at hand for physics projects as a desk ruler.

Numerous teachers, researchers, hackers and DIY enthusiasts have written about hacking game controllers for learning in various contexts (Williams & Rosner, 2010; Lee, 2008; Hill, 2009; Pearson & Bailey, 2007; De Bruyn, 2008; Graves et al., 2007). Perhaps closest to our work, researchers focusing on high school and college-level physics curriculum (e.g., Vannoni & Straulino, 2009; Somers, et al., 2009; Wheeler, 2011) used game controllers to collect data related to phenomena such as the motion of a pendulum, simple harmonic motion in a spring, and linear displacement on a track. Inspired by this work and interested in bridging the gap between 'hands on' project based learning and abstract concepts presented in lectures in middle school science classrooms, we launched a multiyear research project to engage children in hacking game controllers to collect and visualize data related to their design projects (described in detail in Lewis, Acholonu & Ju, 2012). We anticipated that children would glean data from their projects and—perhaps more importantly—that they also would gain a better sense of how physics and math relate to the technologies in their everyday lives.

Throughout the course of the project, we noted consistent talk anchored around the device that traversed boundaries between children's 'social worlds' of gaming and their classroom 'science worlds.' In this paper we briefly provide a description of activities with students, and then present snippets of classroom interaction that exemplify the overall *technological* and *social* affordances that contribute to the blending of these worlds. We conclude with larger questions for CSCL about how the commonplace "smart" devices of childhood might contribute to math and science learning across contexts. Currently everyday technologies are designed to be easy to use, their very form factor inviting the user how to learn to manipulate them to accomplish a goal (Norman, 2002). We suggest that, with more people learning in distributed, informal contexts, the technologies around us, such as cell phones, depth cameras, GPS technologies, and other sensor-driven personal devices, might be altered to lead people toward learning not only *how to* accomplish something, but also toward more fundamental principles about *why* the world and its technology works as it does.

Project Overview

The interactions presented below are from a three-year design experiment (Brown, 1992) to develop software and curricular tools to support the use of game controllers and mobile phones as inexpensive data acquisition tools for middle-school-level physics activities. The design research (Edelson, 2002) goals of this project are twofold: 1) to design activities and software to help students and teachers harness everyday technologies to support scientific inquiry in science classroom labs and projects and, more broadly, 2) to develop activities and software tools that encourage young people to repurpose these same technologies for their own interests and pursuits in and out of school. In the spirit of participatory design, the project was carried out in close collaboration with the faculty and students of our partner school, with the participants considered co-designers of the overall program. Researchers programmed the software and ran classroom activities, soliciting suggestions and feedback for both. We refined designs in response to faculty and student input, as well as our own observations of our program's impact on classroom dynamics and learning. This cooperative design approach is similar to participatory and collaborative approaches outlined by Inkpen (1999), Druin (1999), and others working in the realm of classroom technology research and design (Rode 2003).

Setting and participants

All activities took place in a 'constructivist' sixth-grade physics class at a private school for academically talented students. The physics teacher was experienced at leading project-based activities, and alternated between open-ended design projects, structured labs, and direct instruction, usually presenting abstract concepts, algebraic formulas and graphical representations of data via interactive lectures only after related hands-on projects. This site was chosen because of its instructional philosophy, because of its commitment to innovation, and because the school issues and supports Apple laptops to all middle school students, enabling teams to easily use the software we provided.

One hundred and eighteen sixth-grade boys and girls, in three classes per year over two years, worked with us on this project. Overall, students drew upon high levels of technological experience, with several bringing engineering and programming skills from robotics clubs or mobile app programming classes. Surveys indicated most had extensive experience with a variety of computational devices, with all having access to or familiarity with computer and video game technologies at home.

Activities

Four hands-on activities with gaming technology were conducted during the course of each school year, all of which were video recorded. After each activity as many teams as possible were interviewed using "artifact based" techniques (Barron 2002). Students used two emerging interfaces that either harnessed data from the force sensor, showing graphs of changing acceleration over time, or tracked the position and duration of interruption of infrared lights aligned with the game controllers IR camera.

To engage students in a common experience and to better understand their prior knowledge, the project started with a warm up activity that involved playing the "Wii Tennis" game in small groups, and followed by group discussion on how they thought it worked. This was followed by a three-week "mousetrap car" design activity (Figure 1) a fun project often conducted in middle school science classrooms. In this activity students are tasked with designing and building a car powered by a mousetrap. They tinker with materials and their naive understandings of the physics of motion to develop an efficient car with maximum acceleration that will travel as far as possible under its own power. In the course of the project, students encounter concepts related to forces, such as friction, mass, velocity and acceleration. Students strapped game controllers onto their cars to visualize changes in acceleration over time and relate those to design factors. Some students disassembled the game controllers to make them lighter. The classroom teacher followed up this activity with students presenting and discussing their findings before offering several direct instruction sessions on Newton's laws of motion.



Figure 1. Students mousetrap car with WiiMote suspended from chassis.

Noting conceptual issues related to "negative acceleration" in students' explanations of their mousetrap cars, researchers next presented students with a "punch" activity. Students were asked to predict the "shape" of

the acceleration graph of a single, extended punch and test their prediction over multiple trials. Students drew their predictions on paper, and then, in groups of two, recorded their trials using camera phones. Holding a string tied to a vertically suspended game controller (Figure 2), students punched and examined the line graph of change in acceleration over time. Results were reported out, which led to a group discussion of ‘negative acceleration.’

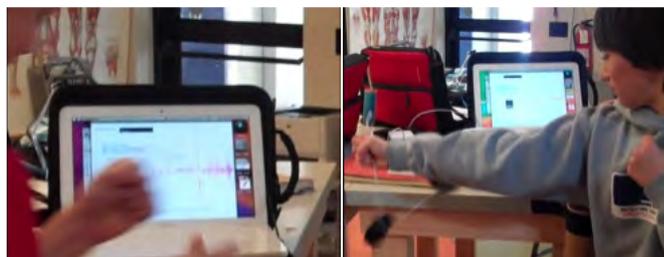


Figure 2. Students punch and watch the screen to explore negative acceleration.

The school year concluded with a marble rollercoaster activity, also a common project in middle school science classrooms. Studying kinetic and potential energy, students built a marble rollercoaster (Figure 3) within specified constraints. Using their knowledge of the workings of the gaming system, and infrared LEDs and IR cameras from game controllers, they developed systems to track the marbles velocity at different points along the track.



Figure 3. Students aligning IR light with Wii IR camera on their rollercoaster track.

Findings

We found it notable that some of the most productive discussions involved students not simply analyzing their team’s data, but talking more broadly in conversations that forged a three-way link between the principles of physics under study, the technical functioning of the game controllers components, and the social context of gaming. The illustrative snippets below were taken from video of classroom activities and post activity interviews across two years of the project. As a whole, they suggest that the gaming origins of the technology made a difference in how students participated in and talked about what they were doing in the classroom. This influence is subtle and likely would have gone unnoticed except that it showed up on tape repeatedly across classes, and often during gaps between more formally organized activities – such as during set up or testing of technology as students were preparing to use the remotes in their projects. Classroom video was revisited and coded for general reference to physics and the technology. Excerpts that contained direct references to *gaming* were further coded, revealing six general ways repurposing helped link children’s social and scientific worlds.

Linking technological affordances to core concepts in science

The tools used in these activities are ultimately useful because of the inherent deep connections between the core functions of game controller components, the action on screen, and core curricular concepts. For example, the force sensors in handheld game controllers such as the Wii Remote measure acceleration across three axes, and take into account the force of gravity. Core mathematics skills involving graphing, algebraic functions, and the mathematics of change are deeply connected to the data stream the game controllers emit, as are core concepts in physics such as position, velocity, and positive and negative acceleration.

Most importantly, this data stream is accessible and flexible. It can be harvested and interpreted according to students’ learning contexts, and readily pairs with their own laptops or other Bluetooth-enabled devices. While commercially available demonstration cars and other pre-formulated technologies display data related to force and acceleration, this technology is expensive (at the time of writing, approximately \$400 per

car) and its display output is limited. Data cannot be saved, aggregated or easily shared. By harvesting data from the game controllers, we are not only able to use custom data visualizations to directly address issues students were struggling with, but also take advantage of hidden ‘hooks’ into the curriculum. For example, gravity is a difficult concept in part because unless something is falling, students don’t ‘see’ it in action. Even after studying gravitational force, students often base their explanations on their everyday intuitions that gravity is something that ‘happens’ when you drop something. The force sensor in the game controller measures gravity along the Z-axis, of course whether or not the device is moving. The following vignette illustrates how this offers an accidental learning opportunity. In the interaction below, one student, while pairing the Wii remote with his computer for the mousetrap car activity, noticed a line marking “1” along the Z axis:

Student 1 (assuming he had paired with another team’s Wii): “But I’m not moving it. Why does it say one? Wait, stop shaking the table. Maybe that is someone else’s?”

Student 2 (grabbing the Wii and shaking it): “No. Look!”

Student 1: “It’s stuck... no...”

Teacher (passing by): “You aren’t touching it, but are any forces acting on it?”

Student 2: (after some time playing to figure out what the Z axis represented): “Gravity?”

Student 2: “No, but it’s not falling ...”

Teacher (swinging the Wii like a tennis racquet): “Ok. The force sensor measures the force you move it with, so it can know how hard you hit it but also any other forces acting on it.

Is gravity acting on it right now? ... What else is acting on it? Why isn’t it falling?...

This interaction led to a rich discussion of the force of the table exerted on the device, the general concept of the ‘normal force,’ and to how the Wii tennis game takes gravity into account. Exploration with the tool and a mathematical representation of its data stream led students towards exploration of core ideas of physics and the mathematics of change.

Linking social worlds to science worlds

Certainly other technologies, such as sensors attached to Arduino boards or photogate systems described in physics education catalogs are also available and similarly useful for harvesting and representing data. Besides the convenient form factor and cost savings, what is the advantage of using everyday familiar technologies such as game controllers in classrooms? While visualizations of the output of the technology’s components led students toward discussion related to math and science, the social gaming context of the device led to students forging connections across contexts that we think led to further encounters with these ideas in action in their everyday lives. The choice to repurpose a well-known game controller as the device to reveal data related to force, acceleration, and velocity mattered to students’ interactions. We have observed that while the activities themselves presented students with data, graphs, and ‘discrepant events’ (Nussbaum & Novick, 1982) that helped them confront and develop their lay opinions and intuitions about physics, often *the way* in which they came to interpret these events was by synthesizing their new observations with their prior experiences of using the technology for gaming. This cross-contextual synthesis appears to have played a role in students’ everyday thinking with scientific principles. The following general categories emerged as relevant to the value of the ‘everydayness’ of these tools for math and science learning.

Expertise is connected to ‘felt’ experiences and intuitions

Students’ experiences with game controllers are for the most part felt experiences; they learn to swing an on-screen tennis racquet by feeling the relationship between the motion of the controller in their hand and the visual feedback they receive onscreen. Through felt experiences they develop intuitions about the way the game works as well as the way the Wii Remote works. For example, in Wii gaming, expert players often make only small, quick movements with the device, rather than the broad sweeping strokes of a novice swinging an on screen racquet. While experienced players have a felt sense of the difference, they often don’t “know” why their techniques work. This presents rich fodder for discussions that link everyday intuitions and experiences with core principles:

Student: “No, it’s like (flicking the Wii remote and pointing to the screen that shows a graph of change in acceleration over time)... it’s not how far. It’s like how much it’s changing ... see... (moving hand steadily at a constant rate, producing a relatively flat line). See... cause it’s flat... cause it’s not changing it’s just moving steady. Like when you swing... you know you can just go... (flicks it again, causing a sharp peak on screen) ... ‘cause it’s not how big the swing is, it’s just like, how hard you do it, how quick (flicking several times). Like, I mean in tennis, if you know what you’re doing you just flick it to hit it hard.”

Ready at hand, easy to use, and draws on familiar metaphors

Gaming technologies have been designed to be ready at hand for students' play. They are easy to set up, easy to use, easy to remember. Students feel they "know" them, and are not concerned with breaking them or worried about learning anything new to set them up. Because of their prior experiences, in general students had an orientation towards efficient set up and troubleshooting. Generally, their familiarity with the metaphors of 'pairing' and 'sensing' provided anchors for getting everyone on board. However, some metaphors provided fodder for rich discussion when they broke down. Although facilitating the practical set up of a game console, some terms, like 'sensor bar,' are misleading when it comes to science and engineering. Discussion of the 'sensor bar' provided an opportunity to analyze not only how the game worked, but the forces at play in its marketing:

Instructor: "So, what is this?"

Students: "A sensor bar!"

Instructor: "Sensor? What does it sense?"

Student 1: "LEDs."

Student 2: "Like, the position of the Wii thing."

Instructor: "So, where is the LED?"

Student 3 (eventually, after some disagreement): "In the sensor bar!"

Instructor: "Ok, in the 'sensor bar.' So where is the sensor?"

Student 1: "In the Wii!"

Instructor: "So, there is a camera in the Wii remote, and infra red LEDs in the 'sensor bar.'

Why do you think Nintendo calls it a sensor bar?"

While generally the culture of gaming provided analogies for students to draw on in their learning of physics, in this case, terms used in the culture of Wii gaming obfuscated the functionality of the device. Metaphors, as cultural tools, needed to be explicitly redefined in order for students to understand how the game and the photo gate system worked. The students agreed to use the term "LED bar" in class rather than sensor bar. Several students reported telling their friends outside of school not to call it a sensor bar as well, thus bringing ideas from class back into a gaming context.

Socially situated and meaningful to students' interests

Handheld game controllers are 'social' tools that students use in groups. They are closely connected to the interests and concerns of students, and students spend a great deal of time learning the values and culture of game play from each other. This gives them a positive social valence and connects 'science' to students' social worlds outside of school. Availability in out-of-school times and contexts helps blur the boundaries between learning and play; students can hack their game controllers at home as well as at school. For those with strong interests in gaming (or hacking), this can lead to discussions of the math or science of game controllers with parents and friends in multiple contexts.

While most gaming is social, 'embodied' gaming via handheld game controllers is perhaps more social than most. Students move dynamically in teams, often in direct or peripheral physical contact with each other. They value being able to play well, and keep track of details of set up, scoring points, and techniques that show expertise. In classroom conversations we found this sociality bled into physics talk – students frequently referenced or mimed game play in the classroom while talking about science. In addition, students reported talking about science with siblings, parents and friends while setting up or playing the game at home. This blurring of boundaries between students' social and scientific worlds we think is likely productive for future learning. If knowing the science behind gaming devices becomes part of their gaming cultures, students may be more likely to integrate science into their everyday worlds and future plans. Examples of this blurring of boundaries of time, place and social context show up repeatedly in video recordings of students' talk in class. Two examples:

Student: "Hey, maybe we could all come over to my house. I have a big TV and I live just over there. We could bring science class over and we could like, do Wii gaming and stuff at home after school. 'Cause we'd have more time and it would be more fun to, you know, do it outside of here."

Researcher: "That's an interesting idea. We could do some things after school here if you think kids would like that."

Student: "Yeah, but at home we have snacks and my other friends could come. It would be more fun."

Student: (smiling) "Hey, you guys kind of ruined the game for me. ..."

Researcher: “Why?”

Student: “Because now every time I swing, its like I don’t really see the racquet move, I see acceleration in my mind. Like, I swing to hit the ball, but in my head a graph shows up. It’s really distracting. And, like, my friends don’t really want to hear all about acceleration when they’re playing.”

Researcher: “Do you talk about it with your friends?”

Student: “Well, I showed them on the laptop what we were doing and we played... They [thought it was] cool, like, that I could connect it to my laptop.”

Event based and therefore evocative of storytelling

Things ‘happen’ in games. Games have beginnings, middles, ends, as well as heroes and heroines. Experiences of gaming are memorable and get repeated in stories of great achievements and defeats. Students draw on these stories and memories of past experiences when encountering the technology in new contexts, providing rich fodder as supporting analogies for learning.

In the classroom, students told stories about their gaming achievements and adventures, occasionally reinterpreting them using concepts they are learning in science class. These stories served as cultural resources, so that the gaming technology supporting not only the goal of activity, but also ongoing cultural change as dramatic stories played a role in knitting together students’ social memories and emerging understandings.

Student: “One time, when I was beating my brother, he came after me and threw it at my head. But, like when he did that, it whizzed across the room but in the right time, so like on the screen he scored the winning shot. My family always jokes about that... he had to throw it at me to win. I guess it, like, accelerated just right?”

Elicits more general ‘imagineering’

The act of repurposing itself is an inventive and creative act. It requires the re-envisioning of one thing for completely new purposes. This kind of deconstructive and reconstructive thinking often leads to further episodes of ‘imagineering’ and design. Because the project-based classroom we were working in had a focus on design, we conducted several brainstorming sessions with students, asking them to invent other ways they could use the technology for their own ends. Students came up with creative answers such as using the IR camera for a “mom detector,” an “automated pet feeding mechanism,” and a “Halloween candy counter.” They used the accelerometer to design a means of determining a pet’s activity level, and a “little brother running in the house” alarm, etc. Several technically minded students rewrote sections of our code to change the interface, and got intrigued with the practical possibilities for building some of these ‘imagineered’ designs. Sometimes students expressed changing identities in relation to imagineering. For example, in a casual debriefing interview toward the end of the school year, a student was asked what she learned:

Student (laughs teasingly)...: well... to play the game better!

Researcher (teasing): “Well, at least something we did together was useful then...”

Student: Well, really I wasn’t into it. I wasn’t very good before. At tennis, I mean... (long pause)... I think I’ll make a game one day. Do you know the light saber one?”

Researcher: (shakes head no)

Student: “I found it online. Like... well you can pair your Wii remote to this game a guy made. Then your Wii acts like a light saber. It’s.. well real simple... like someone just made it. You don’t have to be a big company, really, if you have you know, the stuff. I mean, if you know how to get into it... and anyone can download it online for free.”

Reveals everyday invisible processes/ data

Games rely on the interpretation of a data stream to construct visualizations that enable play. Students who hack gaming technology gain an understanding of what data is, how it is useful, and why it’s valuable to be able to collect lots of it over time. They start to ‘see’ the data that drives much of the technology that makes up their mediated worlds, and through looking at the computer code that processes it, start to understand why digital things look and act the way they do. One simple example among many: A student asked, “So, what about the remote control to my TV? Is that Bluetooth too? And what is it sending... like a number or something for the computer in the TV to change the channel?”

Discussion

These interactions illustrate why, in certain contexts, repurposing everyday devices for science and math learning may make more sense than using specialized, unfamiliar technologies. This may be particularly true for middle school age children, who generally are constructing and asserting both social and academic/

mathematical/scientific identities that carry through into their high school years. By bringing artifacts and context from children's social worlds into the classroom, we are able to draw on their interests, cultural supports and expertise to support inquiry. Of course, in so doing children begin to construct new cultural experiences, ones imbued with science, that also spill over into their social worlds outside of school. Not only does harnessing familiar technologies offer students opportunities to informally synthesize prior experiences, but it also offers them 'prior' experiences to draw upon when they next encounter the tool in a gaming context. Tapping into the game controller as a cultural as well as scientific tool offers cultural groundings for the scientific understanding of future informal gaming experiences. While the one student joked that we "ruined" the game for him, his experiences appear to have opened up, or perhaps reinforced, pathways towards a socially supported identity as not only a gamer, but also as someone who knows science and knows how to code.

For those interested in learning design for CSCL, this raises questions about the plethora of devices beyond game controllers that technological affordances and hold potential for tapping into social contexts to support children's exploration of fundamental STEM concepts. With what supports and interfaces, and in what contexts, could other everyday data-enabled devices be positioned to make science concepts more "culturally available" (Roth, 1994, 1999) in everyday experience? Our study of learning with game controllers indicated that several factors might be worth noting when considering harnessing other everyday technologies for learning. Engaging with tools that are socially situated, that tap into felt experiences, that draw upon prior technical expertise and narratives, that bring hidden processes to light and that call these into future possibilities via 'imagineering' sketching and talk might offer children opportunities to more deeply connect their social identities with math and science.

Although this study indicates that adapting everyday technologies in the interests of math and science curricula is useful, at a technical level it isn't yet easy. It is only recently that three forces have come together to make innovation work. A growing number of manufacturers have opened up APIs to software developers, who have enthusiastically developed middleware and published "how to" videos and articles to support a more general, and growing, DIY (Do It Yourself) movement that includes creative individuals, component retail outlets, art organizations, and tech enthusiasts. Open scripting languages such as Processing that facilitate the rapid generation of code have helped interested people develop useful software quickly and inexpensively. Even more critical, easily downloaded middleware such as "Osculator," has enabled Bluetooth pairing of Wii Remotes to a computer so data is accessible.

While these efforts have launched a hacking movement that has started to be picked up by teachers, there is still little ongoing communication between classroom teachers and those enthusiasts who are developing tools that make everyday devices accessible. Having spent several years conducting these activities in classrooms with teachers and students, we would like to suggest that applying learning design, and not just usability design, would help make it easier for classroom teachers to take advantage of the technologies that are available in the interests of math and science. Several basic technical design considerations would be helpful for supporting learning. Among these of course include prioritizing social and collaborative features; building with transparency in mind; bringing cultures of science (terms, metaphors and analogies) into 'play' to map affordances to core science and math concepts; and offering alternative interfaces to support multiple representations of data, from game characters in motion to graphs that represent change over time.

Conclusion

One of the major struggles in science education is creating contexts in which all students, regardless of gender or social background, can see themselves as connected to science. Facilitating "science talk" (Lemke, 1990) and "transformative conversations" (Polman & Pea, 2001) in everyday interaction can help students forge an identity as investigators and inventors in science and engineering. Using gaming or other social technologies for scientific inquiry holds promise for integrating students' playful identities with their emerging identities as scientific investigators, thinkers, and inventors. In addition to integrating their social worlds into gaming, they also brought their experiences of science back into their social worlds. We see this as a kind of identity work by which students expand their sense of possible future selves.

We'd like to conclude by raising a large open set of possibilities for CSCL research. Computational and sensing technology is increasingly embedded in everyday items. This means that everyday items contain components with deep affordances for teaching science and mathematics. As increasing numbers of students are learning outside of school, via online courses, via their own online research, or by tinkering and hacking on their own, this raises large opportunities and questions for the field of technology design. If game controllers, with the addition of learning design interfaces, hold the possibility to lead students toward physics, what about other devices? What would a hammer look like that was not only designed for usability to be ready-at-hand for hammering, but also designed for learning about kinetic and potential energy? What if it could lead a user not only how to hammer efficiently, but reveal the fundamental principles of physics on which such motion functions? Or, what design factors could be added to a merry-go-round to help riders understand centrifugal force? What collaborative features could be built in? For DIY enthusiasts interested in learning, what interfaces

and components would make everyday objects embedded in our social worlds not only usable, but deeply conceptually and socially instructional? As core education is increasingly distributed in out of school contexts, it is time to start considering how everyday things might lend themselves to teaching the fundamentals students need to know. After all, they will be using science and mathematics to invent the next, newest everyday things.

References

- Barron, B., Martin, C., Roberts, E., Osipovich, A., & Ross, M. (2002). Assisting and assessing the development of technological fluencies: Insights from a project-based approach to teaching computer science. In *Proceedings of the Conference on Computer Support for Collaborative Learning*, 668-669.
- 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.
- De Bruyn, L., Leman, M., and Moelants, D. (2008). Quantifying childrens' embodiment of musical rhythm in individual and group settings. In *Proceedings of the 10th International Conference on Music Perception and Cognition*, 662-667.
- Druin, A. (1999). Cooperative inquiry: developing new technologies for children with children. In *Proceedings of the ACM conference on Human Factors in Computing Systems*, 592-599.
- Edelson, D.C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences*, 11(1), 105-121.
- Graves, L., Stratton, G., Ridgers, N.D., and Cable, N.T. (2007). Comparison of energy expenditure in adolescents when playing new generation and sedentary computer games: cross sectional study. *BMJ* 335(7633), 1282-1284.
- Hill, C.S. (2009). Developing psychomotor skills the Wii way. *Science* 323(5918), 1169.
- Hornecker, E., and Buur, J. (2006). Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 437-446.
- Inkpen, K.M. (1999). Designing handheld technologies for kids. *Personal and Ubiquitous Computing* 3(1), 81-89.
- Iwata, S. (2010). Corporate management policy briefing semi-annual financial results briefing for FY ending March 2001. Public presentation, October 29, 2010, Tokyo Japan.
<http://www.nintendo.co.jp/ir/en/library/events/101029/index.html>
- Lee, J.C. (2008). Hacking the Nintendo Wii Remote. *Pervasive Computing*, IEEE, 7(3), 39-45.
- Lemke, J. L. (1990). Talking Science: Language, Learning, and Values (Vol. 1). Ablex Pub.
- Lewis, S., Acholonu, U., & Ju, W. (2012). Using low cost game controllers to capture data for 6th grade science labs. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, 1115-1124.
- Norman, D. (2002). The Design of Everyday Things. Basic books.
- Nussbaum, J & Novick, S. (1982). Alternative frameworks, conceptual conflict and accommodation: Toward a principled teaching strategy. *Instructional Science*, 11(3), 183-200.
- Polman, J. L., & Pea, R. D. (2001). Transformative communication as a cultural tool for guiding inquiry science. *Science Education*, 85(3), 223-238.
- Pearson, E. and Bailey, C. (2007). Evaluating the potential of the Nintendo Wii to support disabled students in education. *ICT: Providing choices for learning: Proceedings ascilite Singapore 2007*.
- Rode, J.A., Stringer, M., Toye, E.F., Simpson, A.R., and Blackwell, A.F. (2003). Curriculum-focused design. *Proceedings of the 2003 Conference on Interaction Design and Children*, 119-126.
- Roth, W.M. (1994). Experimenting in a constructivist high school physics laboratory. *Journal of research in Science Teaching*, 31(2), 197-223.
- Roth, W.M., McGinn, M. K., Woszczyna, C., & Boutonne, S. (1999). Differential participation during science conversations: The interaction of focal artifacts, social configuration, and physical arrangements. *The Journal of the Learning Sciences*, 8(3-4), 293-347.
- Somers, W., Rooney, F., and Ochoa, R. (2009). Introductory Physics Experiments Using the Wiimote. *Bulletin of the American Physical Society*, 54.
- Vannoni, M. and Straulino, S. (2007). Low-cost accelerometers for physics experiments. *European Journal of Physics* 28(5), 781.
- Wheeler, M.D. (2011). Physics experiments with Nintendo Wii controllers. *Physics Education*, 46(1), 57.
- Williams, A., and Rosner, D.K. (2010). Wiimote hackery studio proposal. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction*, 365-368.

Acknowledgments

We wish to acknowledge Terry Winograd and the Wallenberg Foundation for their generous assistance with this research, as well as the teachers and students for their time, creative ideas and contributions.

Delaying Instruction Alone Doesn't Work: Comparing and Contrasting Student Solutions is Necessary for Learning from Problem-Solving prior to Instruction

Katharina Loibl, Nikol Rummel, Institute of Educational Research, Ruhr-Universität Bochum
Email: katharina.loibl@rub.de, nikol.rummel@rub.de

Abstract: Recent studies have shown benefits of problem-solving prior to instruction. However, it is unclear whether these benefits are based on the cognitive processes related to the problem-solving activity prior to instruction or originate from comparing and contrasting students' solutions to the canonical solution during subsequent instruction. To separate these effects, we conducted a quasi-experimental study with 240 students varying the two factors *timing of instruction* (problem-solving prior to instruction versus instruction prior to problem-solving) and *form of instruction* (standard instruction versus instruction that compares and contrasts typical student solutions). Our results indicate that comparing and contrasting typical student solutions is a prerequisite for the effectiveness of problem-solving prior to instruction. Problem-solving prior to instruction combined with instruction where student solutions were compared and contrasted to the canonical solution outperformed all other conditions. Problem-solving prior to standard instruction was no more effective than standard instruction prior to problem-solving.

Introduction

Can learning be best promoted by providing or by withholding instructional support? This so-called assistance dilemma (Kapur & Rummel, 2009; Koedinger & Aleven, 2007) targets the question of how to best balance the amount and timing of the instructional support given to learners. The potential benefits of delaying instruction have been shown in recent studies (Kapur, 2010, 2012; Kapur & Bielaczyc, 2012; Roll, Aleven, & Koedinger, 2009, 2011; Schwartz & Martin, 2004). In these studies students who first solved problems to a yet unknown concept before receiving instruction outperformed students who received direct instruction (i.e. instruction without previous problem-solving). It seems that solving problems which require the application of a yet unknown concept prepares students for understanding the concept in the subsequent instruction (Schwartz & Martin, 2004). It has been argued that problem-solving prior to instruction allows students to activate their prior knowledge about the domain (e.g. Kapur & Bielaczyc, 2012; Schoenfeld, 1992).

In most of the studies cited above, students worked in small groups. When solving problems prior to instruction collaboratively, students can co-construct a shared understanding that goes beyond the understanding of each individual (Moschkovich, 1996). Indeed, Sears (2006) could show that students who collaboratively solved problems prior to instruction engaged in knowledge-sharing behavior. Furthermore, they outperformed students who solved problems individually on transfer problems. The potential benefits of collaboration during problem-solving relate to the general finding that collaborative learning can promote deeper elaboration (Teasley, 1995). However, during collaborative problem-solving prior to instruction students usually invent non-canonical and incomplete solutions (Kapur & Bielaczyc, 2012). Therefore carefully designed subsequent instruction is needed to lead students towards the canonical solution.

Most research on problem-solving prior to instruction has focused on designing the problem-solving phase (e.g. with or without collaboration, Sears, 2006; with or without support, Roll, Holmes, Day, & Bonn, 2012; Westermann & Rummel, 2012a, 2012b), while the instruction phase has received less attention (Collins, 2012). As students usually fail to invent the canonical solution themselves during collaborative problem-solving, instruction is needed to ensure that students learn the correct solution method in the end. How the *form of instruction* contributes to the effectiveness of collaborative problem-solving prior to instruction has not been investigated so far. Upon closer inspection of the instruction provided in the studies by Kapur (e.g. 2010, 2012), it becomes apparent that the form of instruction might indeed be a relevant aspect: In the instruction prior to problem-solving control condition (called Direct Instruction, DI) the teacher directly presented the canonical solution (with or without explaining the structural relevant features of the formula, see Kapur & Bielaczyc, 2011). In the problem-solving prior to instruction condition (called Productive Failure, PF) the teacher compared typical student-generated solutions and contrasted them to the canonical solution during instruction in a classroom discussion. Thus, when comparing instruction prior to collaborative problem-solving and collaborative problem-solving prior to instruction, the two variables *timing of instruction* and *form of instruction* were confounded. There is reason to believe that the confounded variable (i.e. the form of instruction) is relevant for the results of problem-solving prior to instruction: When problem-solving prior to instruction is compared to an augmented instruction prior to problem-solving condition where the teacher explains the structural relevant features of the canonical solution (called Strong-DI condition), the learning differences

between problem-solving prior to instruction and instruction prior to problem-solving conditions are reduced (Kapur & Bielaczyc, 2011). One might infer from this result that the form of instruction may play a crucial role to explain the beneficial effect of problem-solving prior to instruction. Although this study presents an attempt to align the instruction in both conditions, the instruction in the augmented instruction prior to problem-solving condition did not build on student-generated (i.e. erroneous) solutions. Thus, the two variables timing of instruction and form of instruction were still confounded.

We argue that comparing non-canonical student solutions to the canonical solution during instruction may help students to detect differences between their own prior ideas and the canonical solution. This process of detecting differences by comparisons is analogous to learning with contrasting cases that fosters students to distinguish between cases (e.g. Rittle-Johnson & Star, 2011; Schwartz & Martin, 2004). Detecting differences between cases or solution approaches can guide students' attention to the structural relevant features of the new content (on the effectiveness of comparing erroneous and correct examples see Durkin & Rittle-Johnson, 2012; Große & Renkl, 2007). Against this background, a classroom discussion about typical erroneous solutions may also be fruitful in instruction prior to problem-solving conditions: In such a classroom discussion the teacher can meet students at their level of knowledge and understanding (for the importance of meeting students at their level of understanding see Wittwer & Renkl, 2008) and make discrepancies between the canonical solution and possible erroneous ideas explicit (Smith, diSessa, & Roschelle, 1994). Research demonstrated that students process the canonical solution more deeply when they realize impasses and errors (van Lehn, Silver, Murray, Yamauchi, & Baggett, 2003) and that the realization of an impasse can be triggered by warning about possible errors before presenting the instructional explanation (Acuña, García-Rodicio, & Sánchez, 2010; Sánchez, García-Rodicio & Acuña, 2009).

Taking these findings together, it seems an important next step to investigate the role of taking up typical student-generated (i.e. non-canonical) solutions during instruction in problem-solving prior to instruction settings. The studies cited above indicate that students activate their prior knowledge during problem-solving which prepares them for subsequent instruction (e.g. Kapur & Bielaczyc, 2012; Schwartz & Martin, 2004). We argue that in addition to the cognitive processes related to the problem-solving activity, the form of instruction merits attention: Comparing student solutions and contrasting them to the canonical solution during instruction might be a necessary component for the effectiveness of problem-solving prior to instruction. Activating prior knowledge during problem-solving can only be effective, if students connect their prior knowledge to the new content and realize differences. Contrasting student solutions to the canonical solution helps students to connect their prior ideas to the new content and to focus on the distinguishing features, which in turn may foster the acquisition of conceptual knowledge. By contrast, problem-solving prior to instruction might be less productive for fostering procedural skills (Sleeman, Kelly, Martinak, Ward, & Moore, 1989) as it might reduce the time needed for acquiring procedural skills through practice (Klahr & Nigam, 2004; Rittle-Johnson, Siegler, & Alibali, 2001).

Against this background we hypothesized that collaborative problem-solving prior to instruction combined with subsequent instruction where student solutions are contrasted to the canonical solution is most effective to acquire conceptual knowledge, but instruction prior to collaborative problem-solving may lead to better procedural skills.

Methods

Study Design

Table 1: Experiment design with final sample size

		Form of instruction	
		Standard instruction	Instruction that compares and contrasts typical student solutions
Timing of instruction	Problem-solving prior to instruction	PS-I (<i>N</i> = 51, 3 classes)	PS-I _{contrast} (<i>N</i> = 56, 3 classes)
	Instruction prior to problem-solving	I-PS (<i>N</i> = 62, 3 classes)	I _{contrast} -PS (<i>N</i> = 71, 4 classes)

To separate the effects of the sequence of problem-solving and instruction, and of comparing and contrasting student solutions to the canonical solution during instruction, we conducted a quasi-experimental study with two

factors: we varied *timing of instruction* (problem-solving prior to instruction versus instruction prior to problem-solving) and *form of instruction* (standard instruction focusing on the canonical solution versus instruction that compares and contrasts typical student solutions to the canonical solution). Table 1 gives an overview of the conditions. Participants were 240 10th graders (13 classes) recruited from four secondary schools in Germany. For practical reasons, classes were randomly assigned to conditions as a whole. The resulting conditions did not differ significantly concerning prior knowledge as measured by a pretest ($F[3,234] = 0.47, p = .71$).

Learning Material

To be able to compare the results of our study to those of other studies on problem-solving prior to instruction (e.g. Kapur, 2012; Roll et al., 2009; Schwarz & Martin, 2004), our learning material addressed the same concept that has been targeted in those studies: the concept of variance. Students in grade 10 of German secondary schools have not covered this topic yet.

The learning task was aligned to the task used by Kapur (2012) and was the same in all conditions: At the beginning of the first learning phase (i.e. instruction for I-PS and I_{contrast}-PS, problem-solving for PS-I and PS-I_{contrast}) students were provided with a table listing the number of goals that three fictitious soccer players had scored in the last 10 years. Students were asked to answer the question who the most consistent goal scorer was. Range and mean of the number of goals was the same for all three players to force students to think about strategies beyond their formal prior knowledge.

Experimental Procedure

Instruction and problem-solving phases respectively took place during a lesson of 45 minutes on two consecutive days. During problem-solving, students worked in groups of three in all conditions. The same experimenter gave the instruction in all conditions. In the instruction phase of all conditions, the experimenter explained the concept and the canonical solution using the example of the goal scorers. Prior to both learning phases (i.e. prior to the first lesson) students completed a pretest on related content (e.g. mean, range, box plot, graphical representations). After both learning phases (i.e. after the second lesson), students completed a posttest.

In the problem-solving prior to instruction conditions (PS-I and PS-I_{contrast}), students dealt with the task to identify the most consistent goal scorer during the first lesson. During this problem-solving phase, the task asked them to invent as many solutions as possible. Students used tablet PCs to generate and exchange solution ideas. The use of tablet PCs allowed students to work individually as well as to share their ideas and focus the group's attention on selected ideas. During the second lesson, students received instruction.

In the instruction prior to problem-solving conditions (I-PS and I_{contrast}-PS), students first received instruction. The problem-solving phase took place during the second lesson where students solved problems isomorphic to the one discussed during instruction.

In the standard instruction conditions (PL-I and I-PL), the experimenter first presented the problem of the three soccer players and discussed the meaning of consistency with the class. This introduction was followed by a presentation of several approaches (graphical approaches, range, mean absolute deviation, and standard deviation). The class discussed the advantages and disadvantages of the different approaches (e.g. graphical approaches might be imprecise, range is sensitive to outliers). Finally the experimenter explained the structurally relevant features of the canonical solution.

In the conditions with instruction that compares and contrasts typical student solutions (PS-I_{contrast} and I_{contrast}-PS), the experimenter presented and compared typical student-generated solutions (e.g. graphical approaches, range, number of times the soccer player scored at the mean, deviation from one year to the next with or without absolute values) and discussed whether these approaches were suitable to solve the problem by contrasting them to the canonical solution. It should be stressed, that the solutions were not the very solutions generated by students during problem-solving in this study. Rather, the solutions were *typical* student-generated solutions (taken from previous studies and pilots) that matched the solution types most often generated. Notably these solution types were similar to the ones usually generated by students in Singaporean classes in previous studies (cf. Kapur, 2012). Finally the experimenter explained the structurally relevant features of the canonical solution.

Dependent Variables

A posttest assessed the learning outcomes after the second lesson. It included items testing for procedural skills and items testing for conceptual knowledge. Students had 30 minutes to answer the posttest items. All students finished the posttest in time.

The items testing for *procedural skills* required students to solve problems isomorphic to the one discussed during instruction. Students received 1 point for each correct calculation with a deduction of 0.5 point for computation errors. They received 1 additional point in cases where they had to compare two deviations.

Students could achieve a maximum of 4 points (i.e. 1 item required a single calculation, 1 item required the calculation of two deviations including a comparison).

The items testing for *conceptual knowledge* required students to decompose the canonical solution into its structurally relevant features (cf. Roll et al., 2011) and to translate between graphical and algebraic strategies: Two items presented incorrect solutions and asked students to detect the errors and to reason mathematically. For the reasoning, students had to decompose the canonical solution and refer to these structurally relevant features of the canonical solution. Students received 0.5 point for the detection of each error. They received an additional 0.5 point per detected error for correct reasoning about the structurally relevant feature. Figure 1 presents one example. Two other items required sense-making using both graphical representations and the structurally relevant features of the canonical solution. Students received 0.5 point for each structural feature correctly represented in the graphical representation. Taking all conceptual knowledge items together, students could achieve a maximum of 7 points (3 points for the first type of items, 4 points for the second type of items).

One student calculated the consistency the following way.

How did he calculate consistency? Is the method suitable to measure consistency? Explain why or why not.

$$\frac{(x_2 - x_1) + (x_3 - x_2) + (x_4 - x_3)}{N} = \frac{(50 - 30) + (90 - 50) + (70 - 90)}{4} = 10$$

Error: Deviation from one value to the next instead of deviation from the mean. (0.5 point)

Reasoning: Sensitive for sequence of data points as there is no fixed reference point. (0.5 point)

Error: No absolute or squared values; deviations may be negative. (0.5 point)

Reasoning: Positive and negative values might cancel out. (0.5 point)

Figure 1. Example of one item testing for conceptual knowledge with solution and coding.

In the problem-solving prior to instruction conditions (PS-I and PS-I_{student}) students used tablet PCs to invent their solutions during problem-solving. This enabled us to collect audio and screen recordings of students' collaborative problem-solving prior to instruction. We are currently analyzing the process of inventing and discussing solution ideas in the small groups as well as coding the quantity and quality of the invented solution ideas.

Results

We performed a two-factorial MANOVA with the factors *timing of instruction* and *form of instruction* and the outcome variables procedural skills and conceptual knowledge. Table 2 shows the means and standard deviations.

Table 2: Means and standard deviations of the posttest results.

Conditions	Procedural skills	Conceptual knowledge
PS-I	3.24 (0.99)	1.29 (1.02)
PS-I _{contrast}	2.99 (1.27)	2.63 (1.53)
I-PS	3.27 (1.02)	1.17 (1.23)
I _{contrast} -PS	3.41 (0.91)	1.68 (1.35)

For *procedural skills*, we found only a marginally significant effect for timing of instruction favoring instruction prior to problem-solving ($F[1,236] = 2.81, p = .095, \eta_p^2 = .01$). Neither the form of instruction ($F[1,236] = 0.16, p = .69$) nor the interaction of timing and form ($F[1,236] = 1.93, p = .17$) was significant.

For *conceptual knowledge* we found a small significant effect for timing of instruction favoring problem-solving prior to instruction ($F[1,236] = 10.02, p = .002, \eta_p^2 = .04$) and a large significant effect for form of instruction favoring instruction based on typical student solutions ($F[1,236] = 29.35, p < .01, \eta_p^2 = .11$). We further found a significant interaction ($F[1,236] = 5.90, p = .02, \eta_p^2 = .02$) indicating that the form of instruction has a higher effect in the problem-solving prior to instruction conditions. In order to compare the effects of the two factors and their combination, we additionally calculated posthoc comparisons (LSD) between all conditions. The pair-wise comparisons revealed significant differences between the $I_{\text{contrast-PS}}$ condition to the I-PS ($p = .02$) condition and to the $PS-I_{\text{contrast}}$ condition ($p < .01$), that is, the $I_{\text{contrast-PS}}$ condition significantly outperformed the condition that received standard instruction first (I-PS), but was outperformed by the $PS-I_{\text{contrast}}$ condition that combined problem-solving prior to instruction with instruction where student solutions were compared and contrasted to the canonical solution. The comparison between the I-PS condition and the PS-I condition was not significant ($p = .61$), that is, the timing of instruction had no effect when combined with standard instruction. In other words, for conceptual knowledge, problem-solving prior to instruction was only more effective than instruction prior to problem-solving if student solutions were compared and contrasted during the instruction.

Discussion

Previous studies have shown benefits of problem-solving prior to instruction for the acquisition of conceptual knowledge. These benefits may stem from the cognitive processes related to the problem-solving activity prior to instruction or they may originate from the specific form of subsequent instruction that compares and contrasts students' solutions to the canonical solution during instruction. In our study we aimed at separating the effects of *timing of instruction* (problem-solving prior to instruction versus instruction prior to problem-solving) and of *form of instruction* (standard instruction focusing on the canonical solution versus instruction that compares and contrasts typical student solutions to the canonical solution). We tested for learning effects on conceptual knowledge and procedural skills.

Our findings support the notion that problem-solving prior to instruction can prepare students for the acquisition of *conceptual knowledge* from subsequent instruction as indicated by the main effect for timing of instruction. In this regard, our study replicates the beneficial effect of problem-solving prior to instruction found by others (e.g. Kapur, 2009, 2012; Roll et al., 2011; Schwartz & Martin, 2004).

Moreover, the form of instruction appears to be of central relevance: Comparing typical non-canonical student solutions and contrasting them to the canonical solution during instruction may guide students' attention to the structurally relevant aspects of the content and thereby promotes learning. As indicated by the main effect of form of instruction and the pair-wise comparisons, comparing and contrasting typical student solutions to the canonical solution is beneficial in both settings: problem-solving prior to instruction and instruction prior to problem-solving. Similar to these results, we already showed in an earlier study (Westermann & Rummel, 2012b) that even in an instruction prior to problem-solving setting it is beneficial for learning if instruction builds on typical student solution in comparison to standard instruction.

The most interesting finding of our study is the interaction effect showing that the beneficial effect of problem-solving prior to instruction only comes to bear if the teacher (or in our study the experimenter) compares typical student solutions and contrasts them to the canonical solution during instruction. More specifically the $PS-I_{\text{contrast}}$ condition, that is problem-solving prior to instruction combined with instruction where student solutions are compared and contrasted to the canonical solution, outperformed all other conditions. This finding suggests a dual learning mechanism: In a first step, problem-solving prior to instruction prompts students to activate their prior knowledge and to generate own solution ideas (cf. Kapur & Bielaczyc, 2012). In a second step, comparing student solutions and contrasting them to the canonical solution during instruction helps students to detect differences between their own prior ideas and the canonical solution. The detection of differences guides students' attention to the structurally relevant aspects of the content (cf. Durkin & Rittle-Johnson, 2012). Focusing the attention on the most important aspects in turn helps students to process these aspects deeply and fosters the acquisition of conceptual knowledge (Renkl, 2008).

Furthermore the difference between the $I_{\text{contrast-PS}}$ condition and the $PS-I_{\text{contrast}}$ condition suggests that connecting prior knowledge to the new content and detecting differences between non-canonical solutions and the canonical solution works better when students indeed activated their prior knowledge during problem-solving and generated solutions themselves. In the study cited above (Westermann & Rummel, 2012b), we only found a descriptive, but not statistically significant difference between the two conditions with different timing of instruction (before versus after problem-solving) where instruction build on typical student solutions. How do these two studies differ? First of all, the sample size of our study presented here is higher. Secondly, we conducted the previous study (Westermann & Rummel, 2012b) at two schools from the same well-educated neighborhood. Students from these schools might have been higher motivated in connecting the new content to their prior knowledge and therefore prompting them to activate their prior knowledge first might have been less

important. The schools of our current study were located in four different neighborhoods resulting in a more representative sample.

In addition to the learning effects on conceptual knowledge, our findings confirm that time for practicing problem-solving after instruction is needed to foster *procedural skills* (cf. Rittle-Johnson et al., 2001): Both instruction prior to problem-solving conditions (I-PL and I_{contrast}-PS) outperformed both problem-solving prior to instruction conditions (PS-I and PS-I_{contrast}) on items testing for procedural skills. This finding is not surprising as the latter conditions had no time to practice problem-solving after learning the canonical solution. Studies that found no difference between instruction prior to problem-solving and problem-solving prior to instruction on items testing for procedural skills usually allowed practice for students in the problem-solving prior to instruction conditions after students received the canonical solution during instruction (e.g. Kapur, 2010, 2012; Roll et al., 2009). Taken together, our findings underline the importance of defining the learning goal when choosing one instructional approach over the other.

Limitations and Outlook

Although our study yields interesting results, we would like discuss some limitations and give an outlook to future research. Inspired by the in vivo research paradigm advocated of the Pittsburgh Science of Learning Center (Koedinger, Corbett, & Perfetti, 2012), we conducted our study in the field with real learners and real learning content, which promotes the external validity of the study. However, this also yielded some problems: The implementation in schools forced us to conduct a quasi-experimental study for organizational reasons. Thus, prior differences between conditions cannot be completely excluded due to randomizing at the class level.

Another aspect that has to be considered is the fact that the experimenter who taught the instruction in our study was very familiar with the material used during instruction, the structurally relevant features of the canonical solution, and the typical student solutions. This knowledge might relate to student achievement (Tchoshanov, 2011). Tchoshanov showed that teacher content knowledge is associated with lesson quality and student achievement in mathematics. Especially when building on student-generated solutions it seems crucial to be familiar with these solutions. In order to ensure a smooth implementation in the field, teachers might need to be provided with new resources and strategies (Meder, Schüpbach, & Krause, 2011) as building on student-generated solutions imposes high demands on the teacher.

When focusing on the effect of connecting the new content to the prior knowledge it should be noted that the solutions used in the instruction phase of the I_{contrast}-PS condition and the PS-I_{contrast} condition were *typical* student-generated solutions (taken from previous studies and pilots) that matched the solutions most often generated in the problem-solving prior to instruction conditions and not the very own solution of the students. Yet, until this date, it has not been systematically investigated whether using the very own solutions of students in comparison to typical student-generated solutions during instruction would further help students to connect their prior knowledge to the new content and to detect differences between their intuitive solutions and the canonical solution.

Solution approaches invented prior to instruction are generally incomplete or erroneous (e.g. Kapur & Bielaczyc, 2012). Nevertheless, the diversity, that is the number of different solution ideas, seems to have a positive effect on posttest performance (Kapur, 2012; Kapur & Bielaczyc, 2012). While Kapur and colleagues claim that the positive effect of diversity is independent of the quality of the solution ideas, others did find that the quality of the invented solutions matters (Wiedmann, Leach, Rummel, & Wiley, 2012). In accordance with the finding of Wiedmann and colleagues, we hypothesize that the more knowledge components are shared between the invented solutions and the canonical solution, the easier it should be to connect the prior ideas to the new content during instruction. As indicated by our findings, the connection between prior ideas and the canonical solution may lead to deeper processing and in turn promote learning. We recorded process data of the problem solving prior to instruction conditions (PS-I and PS-I_{contrast}) that allow us to code the quantity and quality of the invented solution ideas. For future analysis, we aim at testing for possible relations between these codings and learning outcomes.

References:

- Acuña, S. R., García-Rodicio, H., & Sánchez, E. (2010). Fostering active processing of instructional explanations of learners with high and low prior knowledge. *European Journal of Psychology of education, 26*(4), 435-452.
- Collins, A. (2012). What is the most effective way to teach problem solving? A commentary on productive failure as a method of teaching. *Instructional Science, 40*(4), 731-735.
- Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction, 22*(3), 206-214.
- Große, C. S. & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction, 17*(6), 612-634.

- Kapur, M. (2010). A further study of productive failure in mathematical problem solving: Unpacking the design components. *Instructional Science*, 39(4), 561-579.
- Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, 40(4), 651-672.
- Kapur, M., & Bielaczyc, K. (2011). Classroom-based Experiments in Productive Failure. In L. Carlson, C. Hoelscher, & T.F. Shipley (Eds.) *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2812-2817). Austin, TX: Cognitive Science Society.
- Kapur, M., & Bielaczyc, K. (2012). Designing for Productive Failure. *The Journal of the Learning Sciences*, 21(1), 45-83.
- Kapur, M., & Rummel, N. (2009). The assistance dilemma in CSCL. In: A. Dimitracopoulou, C. O'Malley, D. Suthers, & P. Reimann (Eds.), *Computer supported collaborative learning practices - CSCL2009 community events proceedings*, Vol 2 (pp. 37-42). Berlin: International Society of the Learning Sciences.
- Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, 15(10), 661-667
- Koedinger, K. R., & Alevan, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, 19(3), 239-264.
- 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.
- Meder, L., Schüpbach, H., & Krause, A. (2011). Sind innovative Lehr- und Lernformen für Schüler wie auch für Lehrkräfte vorteilhaft? Internationale Vergleichsstudie zum Zusammenhang von Lehr-/Lernformen, Unterrichtsqualität und psychischen Belastungen der Lehrkräfte [Are innovative teaching and learning methods for teachers as beneficial as for students? International comparative study about the relations between teaching/ learning methods, quality of instruction and psychological stress in classroom teachers]. In Landesstiftung Baden-Württemberg (Ed.), *Programm Bildungsforschung* (pp. 197-223). Stuttgart: Landesstiftung Baden-Württemberg.
- Moschkovich, J. N. (1996). Moving up and getting steeper: Negotiating shared descriptions of linear graphs. *Journal of the Learning Sciences*, 5(3), 239-277.
- Renkl, A. (2008). Lehren und Lernen im Kontext der Schule [Teaching and learning in school contexts]. In Renkl (Ed.), *Lehrbuch Pädagogische Psychologie* (pp. 109-153). Bern: Huber.
- Rittle-Johnson, B., Siegler, R., & Alibali, M. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93(2), 346-362.
- Rittle-Johnson, B., & Star, J. R. (2011). The power of comparison in learning and instruction: Learning outcomes supported by different types of comparisons. In B. Ross & J. Mestre (Eds.), *Psychology of Learning and Motivation: Cognition in Education* (Vol. 55, pp. 199-226). San Diego: Elsevier.
- Roll, I., Alevan, 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. 1169-1174). Austin, TX: Cognitive Science Society.
- Roll, I., Alevan, V., & Koedinger, K. R. (2011). Outcomes and mechanisms of transfer in invention activities. In L. Carlson, C. Hoelscher, & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 2824-2829). Boston, Massachusetts: Cognitive Science Society.
- Roll, I., Holmes, N., Day, J. & Bonn, D. (2012). Evaluating metacognitive scaffolding in Guided Invention Activities. *Instructional Science*, 40(4), 691-710.
- Sánchez, E., García Rodicio, H., & Acuña, S. R. (2009). Are instructional explanations more effective in the context of an impasse? *Instructional Science*, 37(6), 537-563.
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for Research on Mathematics Teaching and Learning* (pp. 334-370). New York: MacMillan
- 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
- Sears, D. A. (2006). *Effects of innovation versus efficiency tasks on collaboration and learning* (Doctoral dissertation, Stanford University, California). Retrieved November 07, 2012, from <http://www.stat.auckland.ac.nz/~iase/publications/dissertations/06.Sears.Dissertation.pdf>
- Sleeman, D., Kelly, A. E., Martinak, R., Ward, R. D., & Moore, J. L. (1989). Studies of diagnosis and remediation with high school algebra students. *Cognitive Science*, 13, 551-568.
- Smith, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, 3(2), 115-163.

- Tchoshanov, M. A. (2011). Relationship between teacher knowledge of concepts and connections, teaching practice, and student achievement in middle grades mathematics. *Educational Studies in Mathematics*, 76(2), 141-164.
- Tasley, S. D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology*, 31(2), 207-220.
- van Lehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, 21(3), 209–249.
- Westermann, K. & Rummel, N. (2012a). Delaying instruction: evidence from a study in a university relearning setting. *Instructional Science*, 40(4), 673-689.
- Westermann, K. & Rummel, N. (2012b). New evidence on productive failure - Building on students' prior knowledge is key! 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* (pp. 266-270). Sydney, Australia: ISLS.
- Wiedmann, M., Leach, R. C., Rummel, N. & Wiley, J. (2012). Does group composition affect learning by invention? *Instructional Science*, 40(4), 711-730.
- Wittwer, J., & Renkl, A. (2008). Why instructional explanations often do not work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist*, 43, 49-64.

Acknowledgements

This research was funded by the Center of Educational Studies of the Professional School of Education at the Ruhr-Universität Bochum, Germany. We would like to especially thank Manu Kapur for providing us with his study materials and many background information concerning his own studies. We are thankful for the participating schools for the organizational efforts. We thank our student research assistants, Katja Goepel, Christian Hartmann, and Andreas Vogel for their help with collecting and coding the data.

Exploring Evolutionary Concepts with Immersive Simulations

Michelle Lui, James D. Slotta,
University of Toronto, 252 Bloor Street West Toronto, Canada, ON,
Email: michelle.lui@utoronto.ca, jslotta@oise.utoronto.ca

Abstract: This paper presents two iterations of our design of an immersive simulation and inquiry activity for exploring evolutionary concepts in a Grade 11 Biology course. Interacting with large projected displays of animated rainforest flora and fauna, students worked as “field researchers” to observe changes in life forms occurring over a 200 million year span. Students gathered evidence of evolution using networked tablet computers that scaffolded their interactions with peers and with the room itself. Improvements from the first to the second design iteration focused on (1) improving content-focused interactions within the simulation; (2) improving the integration of the simulation activity into the overall curriculum; (3) improving embodied interactions of students working within the physical space. Student explanations from the second implementation demonstrated increased variation in evolutionary topics compared to those in the first iteration. Key design features from the two iterations are discussed with respect to the observed interaction patterns.

Introduction

Evolution has been described as a central idea in understanding biology, accounting for fundamental issues about how organisms came to their present form, explaining relatedness among different species, as well as how certain traits are passed down and accumulated over many generations (Kampourakis & Zogza, 2008; National Academies Press, 1998). There are also strong links to understanding evolution and learning about the nature of science (Rudolph & Stewart, 1998). However, science topics of biological evolution are well recognized as being challenging to teach, due in part to their complex systemic nature (Chi & Slotta, 2006), and students’ incoming ideas, which are often inconsistent with the scientific theory (Demastes, Good, & Peebles, 1995; Mayr, 2002). The research literature on conceptual change in students’ understanding of evolutionary biology promotes a constructivist approach that takes into account students’ epistemic positions (see for example, Alters & Nelson, 2002; Anderson, 2007; Sandoval, 2003), yet it remains a challenge in determining how best to do achieve this.

One early example is that of the Biology Guided Inquiry Learning Environment (BGuILE), where students were presented with a scientific challenge concerning a Galapagos island ecosystem, where the task was to find out what was killing some of the finches on the island (Reiser et al. 2001). A technology environment prompted students to formulate an evidence-based argument, helping them articulate questions and support their explanations with data. BGuILE examined the causal claims made by students and how they warranted these claims. Results showed that students were able to adopt explanatory goals and that scaffolding students’ attention to epistemic practices helped them to focus on evidence (Sandoval, 2003).

Expanding on these ideas, Chinn and Buckland (2012) advocate a model-based inquiry approach, as well as a stronger focus on macroevolution (i.e., evolution on a grand scale, as opposed to the smaller scale processes within microevolution, such as allele frequency changes). However, such evolutionary phenomena are not easily accessible to student manipulations within a classroom setting. The present study seeks to leverage technology-enhanced learning environments in support complex and participatory forms of scientific inquiry with macroevolutionary concepts. Our research group at the University of [name withheld] has advanced the concept of a “smart classroom,” where the physical environment (e.g., walls, furniture, etc.) is infused with a set of digital tools and materials to support student interactions across physical, social and curricular dimensions. The room, together with various server and client technologies, serves to scaffold collaboration, enhancing real-time face-to-face interactions, capturing and representing the collective contributions of the entire class. Inspired by research in immersive virtual worlds, such as River City (Dede, 2009) and Second Life, we are investigating an educative role for such immersive simulations, where students are immersed within a room-sized simulation, and conceptual content is distributed across a spectrum of embedded technologies to support learning activities.

Reminiscent of how students adopt “avatars” within online immersive environments, participatory simulations also allow students to be embodied within particular roles. For example, students may become one element of a complex system, so that the emergent behavior of the system might be directly observed or experienced (Wilensky & Stroup, 1999). Such participatory role-playing can be augmented with networked technologies, such as wearable computers (e.g. “Thinking Tags”) to help provide information to the participant during the simulation (Colella, 2000; Resnick, 1996). In Colella’s work, wearable “Thinking Tags” transformed students into potential virus carriers whose mission was to greet as many people as they could without getting “sick.” By participating in the process of viral transfer, Colella hoped that students could come to a deeper understanding of the underlying concepts (i.e., of disease progression). In another approach called Embedded

Phenomenon, a persistent scientific simulation is embedded within the walls or floor of the classroom (Moher, 2008). Students are tasked with monitoring and manipulating the state of the simulation, requiring physical interactions within their learning environment: observing and measuring aspects of the simulation, forming hypotheses, and gathering evidence to solve problem or answer questions

Our own design of an immersive simulation builds upon this previous research, incorporating aspects of participatory simulations within our learning activities and a sense of full-body immersion through our projected displays (together with audio and other ambient media). The goal is to help students deeply engage in scientific inquiry, providing them with opportunities to experience evolutionary phenomena that would be geographically (Borneo) and temporally (200 million years) inaccessible to them, otherwise. This paper addresses the following research question: *How can immersive environments and embodied interactions support a co-located group of students to collaboratively develop their understanding of evolutionary concepts?* We designed EvoRoom, an immersive simulation of evolutionary biology in a Borneo rainforest, where students can observe changes in flora and fauna over a 200 million year time period. In addition to the immersive environment itself, we designed a set of learning activities for use within and outside the immersive simulation, and worked closely with the teacher to tailor her curriculum so that the activities fit well in the sequence of topic coverage (i.e., that the time spent in EvoRoom played a meaningful role in the curriculum sequence). The smart classroom technology environment served to orchestrate our complex interaction design, delivering all materials, collecting student interactions, and supporting collaborations within EvoRoom as well as at home and in the classroom. Here, we report on two iterations of our design-based research project, with findings from the first incorporated into the second. We report on students' inquiry experience, examine the content of their explanations, and discuss the features of our environment and interactions that made it successful.

EvoRoom Design

To help students learn about evolution, we required a rich context that would engage them in the exploration of macro-level evolutionary concepts while allowing enough flexibility to be tailored within our partner teacher's curriculum. We decided on the context of a tropical rainforest, due to its clear connections to our target topics of biodiversity and evolution, as well as the range of interesting features that would be well suited to an immersive environment. Ultimately, we achieved EvoRoom (see Figure 1) where students enter a simulated rainforest as a team of "field researchers," gathering evidence of evolution by comparing simulations from a range of time periods. Working individually and in groups, students observed changes in life forms over time, consolidated their findings as a collective community, and developed hypotheses about the evolutionary changes that might have taken place. Students observations (e.g., of ancestral relationships or patterns amongst species), their consultation of field guides, their written reflections, and other activities were scaffolded throughout the activity using tablet computers that and custom software application. At the front of the room were located two interactive white boards, where we aggregated observations from all students, at all time periods, for purposes of student reflection and teacher-led discussions.



Figure 1. Large screen projections around the room displaying the immersive simulation, as well as audio tracks of natural rainforest sounds transform a smart classroom into a rainforest in Southeast Asia.

Methods

Following a design-based methodology (Brown, 1992; Design-Based Research Collective, 2003), the immersive simulation was designed and evaluated over two iterations as part of a Grade 11 Biology course. Using a co-design method (Penuel, Roschelle, Shechtman, 2007), our team of researchers, designers, technology developers, and a high school teacher met regularly since January 2011 to develop curriculum activities and the

immersive simulation itself. The first iteration was evaluated in June 2011 as a pilot study, with the second iteration implemented in the following Fall/Winter (2011-2012) semester as part of the Biology course.

Participants

The first iteration was conducted with eight high-school student volunteers who had completed Grade 11 Biology. The second iteration was an evaluative study, including 45 students from two class sections of Grade 11 Biology (taught by our co-design teacher). For both design iterations, students completed pre-/post-activity questionnaires. During the activity, video recordings captured student interactions, while knowledge artifacts created by students (e.g., notes) were collected as measures of the quality of student ideas.

Procedure

The first study was conducted in a single 2-hour session, which took part in the smart classroom one week after the end of the academic school year, while the second study lasted 12-weeks and included three visits to EvoRoom, along with a set of in-class and homework activities (Table 1). The full design and expanded the curriculum are detailed in a separate paper (AUTHORS, 2012). For the purposes of examining student ideas about evolutionary concepts (and for linking ideas made in the first study), the present analysis focuses on the relevant EvoRoom sessions, each of which lasted approximately 45 minutes.

Activity Design

Iteration 1 (implemented in June 2011) was quite basic, with students entering the smart classroom to find the large displays set up as *Sundaland*, a region in Southeast Asia predating Borneo and Sumatra, about two million years ago. After the premise of the activity was introduced, and the historical context of the rainforest environment explained, students were scaffolded by their tablets to observe individual species and use the Field Guide application.

After approximately 15 minutes, the teacher used her own tablet computer (with teacher controls) to “accelerate time,” revealing a sequence of geologic events that affected the Sundaland landscape over the span of two million years. On the interactive white boards at the front of the room, students observed changes in sea level that broke Sundaland’s central landmass into a peninsula and several islands, including Borneo and Sumatra. When the teacher then set the room’s timeline to “present day”, the two sides of the simulation were updated: one side of the room (3 large projectors “stitched” together, as shown in Figure 1) now showed Borneo’s ecosystem, while the other side showed Sumatra’s. Students spent another 15 minutes making observations in this new context. At the end of the observation phase, students were divided into two field researcher teams: Borneo and Sumatra. Each group answered a set of questions designed to have students review and compare notes about their individual observations (e.g., in the Borneo group, students were asked *What common species were found in both Sundaland and Borneo?*).

In the final step, the two teams came together to collectively document evidence of evolution. Students were encouraged to discuss their ideas with others and to post ideas about evolution concepts. The posts were aggregated to the interactive white board, which served to visibly represent the collective knowledge base of the students at the end of the activity. The teacher was able to use the content of this display to lead a synthesis discussion to close the activity.

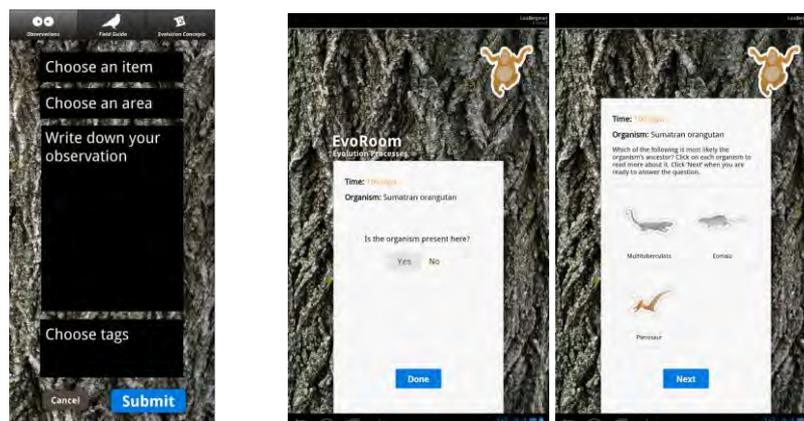


Figure 2. Tablet computer screens for iteration 1 (left) and iteration 2 (right). Note the open-ended nature of tasks given in the first iteration versus more structured format in the second.

The second iteration of the curriculum was informed by our observations and analysis of student interactions within the first. In particular, the EvoRoom activities were more deeply integrated within the broader curriculum, and interactions refined to focus on topics of evolution and biodiversity (see Table 1).

Moreover, additional effort was placed in mapping particular inquiry objects to different areas of the room. Students were assigned to one of four specialist categories (i.e. plants & insects, birds, primates, and other mammals), which they held for the duration of the curriculum. Two EvoRoom sessions were developed. For the first session, we greatly extended the timeline, such that students examined the Borneo rainforest as it may have appeared at nine different time periods (i.e., 200, 150, 100, 50, 25, 10, 5, 2 million years ago and present day), as opposed to just two (i.e., 2 million years ago and present day). Students were asked to go to each station (from 200 to 2 million years ago) and look for their assigned specialty species as part of a larger team consisting of different specialists. If the species were not present, they were asked to identify their evolutionary predecessors from a short list that popped up (scaffolded by a Field Guide application). Their answers were recorded by the tablets and – via the smart room software – aggregated in real-time on the interactive white board at the front of the room, resulting in an interactive cladogram (a diagram showing descendancy relations amongst species over time). In the second session, students again focused on evolution, working as a team in their assigned species groups on activities with similar goals as in the first iteration, but with more structured and scaffolded tasks than had occurred in the first iteration (Figure 2).

Table 1. Summary of the activity sequence for iteration 2: in-class (I), homework (H), and smart classroom (S)

Week	Description	Curricular goals
1	<ul style="list-style-type: none"> • Introduction (I) • Assign groups and specialty categories (i.e., plants & insects, birds, primates, and other mammals; I) • Review field guide (H) • Zoo field trip group assignment (I) 	<ul style="list-style-type: none"> • Become familiar with assigned organisms • Understand scientific connections (e.g. taxonomy and phylogeny) between related species
2	<ul style="list-style-type: none"> • Collaborative food web activity (I) • Assign environmental impact variable (I) • Prediction analysis group assignment (H) 	<ul style="list-style-type: none"> • Understand relationships among a set of species (e.g., in the Borneo rainforest) • Understand how environmental factors (e.g., high/low rainfall, tsunamis, earthquake) affect ecosystems
3	<ul style="list-style-type: none"> • EvoRoom: Biodiversity activity (S) • EvoRoom debrief discussion (I) • Personal reflection (H) 	<ul style="list-style-type: none"> • Improve understanding of complex interrelationships within an ecosystem and implications of environmental factors on biodiversity
4	<ul style="list-style-type: none"> • Traditional teaching on the origin of life, contributions to the theory of evolution 	
5	<ul style="list-style-type: none"> • Traditional teaching on molecular evidence of evolution and microevolution 	
6	<ul style="list-style-type: none"> • Traditional teaching on variation, selective advantage, natural selection 	
7	<ul style="list-style-type: none"> • Traditional teaching on mechanisms of evolution, including sexual selection, gene flow, genetic drift 	
8	<ul style="list-style-type: none"> • Understanding of evolution survey (H) 	<ul style="list-style-type: none"> • Reflect on personal understanding of evolution
9	<ul style="list-style-type: none"> • Relatedness of species in Borneo and Sumatra assignment (H) 	<ul style="list-style-type: none"> • Understand concept of “relatedness” and how assigned species are related to each other
10	<ul style="list-style-type: none"> • EvoRoom: Evolution processes day 1 (S) • Evolutionary mechanisms tagging (H) 	<ul style="list-style-type: none"> • Make connections between evolutionary mechanisms (learned in class) to the organisms in a specific ecosystem • Improve understanding of different organisms' lineages with respect to evolutionary forces over millions of years
11	<ul style="list-style-type: none"> • EvoRoom: Evolution processes day 2 (S) 	
12	<ul style="list-style-type: none"> • Personal reflection (H) 	

Findings

Student Observations

In iteration one, students were asked to make free-form observations about any organism shown in the simulation. A total of 157 observations were made, with 49% about the species at two million years ago, 27% about those in present day Borneo environment and 24% about the species in Sumatra. Students wrote an average of 13 words per observation ($SD=24$). These notes were analyzed following Chi's (1997) method for content analysis. Using the “observation posting” as a unit of analysis, we coded for content type and nature of the content. An inter-rater reliability analysis using the Kappa statistic was found to be $Kappa = 0.80$ ($p < 0.001$), indicating substantial agreement. The notes tended to be about physical characteristics of certain organisms (41%) or about the animal's behavior (57%) – see Figure 3 for a complete distribution of coded

categories. In iteration two, students made structured observations about whether their assigned organisms were present at different time points, and if not, which ancient is most likely its predecessor. These observations were scored for accuracy: with a total of 1047 entries, 81% ($SD=10.33$) were correct.

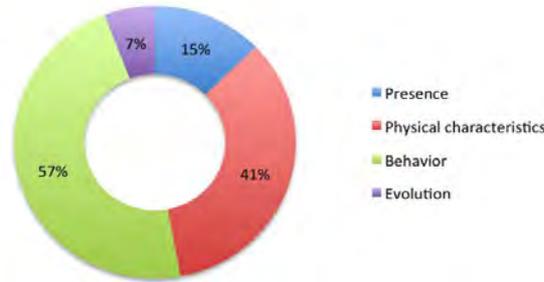


Figure 3. Content distribution of iteration one’s observations may be categorized as Presence (e.g., about the presence of species in a specific location), Physical characteristics, Behavior, or Evolution.

Students’ Conceptual Learning: Explanations of Evolution

At the end of the activity, in both iterations, students were asked to contribute to the following question: *What evolutionary forces do you think were at play in this environment?* Students were asked to choose an evolution concept from a predefined list and explain their answers with sufficient evidence. 14 explanations were collected from the first study, while the second yielded 43 (Table 2). Figure 4 shows the distribution of evolutionary concepts that the explanations attempted to address. Explanations from the first iteration were predominately about adaptation (36%), with topics from the “Other” category comprising of: coevolution (21%), sexual selection (21%), and reproductive isolation (14%). While explanations from iteration two covered a wider range of evolutionary concepts, with the highest levels of explanations focused on natural selection (33%) and adaptation (26%). Topics from the “Other” category comprised of: sexual selection (12%), coevolution (7%), reproductive isolation (7%), gene flow (5%), and miscellaneous topics (12%).

Table 2. Descriptive summary of student explanations to the question, *what evolutionary forces do you think were at play in this environment?*

	Iteration 1 (n=8)	Iteration 2 (n=45)
Number of explanations	14	43
Average word count	24 ($SD=14.58$)	33.28 ($SD=29.51$)
Average KI score	2.36 ($SD=0.75$)	2.72 ($SD=1.05$)

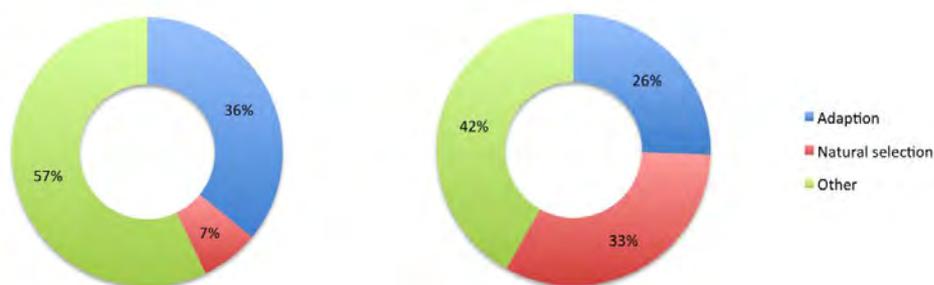
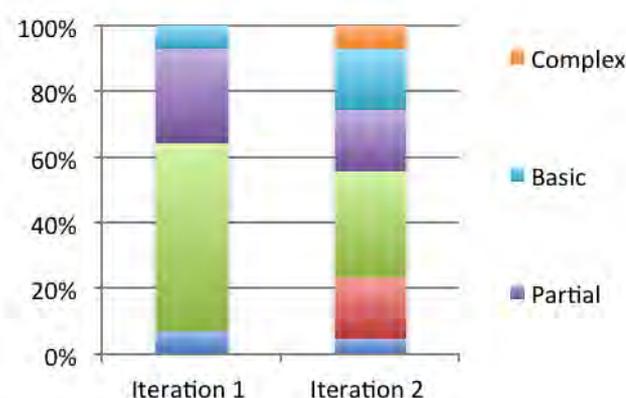


Figure 4. Distribution of evolutionary concepts that the explanations from the first (left) and second (right) iterations attempted to address.

The explanations were scored using a 0-5 Knowledge Integration (KI) scale that rewards valid scientific connections between concepts (Table 4; Linn & Elyon, 2011). The explanations from iteration two attained higher average scores ($M=2.72$, $SD=1.05$) than those from iteration one ($M=2.36$, $SD=0.75$), although no significant difference was found. In general, there was an increase in the complexity and sophistication of explanations from iteration one (34%) to iteration two (43%). Figure 5 displays the distribution of explanations based on their KI scores.

Table 4. KI rubric used to score student explanations. From Linn & Elyon, 2011.

Score	KI level	Description
0	No answer	
1	Off task	<ul style="list-style-type: none"> • Response is irrelevant or “I don’t know” • Student writes some text, but it does not answer the question being asked
2	Irrelevant/Incorrect	<ul style="list-style-type: none"> • Have relevant ideas but fail to recognize links between them • Make links between relevant and irrelevant ideas • Have incorrect/irrelevant ideas
3	Partial	<ul style="list-style-type: none"> • Have relevant ideas but do not fully elaborate links between them in a given context
4	Basic	<ul style="list-style-type: none"> • Elaborate a scientifically valid link between two ideas relevant to a given context
5	Complex	<ul style="list-style-type: none"> • Elaborate two or more scientifically valid links among ideas relevant to a given context

**Figure 5.** Distribution of student explanations' KI scores.

Discussion

From “Free Formed” to Structured Observations

When we first reviewed the observations from the first iteration, we noticed them to be rather basic. An example of an observation that focused on behavior was, “*There are two tapirs, one walking really slowly and one drinking from a shallow pool.*” Observations that focused on physical characteristics were also superficial, e.g., “*The fig wasp has purple wings, long antenna, and a striped body.*” To help promote deeper explanations, we designed scaffolding for iteration 2, in the form of more structured observations. For example, students to answer simply *yes* or *no* to the question, “*Is the organism present here?*” and then reflect more deeply on the larger patterns. Since they only observed their own assigned species, students relied on the work of their peers to understand the complete picture of how all the organisms evolved over time. Their answers were aggregated to the interactive white board at the front of the room (Figure 6) and reviewed in teams of four to six. With students providing structured observations, we were able to assess more easily whether they were paying attention to the correct pieces and seemed to yield positive results (with over 80% accuracy).

Increased Variation of Evolutionary Topics

At first glance, the student explanations from iteration one do not differ significantly from the explanations written from iteration two, given their comparable KI scores. However, improvements to the activity in iteration two may be demonstrated in the increased variation in the types of evolutionary concepts that students addressed. Figure 7 demonstrates the nature of explanations with a visual representation (i.e., Wordle - <http://www.wordle.net/>), where words with the highest frequency are given greater prominence. The nature of the explanations in iteration one tended to be about surface features of the species observed, while the explanations in iteration two focused more so on the processes of evolution.



Figure 6. Interactive cladogram created from the collective inputs of 16 students’ structured observations.



Figure 7. Wordle of student explanations from iteration (left). Words, such as camouflage (6x), curtain figs (6x), different (5x), and wasps (4x), were most frequently encountered. Wordle of student explanations from iteration two (right). Words such as evolved (19x), selection (14x), adapted (12x), species (11x) appeared most.

Current Progress & Future Directions

At the time of writing, audio and video analysis of student interactions are in progress. We expect that results will glean important insights about students' thinking behind their written explanations. We will continue to analyze students' biological understanding by coding the various elements, particularly from our second iteration, which was embedded within a much larger curricular sequence. From the early results presented here, we are already making progress in designing our next iteration. We understand the need to better address students' preconceptions about evolution, as well as the need to encourage increased complexity of student thinking. We will look for opportunities to incorporate these ideas into our designs in a more seamless and meaningful manner.

References

- Alters, B. J., & Nelson, C. E. (2002). Teaching evolution in higher education. *Evolution*, 56(10), 1891–1901.
- Anderson, R. D. (2007). Teaching the theory of evolution in social, intellectual, and pedagogical context. *Science Education*, 91(4), 664–677.
- 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.
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, 6: 271-315.
- Chinn, C. A., & Buckland, L. A. (2012). Model-based instruction: Fostering change in evolutionary conceptions and in epistemic practices. In K. S. Rosengren, S. K. Brem, E. M. Evans, & G. M. Sinatra (Eds.), *Evolution Challenges: Integrating Research and Practice in Teaching and Learning About Evolution*. Oxford University Press.
- Colella, V. (2000) Participatory simulations: Building collaborative understanding through immersive dynamic modeling, *Journal of the Learning Sciences*, 9(4), 471-500.
- Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, 323, 66–69.
- Demastes, S., Good, R., & Peebles, P. (1995). Students' conceptual ecologies and the process of conceptual change in evolution. *Science Education*, 79(6), 637–666.
- Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5-8.
- National Academies Press, 1998
- Kampourakis, K., & Zogza, V. (2008). Preliminary evolutionary explanations: A basic framework for conceptual change and explanatory coherence in evolution. *Science & Education*, 18(10), 1313–1340.
- Linn, M., & Eylon, B. (2011). *Science learning and instruction: Taking advantage of technology to promote knowledge integration*. New York: Routledge.
- Mayr, E. (2002). *What Evolution Is*. New York: Basic Books.
- Moher, T. (2008). Learning and participation in a persistent whole-classroom seismology simulation. *Proceedings of the International Conference of the Learning Sciences (ICLS) – Volume 2*, 82-90.
- 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.
- 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), 51-74.
- Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver, & D. Klahr (Eds). *Cognition and Instruction: Twenty-Five Years of Progress*, (pp. 263–305). Mahwah, NJ: Erlbaum.
- Resnick, M. (1996). Beyond the Centralized Mindset. *Journal of the Learning Sciences*, 5(1), 1-22.
- Rudolph, J. & Stewart, J. (1998). Evolution and the nature of science: On the historical discord and its implications for education. *Journal of Research in Science Teaching*, 35(10), 1069-1089.
- Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *The Journal of the Learning Sciences*, 12(1), 5-51.
- Slotta, J. D. (2010). Evolving the classrooms of the future: The interplay of pedagogy, technology and community. In K. Mäkitalo-Siegl, F. Kaplan, J. Zottmann, & F. Fischer (Eds.), *Classroom of the Future: Orchestrating collaborative spaces* (pp. 215-242). Rotterdam: Sense.
- Wilensky, U., & Stroup, W. (1999). Learning through participatory simulations: Network-based design for systems learning in classrooms. *Proceedings of the conference on Computer Support for Collaborative Learning*, Palo Alto, California.

Designing for Group Math Discourse

Rachel M. Magee, Drexel University, Philadelphia, PA, USA, rachelmagee@drexel.edu
Christopher M. Mascaro, Drexel University, Philadelphia, PA, USA, cmascaro@gmail.com
Gerry Stahl, Drexel University, Philadelphia, PA, USA, Gerry@GerryStahl.net

Abstract: We are developing a socio-technical system to support group cognition among math students in the form of significant mathematical discourse about dependencies in dynamic-geometry constructions. Analysis of a pilot trial in a typical early cycle of our design-based-research approach revealed barriers to group success from both software and mathematics issues, and demonstrated that participants “cycled” between these types of issues. We are responding by developing a curriculum to address the uncovered technical and cognitive issues. We present the findings of our pilot study and the curriculum design criteria that are emerging from continuing cycles of re-design, prototyping, testing and analysis.

Introduction

We are interested in promoting group cognition (Stahl, 2006) among math students learning high-school geometry by enhancing their ability to engage in significant mathematical discourse. Recent research on math learning points to the central role of language, enabling articulation and verbal reflection about mathematical relationships (Sfard, 2008; Stahl, 2008). Because we want to exploit the computational power of computers and the advantages of networking to support online collaboration incorporating math discourse, we necessarily face the dual design constraints of technical software development and social-practice scaffolding.

Increasingly, high school students are learning online, with home schooling, resources like the Khan Academy of math YouTube videos and virtual high schools. The problem with online learning is that the current models for this are often lacking in social interaction and collaborative learning. This is, of course the motivation for CSSL research and innovation.

Discourse is fundamentally a group process, so we want to provide support for small groups of students to engage together in math discourse. This is complicated in terms of both the discourse and the technology as there are multiple facets to “significant mathematical discourse” (Stahl, 2013d). Furthermore, we are interested in taking advantage of networked computers to allow groups of students to discuss math and to work on mathematical tasks together online. We want to supply computer support for their math work and computer recordings for maintaining persistence of their discourse—which raises technological barriers to students navigating the interface.

As a research project, we approach this task with the idea of combining VMT (Stahl, 2009; Stahl, Mantoan & Weimar, 2013)—a generic computer environment for collaborative learning by “virtual math teams”—with GeoGebra (www.geogebra.org)—a popular open-source application for dynamic geometry. This involves enhancing VMT and transforming GeoGebra from a single-user application to a multi-user client integrated into VMT. When developing a socio-technical system, in addition to the technical development we need to guide the group-cognitive work by providing helpful resources and scaffolding group practices.

To get a realistic sense of how groups of students will interact within the environment we are designing, we need to conduct pilot tests throughout our design process. In order to try out our system in naturalistic settings as part of our socio-technical, design-based-research approach, as well as to provide a basis for eventual deployment, we have developed relationships with two professional education schools, where we will eventually deploy our system with practicing math teachers.

In our preliminary stage, we have run informal pilot tests with available groups. Our findings showed that these students encountered significant problems due to a lack of preparation for using the technology and for engaging in the mathematics. As a result of the analysis of these sessions—as discussed below—we realized that we would have to carefully craft a curriculum, which the teachers could follow and then adapt for their own classrooms. This curriculum would need to incorporate not only math lessons, but also tutorials about the technological environment. We started to sketch out a curriculum based on existing best practices and theories. We were fortunate that the Common Core State Standards for Mathematics (CCSSM, 2011) had recently been released and adopted by most states in the US. This provided an up-to-date, research-based outline of content for a geometry course, which was widely accepted.

As we looked at results of the initial trials analyzed below, we realized many problems needed to be addressed. These involved design issues in extending VMT, in making GeoGebra multi-user, in supporting collaboration around the activities, in teaching the deep conceptual ideas in geometry, in taking advantage of computer-supported dynamic math and in promoting significant math discourse (Stahl, 2013d). We ran several cycles of additional trials within our research group and with available college students. In each cycle, we

revised the curriculum, revised the software, ran the trial and analyzed the behaviors. Generally, there were clear lessons from each trial, which led to the next cycle.

Gradually, a set of design criteria for the curriculum was formulated and evolved. In this paper, we report findings from the early session without curriculum to identify challenges faced by technologically adept individuals when attempting to engage in significant math discourse within the GeoGebra environment. Then we review some of the lessons for the technology and some of the aspects of the discourse that we believe are important. Based on these lessons, we are now developing a curriculum based around online, small-group activities. This paper discusses the criteria for the design of that curriculum, as it is emerging from testing of trial curriculum drafts.

From a socio-technical standpoint, the curriculum is central because it mediates between the people and the technology. It tells the people what activities they should be engaging in while communicating through and working within the technology. It also models for them how to talk about math. For an online course, in which there is no teacher present to orchestrate activities and interaction, the textual curriculum provides the major scripting of collaborative sessions and the primary scaffolding of the group cognition.

Relevant Literature

Our approach to online dynamic-geometry education is based on previous research by our own team and by others in the fields of groupware design, collaborative learning and mathematics education.

Dynamic-Mathematics Software

The research on dynamic-mathematics software—such as Geometer's Sketchpad, Cabri and GeoGebra—is new and limited. Much of it merely popularizes the availability and the novelty of the approach. However, there are some important studies of aspects such as the dynamic dragging of geometric objects and the implications of dynamic visualizations for student conceptions of proof. A recent review of 37 publications on dynamic mathematics summarized the research to date (Powell & Dicker, 2012). Dynamic geometry can be effective in improving student understanding of geometry through support for visualization and exploration. There is a trade-off between having students do their own constructions versus having them manipulate prepared constructions. While the construction process may deepen understanding, it takes much longer and can be distracting from curricular goals. The ability to manipulate constructions dynamically aides students in making conjectures, exploring them and understanding their significance, but it can be seen as a substitute for deductive proof and can lower student motivation to engage in rigorous proof procedures.

The utilization of dynamic-math environments by teachers has allowed them to extend traditional materials found in textbooks, allowing for better interaction with their students in both the classroom and through technological mediation (Hohenwarter, Preiner & Yi, 2007). These dynamic-math environments have been found to make mathematical tasks more efficient and allow for more interaction and application of the theoretical knowledge associated with the mathematical task (Laborde, 2001; Öner, 2008). This success of dynamic-math environments in the classroom setting is heavily influenced by the given tasks and the interaction with the instructor who is leading the exercise (deVilliers, 2004; Mariotti, 2001; Sanchez & Sacristan, 2003). The research in dynamic math is limited to specific pedagogical approaches and needs to be developed further. In particular, previous studies focus on individual learning. This is at least in part because until now dynamic-math applications have been designed for single users. Another weakness in the literature is the lack of focus on *dependencies*, which we feel are central to understanding dynamic geometry (Stahl, 2013c).

Online Math Collaboration

The ability for students to co-construct knowledge using technology together has been studied for decades. Depending on the context, the students and teachers play different roles (Jonassen, Peck & Wilson, 1999). While technology adoption in the classroom has met with varying levels of success, using small groups for learning and co-constructing knowledge has been illustrated to be productive through all levels of education (Springer, Stanne & Donovan, 1999).

Researchers have approached studying math discourse and cognition in face-to-face media through the utilization of technology (Dion, Jank & Rutt, 2011). Research on Group Scribbles use in a primary science classroom in Singapore illustrates a transitional stage between the physical classroom and a strictly online context (Chen & Looi, 2011). The Group Scribbles environment provides similar capabilities to the VMT environment, but the interaction occurs in a classroom through tablets. The students' interactions are technologically mediated, and the teacher in the classroom provides physical mediation, allowing for technology problems to be quickly overcome so participants may focus on the problem at hand. In a series of tasks carried out using Group Scribbles, it was found that students had more agency and were given more participation opportunities compared to traditional approaches. This was found to particularly benefit passive students (Chen, Looi & Ng, 2009).

Technology has also taken the place of moderating a learning environment in an effort to facilitate more discourse and reduce direct teacher involvement in student problem solving. In an attempt to automate the support of group math cognition in the VMT environment, research has been initiated to understand how conversational agents could be used (Cui et al., 2009). These agents are used to encourage academic discourse and accountable talk (Michaels, O'Connor & Resnick, 2008), but have been met with only limited success so far (Stahl, 2013a; Stahl et al., 2010).

Understanding the technological environment of the student and how this contributes to successful or unsuccessful learning is integral to the analysis of the learning and the design of an online system (Suthers & Medina, 2010). As evidenced by prior research (Valentine, 2002) and our findings below, technology use in a learning scenario can harm the experience of students, hindering communication as much as facilitating it. Understanding the extent of barriers and modes of facilitation of math discourse in a dynamic-math environment is still limited.

Math Discourse

The theory of math learning through participation in math discourse (Sfard, 2008) specifies important mathematical discourse moves, such as encapsulation, reification, naming, routines, deeds, explorations and rituals—all defined, systematized and passed down through the community, culture, tools, procedures, language and traditions of mathematics. These interactional resources can traverse levels between individual learning, group cognition and community knowledge (Stahl, 2012; 2013b; Stahl & Öner, 2013). The theory of accountable talk (Michaels, O'Connor & Resnick, 2008; Resnick, O'Connor & Michaels, 2007) specifies discourse moves that promote accountability to the group, to standards of math reasoning and to the characteristics of the math objects. Speaking meaningfully in math discourse “implies that responses are conceptually based, conclusions are supported by a mathematical argument and explanations include reference to the quantities in the problem context as opposed to merely describing the procedures and calculations used to determine the answer” (Clark, Moore & Carlson, 2008, p. 298). Socio-mathematical norms include what counts as an acceptable, justifiable, easy, clear, different, efficient, elegant and sophisticated explanation (Yackel, 1995; Yackel & Cobb, 1996). Mathematical practices emerge from interaction, are taken up by participants and are applied repeatedly (Medina, Suthers & Vatrappu, 2009). Though Sinclair and Yurita (2008) study of how dynamic geometry changes discourse began the process, research into the nature of mathematical discourse in a collaborative dynamic-mathematics environment has yet to be conducted.

While the importance of collaborative learning for online education may be obvious to CSCL researchers (Stahl, Koschmann & Suthers, 2006) and its possible advantages have been well documented in cooperative-learning (Johnson & Johnson, 1989; Slavin, 1980) and CSCL research for decades (Sawyer, 2006), support for collaboration is still not always designed into new educational platforms. For instance, the latest hot approach to university instruction—massive open online courses or MOOCs—are generally based on the lecture paradigm, in which students passively watch talking-head videos of famous professors and are not given any sanctioned opportunities for interaction. Similarly, the acclaimed Khan Academy offers thousands of YouTube videos explaining detailed topics in school mathematics, but students have no support for interactively exploring the topics themselves or discussing them with peers. These technological opportunities are generally not designed to incorporate constructivist learning principles (Bransford, Brown & Cocking, 1999).

Method

In Fall 2011, we examined four one-hour-long chat logs from information-science graduate students taking a course on CSCL using the VMT environment. In these chats, the groups met online and attempted to solve a geometry problem within the GeoGebra environment. The students had used the VMT environment to perform collaborative writing exercises in previous weeks, but had not previously used GeoGebra. These students were enrolled in majors related to technology, suggesting that they were engaged rather than nervous about technology use. As part of the exercise, there was no explicit introduction to the GeoGebra tool or further instructions other than the assigned problem.

We were interested in analyzing these groups' interactions and their strategies for navigating a new online collaborative environment. Each log was examined independently using a thematic analysis approach that revealed themes that were typical stages of conversation. These stages include: social niceties, problem identification, technical discourse, math discourse, design suggestions and future planning. While these are separate stages of conversation, we found that each group moved back and forth between *technological* and *mathematical* discourse, behavior we termed “cycling.”

We examined the logs using our initial categories as a guide to further examine this process of cycling. In our subsequent analysis, we identified the cyclical behavior triggered by individual statements distinctly indicating technical issues (involving software usage issues or software problems) versus mathematical issues and discourse (involving attempts to understand, represent and solve the geometry problem). By examining the chat logs we are able to observe phases of group interaction, how technology affects each phase and how the

technology can both facilitate and inhibit successful completion of the task in an online environment new to group members.

Findings

Analysis of the group chat logs illustrates the presence of a variety of stages of conversation by the members of the groups in the context of the problem-solving task. Each of the chats begins with an orientation, including the exchange of social niceties and resolution of unrelated issues, typically lasting two to three lines per group member.

Following the orientation stage, the groups identified the problem by either explaining it in the chat to the other group members or by referring to the posted problem in another tab. This typically involved a statement to orient the group:

Quick summary – we have to work thru the problem (see topic). Summarize the process in the Summary tab and post a few sentences on the wiki too. We good? (Group 1, line 16)

This quote illustrates some of the important characteristics of this type of focusing statement, including a description of what the “speaker” is going to do with the statement, instruction and then a leading question to ensure the team is on the same page.

Table 1 illustrates the different stages identified in the chat logs of the four groups and the different places in the discussion in which math discourse began. This varied for each group, and even when groups did not start with technical issues, they arose very quickly.

Table 1. Stage identification of each of the groups

	Group 1	Group 2	Group 3	Group 4
Opening Stages	Orientation; Problem Identification			
Intermediate Stages	Role Assignment	Math Discourse	Technical Issues	Technical issues
	Technical Issues	Technical Issues	Math Discourse	Role Assignment
	Math Discourse	Design Suggestions	Math Confusion	Math Discourse
			Technical Issues	Use of Alternative Tools
		Technical Confusion		
Concluding Stages	Summarization of task/experience; social niceties; next steps			

Once Group 3 reached math discourse, they experienced confusion about the mathematical concepts, further compounded by technical issues with the tool that further confused the participants in the group and degraded the quality of math discourse:

I'm trying to figure out how to delete this line... I kind of messed up... do you still see a line on the screen? (Group 3, line 24-25)

This quote highlights a number of issues that were common across multiple groups: not knowing how to delete an object (an option expected by participants), and group members being unsure that they were looking at the same objects as their fellow group members.

Both Group 1 and Group 4 achieved significant math discourse as each of the team members attempted to solve the problem, but did so with the help of outside tools. Group 1 used PowerPoint. Group 4 experienced confusion because the VMT environment did not display the same screen to all group members, so one emailed a screenshot to share the solution. This indicates that use of familiar tools or tools that work intuitively enables groups to more quickly reach effective math discourse that achieves a solution.

Overcoming Technological Barriers

While the technological tool—multi-user GeoGebra—provided many opportunities and options, it often served as a barrier for users. Barriers could be as simple as not being able to undo an action. However, even simple barriers stopped the groups from engaging in fluid math discourse and sometimes even went unresolved as the individuals found ways to work around issues. One example of this is the issue of not being able to easily rename an object. The mathematical problem these groups were attempting involved making an angle ABC. Groups began by playing with the system by adding objects to their GeoGebra screen. However, each group discovered that they were unable to simply rename the points on their screen, and the names they needed (A, B and C) were already in use by the system, though the objects they developed later in the process were better suited to solving the problem. This meant that their refined objects were confusingly named (for example, J, K and L), making math discourse about the objects in relation to the problem statement more complicated:

One thing we can state is how the lettering got messed up... I think that is helping to confuse us. (Group 3, lines 58-59)

Each group experienced this issue; because of their lack of familiarity with the system, none were able to fix it.

Other barriers were easier for the groups to work around creatively. When the tool put up barriers, these groups were quick to try to work around the resulting issues, employing their understanding of other technologies to inform their decision in the VMT environment or solve the problem and move forward to math discourse. While it was unsuccessful, in the above barrier, the group attempted to rename an object when it had difficulty, which is a common solution in other tools. Additionally, many groups wished for an undo option, a common affordance in other tools:

Is there an undo function... not that I could find. That would be nice. (Group 1, lines 97-99)

Because of the nature of the work, as groups overcame tool issues and moved into math discourse, new mathematical objectives (e.g., renaming a point, adding a ray) resulted in a return to the tool and often the discovery of a new technological barrier. Even in the face of such issues with the tool, multiple groups managed to achieve effective math discourse that led to solutions. Each successive cycle of math discourse and tool use also led to difficulty with the mathematical concepts at hand, which we will now discuss.

Discourse about Math Difficulties

The goal of these chats for the students was to experience a new tool, but also to achieve math discourse around the visualization and solution of a geometry problem. Reaching math discourse proved to take some time for multiple groups, despite the fact that they were actively pursuing this goal, often within the first few lines of chat. Typically, the first approach involved developing a shared understanding of what the mathematical problem was, which we termed “problem identification” in our stage-identification process. Groups quickly entered into discussion of technological issues with the tool, but had difficulty returning to the larger goal of mathematical discourse.

Participants often employed a question structure to encourage a return to math discourse, and usually included words like “okay,” “well,” or “so” to bridge from the previous topic, which was typically a technological issue. In Group 4, one participant states:

Ok, we are on the same page now... we need a point in the middle. (line 179-180)

In an attempt to move past the technological barrier of not being able to effectively rename objects and establish common ground among the participants, one participant transitioned with:

Well, anyway, do we all at least see i, j, k? (Group 4, line 83)

In addition to bridging words, participants also employed explicit questions to reorient the group, for example:

Can i start by drawing two lines to create an angle? (Group 3, line 22)

These structures serve to call attention to a reorientation, and to give other participants the opportunity to request a pause in that reorientation to ensure they share understanding with the rest of the group.

Reorienting questions also served to highlight an understanding gap, pulling the group back into math discussion to provide an explanation or confirm an understanding. One example of this math-question reorientation comes from Group 2:

If you try to construct a line EF trying to connect AB and BC, wouldn't that mean A=C. (line 94)

The use of reorienting statements rotates through group members, indicating that it was not always the same participant to return the group to math discourse. Talking about technological issues could quickly grab the attention of the group, but these reorienting statements were effective at refocusing the groups' attention on the mathematical issues. When groups returned to this higher level of math discourse, there were a variety of approaches employed by individuals. Multiple participants displayed something akin to math anxiety, highlighting their lack of experience or inability:

I haven't done geometry in a long time... I'll need the hints. (Group 1, line 18)

Often, members of the group shared in their confusion, as evidenced by Group 3's experience with making the decision to look at the hints during a series of math discourse. The group looked at the hints as a whole, but each member admitted to being more confused after doing so, imagining that it could be their unfamiliarity with math causing the issue:

I'm not sure if its cause I haven't done these types of problems in a while or the hints just aren't that good. (Group 3, line 95)

However, Group 1 and Group 4 were able to achieve math discourse and a solution, notably, with the use of familiar outside technologies.

Cycles of Problems

The analysis of the pilot trial revealed cycles of problems, with the groups having to go back and forth between confronting technical problems with the software and cognitive problems with the mathematics. The cyclic nature of the alternation between technical and mathematical difficulties may have been an artifact of the task and the preliminary state of the software prototype. Though the task was to work on a geometry construction, within the online environment, software problems intervened and distracted the group. Groups tried to quickly get around the technical problems and cycle back to the math. There, they found themselves poorly prepared to tackle a geometry problem. Both the technical and the cognitive problems were consequences of the situation of

pilot-test participants in a design-based-research project. The socio-technical goal of the project was still in the distant future and the necessary supports for the participants were not yet in place. Thus, it is not a surprise that the subjects met with many serious difficulties. The point is to learn from the pilot trial: what are the most important social and technical features to be developed next?

Discussion

The experiences of the groups highlight interesting aspects of group-cognitive processes and how tool and math skills can hinder the ability to solve the problem by otherwise competent users. Clearly, while math discourse was a goal of each group, it proved difficult to achieve in the face of tool issues and feelings of math anxiety. When faced with a technical issue, the individuals blamed the tool for the inability to solve the problem, because they felt they were technically competent in general:

I'm an IT consultant and have to deal with various software programs meaning I'm familiar with how software should be designed and navigating my way around...this was definitely tough. (Group 4, lines 310-313)

On the other hand, when faced with a mathematical concept that they were not familiar with, members of the groups blamed themselves for not being mathematically focused:

My High School Math teachers are furious with me right now I can feel it. (Group 3, line 96)

This dichotomy between technical ability and mathematical inability was identified in each log. While this is an interesting case in our specific dataset pertaining to mathematically oriented online-learning contexts, we suspect that this phenomena may be evident in other collaborative-learning situations. Working to learn both content and the technology used to deliver that content can be overwhelming and may distract from the conceptual intent of the lesson. These difficulties are evident in our analysis as triggers of cycling and may be applicable to many technologically mediated learning situations. Because of these identified issues, it is important to build technological familiarity into any educational groupware environment to overcome technological issues early in the process. We find that in the face of tool adversity individuals defaulted to tools they were comfortable with such as PowerPoint, paper/pencil or screenshot/email. The use of familiar tools allowed the members of the groups to focus on the actual math discourse and problem solving, and isolate the effects of the tool on their productivity.

One of the most striking elements of our analysis is the concept of *cycling* in the group process between *tool* issues and *math* discourse. There was a salient presence of software functionality issues that when coupled with gaps in knowledge derailed mathematical discourse. This derailment and the students' interest in getting back on task led to cycling. Though each group experienced cycling, the groups that were most successful were able to quickly manage technological barriers and return to math discourse for the majority of their chats. We speculate that problems will exist in many groupware situations, including math learning, in which there are gaps in ability to manipulate the technology used for the learning. We believe that these findings may be transferable to other environments and contexts. As highlighted by one of the participants,

The issue with our first attempts was the usability of the tools – and lack of familiarity of the capabilities available within GeoGebra (Group 1, line 109)

An increase in familiarity with the system may reduce cycling; however, further research into groups learning a system is required to determine how this might manifest under different circumstances.

In addition to our analytical findings, each of the groups had recommendations for ways to improve the technology and the process of group math problem solving in the VMT-with-GeoGebra environment. These ranged from calls for an undo option to hopes for a primer or tutorial to alleviate some technological issues.

Curriculum Design Criteria

In response to the analysis of the GeoGebra use sessions, we have been drafting a set of dynamic-geometry curricular activities, interspersed with tutorials of the technology features. Curriculum activities have been designed to promote collaborative learning, particularly as exhibited in significant mathematical discourse about geometry. Collaborative learning involves a subtle interplay of processes at the individual, small-group and classroom levels of engagement, cognition and reflection. Accordingly, the activities are structured with sections for individual work, small-group collaboration and whole-class discussion. It is hoped that this mixture will enhance motivation, extend attention and spread understanding.

The goal of our set of activities is to improve the following skills in math teachers and students:

1. To engage in significant mathematical discourse; to collaborate on and discuss mathematical activities in supportive small online groups.
2. To collaboratively explore mathematical phenomena and dependencies; to make mathematical phenomena visual in multiple representations; and to vary their parameters.
3. To construct mathematical diagrams—understanding and exploring their structural dependencies.
4. To notice, wonder about and form conjectures about mathematical relationships; to justify, explain and prove mathematical findings.

5. To understand core concepts, relationships, theorems and constructions of basic high-school geometry. In other words, the activities seek a productive synthesis of the five areas of: discourse, visualization, construction and argumentation skills applied in the domain of beginning geometry. The set of activities is designed to provide an educational experience in basic geometry to math teachers and students, taking them from a possibly novice level to a more skilled level, from which they can proceed more effectively without such designed, scaffolded activities. By providing activities on different levels for each of the dimensions, we hope to help math teachers and students to increase their relevant skills – in different ways for different people.

Conclusion

Our focus has centered increasingly on facilitating and supporting lessons involving geometric dependencies (Stahl, 2013c). GeoGebra allows one to construct systems of inter-dependent geometric objects. Students have to learn how to think in terms of these dependencies. They can learn through visualizations, manipulations, constructions and verbal articulations. These can all be modeled and these skills can be developed gradually; our pilot study indicates that for successful math discourse to be achieved, supporting these skills must be an explicit priority of the socio-technical system. We are now drafting and piloting versions of curricular activities designed to develop significant mathematical discourse focused on dependencies among geometric objects (Stahl & Öner, 2013). Concomitantly, we are implementing software support for teachers and students to explore the dependencies and assembling materials for professional development to prepare teachers to enact this curriculum with their students (Stahl, 2013d).

Our design work is guided by socio-technical implications of continuing pilot studies as the technology and pedagogy of our project co-evolve. We are countering the problems that caused negative cycling of technical and cognitive distractions by improving the software and testing the curriculum. The curriculum integrates tutorials about using the VMT and GeoGebra interfaces with carefully structured sequences of dynamic-geometry activities for virtual math teams. The activities systematically build up the background knowledge, group practices and problem-solving orientation needed for engaging in mathematical discourse.

References

- Bransford, J., Brown, A., & Cocking, R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Research Council. Web: <http://books.nap.edu/html/howpeople1/>.
- CCSSI. (2011). High school -- geometry. In Common Core State Standards Initiative (Ed.), *Common core state standards for mathematics*. (pp. 74-78)
- Chen, W., & Looi, C.-K. (2011). Active classroom participation in a Group Scribbles primary science classroom. *British Journal of Educational Technology*. 42(4), 676-686.
- Chen, W., Looi, C.-K., & Ng, F.-K. (2009). *Students' active learning leveraged by interactive technologies: Group Scribbles in a primary science classroom*. Paper presented at the AERA. San Diego, CA.
- Clark, P. G., Moore, K. C., & Carlson, M. P. (2008). Documenting the emergence of “speaking with meaning” as a sociomathematical norm in professional learning community discourse. *Journal of Mathematical Behavior*. 27, 297–310.
- Cui, Y., Kumar, R., Chaudhuri, S., Gweon, G., & Rosé, C. P. (2009). Helping agents in VMT. In G. Stahl (Ed.), *Studying virtual math teams*. (ch. 19, pp. 335-354). New York, NY: Springer.
- deVilliers, M. (2004). Using dynamic geometry to expand mathematics teachers' understanding of proof. *International Journal of Mathematics Education in Science & Technology*. 35(4), 703-724.
- Dion, L., Jank, J., & Rutt, N. (2011). *Computer monitored problem solving dialogues*. Paper presented at the Research Experience for Undergraduates. Bloomington, Indiana.
- Hohenwarter, M., Preiner, J., & Yi, T. (2007). *Incorporating GeoGebra into teaching mathematics at the college level*. Paper presented at the ICTCM. Boston, MA.
- Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Edina, MN: Interaction Book Company.
- Jonassen, D. H., Peck, K. L., & Wilson, B. G. (1999). *Learning with technology: A constructivist perspective*. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Laborde, C. (2001). Integration of technology in the design of geometry tasks with Cabri-geometry. *International Journal of Computers for Mathematical Learning*. 6(3), 283-317.
- Mariotti, M. A. (2001). Justifying and proving in the Cabri environment. *International Journal of Computers for Mathematical Learning*. 6(3), 257-281.
- Medina, R., Suthers, D. D., & Vatrappu, R. (2009). Representational practices in VMT. In G. Stahl (Ed.), *Studying virtual math teams*. (ch. 10, pp. 185-205). New York, NY: Springer.
- Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in the Philosophy of Education*. 27(4), 283-297.

- Öner, D. (2008). Supporting students' participation in authentic proof activities in computer supported collaborative learning (CSCL) environments. *International Journal of Computer-Supported Collaborative Learning*. 3(3), 343-359.
- Powell, A. B., & Dicker, L. (2012). *Toward collaborative learning with dynamic geometry environments*. Paper presented at the 12th International Conference on Mathematical Education. Seoul, Korea.
- Resnick, L., O'Connor, C., & Michaels, S. (2007). *Classroom discourse, mathematical rigor, and student reasoning: An accountable talk literature review*: LearnLab, Pittsburgh Science of Learning Center.
- Sanchez, E., & Sacristan, A. I. (2003). *Influential aspects of dynamic geometry activities in the construction of proofs*. Paper presented at the 27th PME international Conference. Proceedings pp. 111-118.
- Sawyer, R. K. (Ed.). (2006). *Cambridge handbook of the learning sciences*. Cambridge, UK: Cambridge University Press.
- Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses and mathematizing*. Cambridge, UK: Cambridge University Press.
- Sinclair, N., & Yurita, V. (2008). To be or to become: How dynamic geometry changes discourse. *Research in Mathematics Education*. 10(2), 135-150.
- Slavin, R. (1980). Cooperative learning. *Review of Educational Research*. 50(2), 315-342.
- Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*. 69(1), 21-51.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.
- Stahl, G. (2008). Book review: Exploring thinking as communicating in CSCL. *International Journal of Computer-Supported Collaborative Learning*. 3(3), 361-368.
- Stahl, G. (2009). *Studying virtual math teams*. New York, NY: Springer.
- Stahl, G. (2012). Traversing planes of learning. *International Journal of Computer-Supported Collaborative Learning*. 7(4), 467-473.
- Stahl, G. (2013a). Interaction analysis of a biology chat. In D. Suthers, K. Lund, C. P. Rosé & N. Law (Eds.), *Productive multivocality*. New York, NY: Springer. Web: <http://GerryStahl.net/pub/multivocal.pdf>.
- Stahl, G. (2013b). Learning across levels. *International Journal of Computer-Supported Collaborative Learning*. 8(1), 1-12.
- Stahl, G. (2013c). Poster: Discovering dependencies: A case study of collaborative dynamic mathematics. Presented at the International Conference of Computer-Supported Collaborative Learning (CSCL 2013), Madison, WI. Web: <http://GerryStahl.net/pub/cscl2013dependencies.pdf>.
- Stahl, G. (2013d). *Translating Euclid: Creating a human-centered mathematics*: Morgan & Claypool Publishers. Web: <http://GerryStahl.net/pub/translating.pdf>.
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences*. (pp. 409-426). Cambridge, UK: Cambridge University Press. Web: <http://GerryStahl.net/elibrary/global>.
- Stahl, G., Mantoan, A., & Weimar, S. (2013). Demo: Collaborative dynamic mathematics in virtual math teams. Presented at the International Conference of Computer-Supported Collaborative Learning (CSCL 2013), Madison, WI. Web: <http://GerryStahl.net/pub/cscl2013demo.pdf>.
- Stahl, G., & Öner, D. (2013). *Resources for connecting levels of learning*. Paper presented at the International Conference of Computer-Supported Collaborative Learning (CSCL 2013). Madison, WI. Web: <http://GerryStahl.net/pub/cscl2013resources.pdf>.
- Stahl, G., Rosé, C. P., O'Hara, K., & Powell, A. B. (2010). *Supporting group math cognition in virtual math teams with software conversational agents*. Paper presented at the First North American GeoGebra Conference. Ithaca, NY. Web: <http://GerryStahl.net/pub/geogebra2010a.pdf>.
- Suthers, D., & Medina, R. (2010). *The temporal development of representational practices: Implications for theory and analysis of situated learning*. Paper presented at the Hawaii International Conference on System Sciences.
- Valentine, D. (2002). Distance learning: Promises, problems and possibilities. *Online Journal of Distance Learning Administration*. 5(3)
- Yackel, E. (1995). Children's talk in inquiry mathematics classrooms. In P. C. H. Bowersfeld (Ed.), *The emergence of mathematical meaning: Interaction in classroom cultures*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*. 27(4), 458-477.

Acknowledgments

This work is supported by the NSF Graduate Research Fellowship Program under Grant No. 2011121873.

MTClassroom and MTDashboard: Supporting Analysis of Teacher Attention in an Orchestrated Multi-tabletop Classroom

Roberto Martinez-Maldonado, Judy Kay, Kalina Yacef, School of Information Technologies, Marie-Theresa Edbauer, Business School, The University of Sydney, NSW 2006, Australia.

Yannis Dimitriadis, School of Telecommunications Engineering,
University of Valladolid, Paseo Belén 15, Spain.

{roberto,judy,kalina}@it.usyd.edu.au, medb1925@uni.sydney.edu.au, yannis@tel.uva.es

Abstract: In spite of the substantial progress in CSCL, there is still some distance between the promise of educational technology for classroom learning and what is readily achieved. Emerging tabletop devices can offer new means to enhance teachers' classroom control and awareness. These technologies can help them orchestrate activities, and capture, analyse and visualise students' collaborative interactions. This paper presents MTClassroom and MTDashboard, that were designed, deployed and tested to support the teacher in orchestrating collaborative learning activities at an authentic classroom. MTClassroom is an enriched multi-tabletop environment that captures aspects of students' activity as they work in small groups. MTDashboard is an orchestration tool displayed at a handheld device, giving the teacher control over classroom activities and providing 'real-time' indicators of participation and task progress of each group. We analysed teacher's attention by triangulating quantitative evidence captured by our environment with qualitative observations and teacher's perceptions. We investigated the affordances of our environment and the impact of the information provided to the teacher through the MTDashboard. The contribution of this paper is the novel approach for providing teachers with key indicators of small-group collaboration in the classroom and analysing their impact on teachers' attention to help them manage their time more effectively.

Introduction and Related Work

Research on Computer-Supported Collaborative Learning (CSCL) has demonstrated that small group collaboration can activate particular learning mechanisms and that educational technology resources can be used to mediate and facilitate such collaborative activities (Roschelle et al., 1995; Stahl, 2006). In spite of this substantial progress in research and practice, there is still some distance between the promise of technology and what has actually been delivered in most classrooms. This issue is particularly important for CSCL due to over-generalisation from a small pilot study's findings and over-expectation of new technology (Dillenbourg et al., 2011a). Cuban et al. (2001) argued that the role of teachers is critical to determine the success of deploying technological innovations in the classroom. This points to the need to consider the potentially key role for creating new mechanisms that can help *teachers* design and orchestrate learning activities (Dillenbourg et al., 2010; Prieto et al., 2011), so that they can successfully use emerging technologies in the classroom.

An obstacle in using personal computers in the classroom is that these tend to make it more difficult to promote face-to-face collaboration due to their small display and single input (Morgan et al., 2009). By contrast, emerging shared devices, such as multi-touch tabletops, offer a large enriched interface that learners can use simultaneously to create artefacts. They also offer access to digital content while students collaborate and negotiate understanding face-to-face with equal opportunities of participation (Dillenbourg et al., 2011a). Our work aims to tackle the issues described above by providing a suite of hardware and software tools for (i) enabling students to work in small groups and build virtual artefacts in the form of concept maps that represent their shared understanding (Figure 1, right), and (ii) enabling teachers to orchestrate the learning activities and teach curriculum content. We present MTClassroom (Figure 1, centre) and MTDashboard (Figure 1, left), which were both deployed and tested in authentic classroom sessions. MTClassroom is an enriched multi-tabletop classroom that captures aspects of students' learning and interaction processes as they work in small groups. MTDashboard on an orchestration tool displayed at a handheld device that allows a teacher to control classroom



Figure 1. MTDashboard (Left), MTClassroom (Centre) and the Concept Mapping application (Right).

activities and obtain live visual indicators of collaboration or progress of each group. The main contribution of this paper is the presentation of an approach that provides indicators of small-groups' performance on the teacher's dashboard, and our study of its impact on the teacher's decisions about the groups needing attention.

To date, full class sets of interactive tabletops have been studied in research contexts, rather than authentic learning environments. One important project that explored the use of tabletops in the classroom is SynergyNet (Mercier et al., 2012). This was a multi-tabletop environment used to investigate the quality of school children's collaboration and the ways teachers can interact with their system. Another similar environment was presented by Do-Lenh (2012), which could track command cards for the teacher to orchestrate the tables and also showed task progress indicators at a wall display that all the class could see. Both environments explored ways that a teacher can use these devices for classroom orchestration, in terms of collaboration and usability, respectively. A third example was provided by a teacher's dashboard proposed by Martinez-Maldonado et al. (2012b), who evaluated a system that offered visual indicators of group work at each tabletop to help teachers decide which group needed more attention over the duration of a class. However, in all this previous work, the studies were not linked to authentic curricula; nor were they prepared by the teacher.

The work we present in this paper builds on principles of classroom orchestration (Dillenbourg et al., 2011b), specifically on the dimensions of regulation and awareness (Prieto et al., 2011). MTClassroom provides an environment that captures live information about each learner collaborating at the classroom and MTDashboard is the interface that provides control functions and indicators enabling the teacher to be aware of each group's progress and activity. Our work goes beyond previous research by showing how the captured data can be used by the teacher in two ways: in class for light-weight indicators of students' progress; and after class, to analyse the ways they allocated their attention between the student groups.

Design of our Educational Technology

The main motivation for designing MTClassroom and MTDashboard is that, as the use of technology in and out the classroom is spreading, large amounts of learner data can be captured and summarised. These summaries of data can be exploited to show information that might otherwise not be easily available. This can be provided to teachers so that they can better decide which students may need timely interventions (Bull et al., 2012) or for later reflection on how their classroom attention was divided. Interactive tabletops are devices that have the potential to support knowledge co-construction in small teams (Dillenbourg et al., 2011a) and also to capture aspects of learners collaborative interactions (Martinez-Maldonado et al., 2012c). Next, we describe the principles of classroom orchestration and awareness that drove the design of the educational technology presented in this paper and then, our learning environments. These principles are as follows:

a) To support the role of the teacher as the main actor in classroom orchestration. The design of the system should primarily focus on providing services to assist teachers' actions and awareness in the classroom (Dillenbourg et al., 2010).

b) To support coordination of planned learning activities. The tools should support the enactment of the activities designed by the teacher, so that the learning objectives can be achieved (Prieto et al., 2011).

c) To support classroom regulation and management. The system should provide the teacher with functions to manage and adapt, to some extent, the macro script of the classroom activity (Dillenbourg et al., 2011b). We also highlight the importance of after-class analysis of the data that can be captured during the learning activities for reflection and evaluation.

d) To provide "light-weight" indicators about learners' progress. The system should be able to automatically capture small-groups' interactions data and present this information to the teacher to enhance their awareness and direct their attention (Bull et al., 2012; Martinez-Maldonado et al., 2012a).

MTClassroom: Multi-Interactive Tabletop Classroom

The MTClassroom is composed of a number of interconnected multi-touch interactive tabletops (4 were used in our study). Each tabletop consists of a 26 inch PQLabs multi-touch layer placed over a high-definition display of the same size. Each tabletop is enriched with an over-head depth sensor that detects the student who is touching the interactive surface at any time (Figure 2, left). In this way, the *host applications* running at the tabletops recognise and log differentiated actions performed by each student. From the *teacher's perspective*, MTClassroom offers functionalities for orchestrating the tabletops through a controller dashboard that allows teachers to send commands to the host applications to trigger actions such as blocking the touch input or moving to the next learning phase. A full description of the design of this tool is provided in the next section. Additionally, the system incorporates a connected wall projector that the teacher can use to display the artefact being created at a determined tabletop to lead reflection at classroom level (Figure 2, right).

MTClassroom can run different learning applications. In this study, the classroom activity consisted of the elaboration of concept maps. Concept mapping is an activity that encourages meaningful learning and, when maps are constructed in small groups, can foster externalisation and negotiation of diverse perspectives (Novak, 1995; Stahl, 2006). This tabletop concept mapping application (Martinez-Maldonado et al., 2010) permits



Figure 2. MTClassroom in action (Left). Reflection driven by the teacher using the wall display (Right)

students to have access to a list of suggested concepts and linking words, or type their own words, in order to build a concept map that answers a question asked by the teacher. Prior to the classroom activity, the teacher uses a desktop concept mapping editor to create the list of suggested concepts and linking words, and generate a *Master Concept Map* with the *crucial or relevant* concepts and links that learners must include in their maps, as well as other *relevant* ones that might be expected.

From a *data capture perspective*, the system automatically differentiates students’ actions at the tabletop according to their seating position. The logging system of each tabletop records all actions to a central repository that can be accessed in real time to generate indicators of group activity to be presented to the teacher. Additionally, *observation consoles* can be directly connected to the repository to capture synchronised qualitative data. In our study, two different observers submitted standardised annotations of the teacher’s attention and interventions. More details about these are described in the next sections.

MTDashboard



Figure 3. MTDashboard: condition 1 (Left), being used in the classroom (Centre) and condition 2 (Right)

The MTDashboard is a multi-platform teacher’s tool that contains both controlling and awareness components (Figure 3, right). In this study, the dashboard was displayed at a handheld tablet device that the teacher carried while walking around the classroom to monitor student progress (Figure 3, centre). The design of this dashboard was driven by the requirements specified by the teacher. The design was also based on principles of classroom orchestration of regulation and awareness (Dillenbourg et al., 2010; Prieto et al., 2011) and inspired by similar technologies applied in related work (Do-Lenh, 2012; Mercier et al., 2012). Figure 4 shows details of the MTDashboard interface that includes the following components. A) *General functions*, commands that the teacher can use with any tabletop. These are, “Start” (Figure 4, A1) and “Finish” (A4) commands, to explicitly mark the boundaries of the activity; a “Send message” (A6) command, so the teacher can send text reminders to all the tables about, for example, the time left for the activity; “Block” (A2) to freeze the table when the teacher wants students’ attention and “Unblock” (A3) commands for the teacher to get students’ attention when needed; and a “Reset” (A5) command to clean up the tabletop interfaces and making them ready for students in next

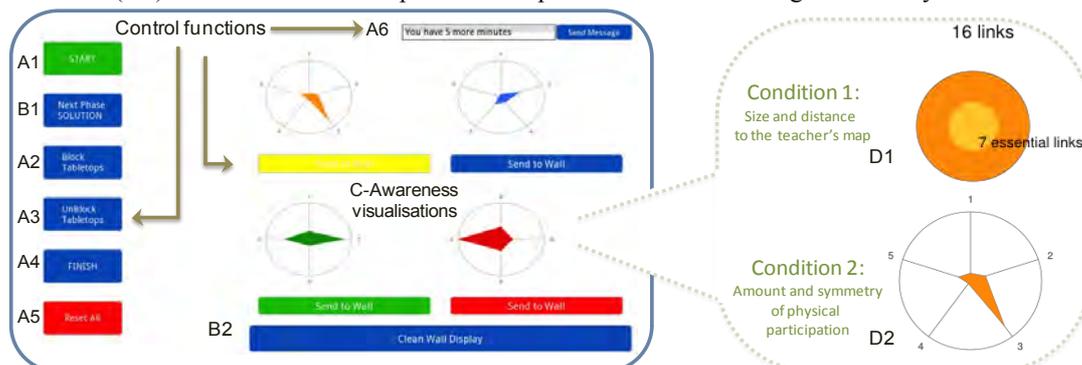


Figure 4. User interface of the MTDashboard (annotated to show Control functions A1-6, B1-2, C, D1-2).

tutorial. B) *Configurable functions*, may be applicable in various activities but their meaning depends on the macro-script definition. These include the “Jump to the next phase” (B1) and “Send to the wall” (B2) commands. Figure 3 (left) shows that in our case, this latter shows a concept map of one tabletop in the wall display. And finally, C) *Awareness visualisations*, which can show key information about each group progress, participation or other indicators that may be coupled to the domain (D1-2 are explained in detail below).

Study Description

Authentic classroom sessions. Two sets of tutorial sessions were taught in Semester 1 and 2, 2012, by the School of Business of the University of Sydney. In the *first set* of 14 classroom sessions, we investigated how a teacher can design and orchestrate small-group activities using an enriched classroom, and subsequently analyse the data to assess that design (Martinez-Maldonado et al., 2012a). The technology used in this preliminary study *did not* have any awareness or control functions available to the teacher. The study informed design of the tools needed to *orchestrate* a classroom, for the second iteration. *This paper focuses on that second set of classroom sessions.* It had 8 tutorials, run in the 6th week of Semester 2, 2012, for a course titled “Management and Organisational Ethics”. In total, 140 students attended these tutorials. Each had 15 to 20 students. The teacher arbitrarily formed four groups, with 4 or 5 students at each table. All students knew each other. The teacher designed a case-resolution activity to cover the set topic as defined in the curriculum for that week.

Activity design. A macro-script was defined by the teacher for the tutorials as follows: 1) *Introduction* (10 minutes): the teacher forms groups, explains the tutorial objective, teaches students how to use the minimalistic concept mapping application and explains the objectives of the first activity. 2) *Activity 1* (10 min.): the teacher uses the MTDashboard to ensure that all groups *start* at the same time. The four tabletops respond by clearing the interface and loading a small scaffolding concept map (5 concepts and 2 links set by the teacher). Students have to complete this map showing how the main actors of the case are linked. 3) *Reflection 1* (5 min.): the teacher *blocks* the tabletops and introduces Activity 2, explaining it and leading class discussion about possible solutions to the case. 4) *Activity 2* (15 min.): for the teacher, this is “*the most important activity of the tutorial from the learning perspective*”. The teacher *unblocks* the tabletops; and students discuss the task and complete their concept map. 5) *Group sharing and final reflection* (10 min.): the teacher *blocks* the tabletops again and then asks each group to share their solution with the class. The teacher uses the function *send to the wall* for each table in turns. After each group has explained their concept map, the teacher summarises the outcomes of the activity and finishes the session. The class time was fixed at 50 minutes.

Visualisations presented to the teacher. Two different conditions of the MTDashboard were used across 8 sessions. For Condition 1, the dashboard (Figure 3, left, and Figure 4, lower right) included the *Group Map Visualisation* that represented the size and distance of each map to the teacher’s map. This information was explicitly requested by the teacher because she wanted this concept map quality measure that is not normally available during the limited classroom time (Figure 4, D1). The second version of the dashboard (Figure 3, right) presented the visualisation *Radar of Physical Participation* that shows the number of touches on the tabletop per student and the equality among group member touches (Figure 4, D2). The design of this visualisation was suggested by the teacher in previous tutorials (Semester 1) who expressed that “*quantitative information about students’ actions would be useful for identifying participation*”. This visualisation was inspired by previous work on group chat communication and physical activity (Martinez-Maldonado et al., 2011). A larger range of visualisations (some more elaborated) were offered to the teacher (Martinez-Maldonado et al., 2012b), but not selected for this study.

Research questions. When teachers orchestrate multiple groups in the classroom, one of their challenges is to identify the group that most needs immediate attention (Dillenbourg et al., 2011b) whilst, concurrently, spending a relatively balanced amount of time with each group, to be fair to all students. This is where MTDashboard can provide awareness support for the teacher, enabling an informed decision about which group to attend next. For this, we sought to address the next questions: *What is the impact of the information provided to the teacher by the MTDashboard during the classroom sessions? Is the teacher attending the “lower achieving” groups according to the information provided?*

Data collection. We collected information from a number of sources to triangulate evidence. These sources included: automated capture of the MTClassroom, notes from an external observer focused on teacher’s actions, notes from a second external observer focused on assessing each small-group work and notes from interviewing the teacher. The *automatically captured data* consisted of synchronised log of the host application at each tabletop (differentiated students’ actions and partial states of the concept maps), logs of teacher’s actions using the MTDashboard, and partial distances of group artefacts from the teacher map. The *manually captured quantitative/qualitative data* consisted of the observed time and duration of the moments when the teacher: i) *attended* or *intervened* a group, ii) looked at the MTDashboard, iii) spoke to the whole group, iv) walked around the class or did not look at any specific group. These observations were captured through a console synchronised with the application logs. The second set of observations consisted of quantitative assessments of perceived qualitative collaboration per group based on an adapted rating scheme (Meier et al., 2007). Our

scheme has 4 dimensions of collaboration, quantified from -2 to 2, for each of (a) mutual understanding and dialogue management, (b) information pooling and consensus (c) task division, time management and technical coordination, and (d) reciprocity. The teacher also assessed groups at the end of each tutorial, using one of three possible values: low, medium or high achieving. Finally, we conducted semi-structured post-tutorial interviews with the teacher to obtain feedback on the functions and visualisations provided for classroom orchestration.

Data exploration. To analyse the teacher’s attention distribution, we first define the terms *attention* and *intervention*. We consider that teachers pay *attention* to a group when their gaze is focused on or they interact with that group. *Intervention* is the *subset* of such *attention* that happens only when the teacher interacts with the group, therefore interrupting their work. We made this distinction based on a previous study in which teachers stated that for some outstanding groups they would “*see what they are doing*” but mostly leave them work by themselves (Martinez-Maldonado et al., 2012b). During the post-tutorial interviews the teacher commented that she “*tried to provide equal attention to all groups*”, while “*focusing on groups that needed more help*”. This means that the teacher dynamically chose the order in which she attended to each group. Having made this distinction, we now describe an example of the teacher’s actions at the MTClassroom. Figure 5 shows a transition diagram where the nodes represent the elements that were at the focus of teacher’s attention. The nodes correspond to *each group*, the *MTDashboard* or the whole *Class*. The latter includes the times when the teacher was not attending to any particular group or gave a message to the whole class. The directed arrows between the nodes represent the transitions recorded by the external observer (45 transitions registered in this example). In this group, the teacher devoted most time to the *red* group (32% of attention and 29% of intervention time) compared with the others (20%, 26%, and 21%). In fact, the teacher assessed the *red* group as the only *low achieving* group in the class, therefore confirming that the attention in this class was not equally distributed. We also observed that the teacher never attended to the *green* group after looking at the dashboard. Coincidentally, the green group also received the fewest interventions. This motivated the analysis of the rest of the cases to find evidence that confirms the impact of the information delivered through the dashboard on teacher’s attention. In other sessions, the accumulated attention was more egalitarian. An analysis of dispersion of *attention* and *intervention* among the sessions showed that the teacher paid attention to all groups largely equally (mean index of dispersion -Gini factor- for attention= 0.12 and intervention = 0.124, where zero means perfect equality). The next section describes our evaluation of the impact of the nature of the information displayed through the dashboard on teacher attention and intervention. The actions that the teacher took after looking at the MTDashboard are the focus of our evaluation (thicker transition lines in Figure 5).

Analysis and Discussion

This section is divided in three parts. The first two tackle our research questions and the last one explores the impact of teacher’s feedback on students to complete our analysis of the orchestration loop at our environment.

Analysis, part 1. For the first question (*What is the impact of the information we provided to the teacher in real-time during the classroom sessions?*), we started by analysing whether there was any relation between the observed *performance* of each of the 32 groups during the tutorials with the accumulated amount of time that the teacher dedicated to attend or intervene each of them. We divided the groups according to the two conditions of the information that was provided to the teacher through the MTDashboard. The two conditions were: (i) *distance to teacher’s map* and (ii) *physical participation*. We performed correlation analyses between *attention/intervention* and *group performance* measured in different levels and from different sources: the external observer that measured collaboration, the artefact that students built and the teacher assessment. Table 1 shows the results of these analyses, where *Attention time* and *Intervention time* are the proportions of the time the teacher dedicated to inspect and interact, or just interact, with specific groups respectively. Regarding the columns of groups’ performance, columns *ob1*, *ob2*, *ob3* and *ob4* correspond to the 4 categories used by the external observer to assess group’s collaboration according to the schema adapted from Meier et al. (2007). Column *Obc* corresponds to the correlations with the cumulative score of these 4. Columns *Size map* and *Dist* correspond to the correlations with, respectively, the size and the distance of groups’ map to the teacher’s map. Finally, the column *Tchr* indicates the correlations with the quality of each group as assessed by the teacher.

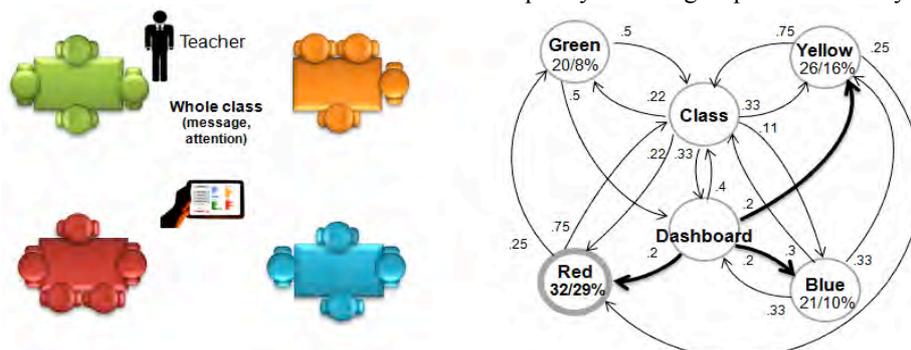


Figure 5. An illustrative transition diagram of the process of teacher’s attention in one classroom

Results showed a difference between the two conditions for the correlation between *observed collaboration* and *attention/intervention time*. For condition (i) *distance to teacher’s map*, we found a significant positive correlation between levels of collaboration and the attention and intervention provided by the teacher (columns *ob2*, *ob3* and *Obc*, left). On the other hand, for condition (ii) *physical participation*, we found a negative moderate correlation (columns *ob1*, *ob3*, *ob4* and *Obc*, right). From a teaching perspective, a negative correlation might appear desirable, since it would mean that the teacher provided more attention to the low groups. However, a perfect correlation is unrealistic, since the teacher cannot totally neglect high achieving students. To explain these findings, we triangulated this evidence with the teacher’s statements during the post-tutorial interviews. The teacher found that the information provided in condition (i) was useful during the class and it was expressed as: “*I looked at the number of relevant links because one group could have 21 links, but how many of them actually matched my map? For a group with 9 linkages with most of them matching my map, I would be satisfied*”. This means that information about the distance to teacher’s map in condition (i) helped the teacher recognise the groups that might have needed more help. The analysis supports this since the only negative correlation of condition (i) was for column *Dist* (-0.3 and -0.16 for attention and intervention). The level of collaboration of groups does not determine the qualitative aspects of their artefacts, therefore there were no negative correlations for *observed collaboration* in condition (i).

Table 1: Correlation analyses between Attention/Intervention and Groups’ performance.

Groups’ performance	i) Distance to teacher’s map							ii) Physical participation								
	Observed collaboration					Size map	Dist	Tchr	Observed collaboration					Size map	Dist	Tchr
	ob1	ob2	ob3	ob4	Obc				ob1	ob2	ob3	ob4	Obc			
Attention time (%)	-0.20	0.33	0.34	-0.10	0.44	0.29	-0.30	-0.10	-0.26	-0.10	-0.33	-0.32	-0.30	-0.10	0.48	0.00
Intervention time (%)	-0.10	0.37	0.34	0.02	0.45	0.34	-0.16	0.02	-0.20	0.05	-0.32	-0.38	-0.27	0.03	0.34	0.15
Correlation	low -	med/low +	med/low +	no	+ p<=0.1	med/low +	low -	no	low -	no	med/low -	med/low -	med -	no	+ p<=0.1	no

For condition (ii), the teacher expressed that the information provided by the *Radars of participation* was good but was not used much because “*a lot of the times groups decided that one only person was going to do the links or I [the teacher] could tell by looking at the table that everyone was discussing but only two or three people were actually moving things around. Then, by looking at the diagrams only, I couldn’t interpret [them] as the group was not working*”. Therefore, during these tutorials for the second condition, the teacher mostly used what she could observe and listen from each group work. We argue that this is the reason why the attention and intervention are more aligned to the *observed* level of collaboration (negative correlations for columns *ob1*, *ob3*, *ob4*, *Obs* in condition *Physical participation*). As the information about the size of the map and the distance to the teacher’s map was not provided in this condition, we found no correlation or positive correlation respectively (columns *Size map* and *Dist*). Finally, the teacher’s assessment seemed independent from her decisions to provide attention (values are close to zero in *Tchr* columns for both conditions). The teacher described that groups’ assessment was primarily based on the explanations that each presented to the class towards the end of the tutorial, and also influenced by the students’ conversations that she could overhear and the groups’ indicators of distance to teacher’s map provided in condition (i). Therefore, the teacher’s assessment was not connected to their evaluation of which groups needed the most help at some point. This suggests that, while the cumulative analysis (*part 1*) is informative in both conditions, we also need to conduct further analysis taking into account the moments when attention was provided to groups.

Analysis, part 2. As groups’ needs for teacher attention vary in time, the teacher needs to continuously monitor groups’ performance to try and keep the levels across groups as close as possible. Here is where our second research question arises: *is the teacher attending the ‘less achieving’ groups according to the information provided?* To answer this, we analysed the decisions made by the teacher right after looking at the dashboard. There were 38 teacher’s actions that were captured by the external observer and synchronised with the MTClassroom’s logs (17 for *distance to teacher’s map* and 21 for *physical participation* conditions).

Condition (i). For each moment when the teacher looked at the dashboard and for each group in the classroom, we calculated the quantitative indicators of size and distance of the map provided by the *Group map visualisation* at that exact moment. Then, the groups were ranked from the smallest and furthest map to teacher’s map to the biggest and closest map at that point in time. There were 3 possible ranks: *furthest behind* group(s), the *strongest* group(s), and the groups in between. The strongest group at a determined moment was the one with more relevant links and less irrelevant links according to the teacher’s map. Then, we identified the group that the teacher chose to attend next. After this, we assessed the category of the group chosen by the teacher, for example, if the teacher chose a furthest behind group or a strong one. Table 2 shows the results of this analysis. Column A corresponds to the 17 cases of teacher’s attention after inspecting the dashboard of the condition under analysis (i). Column B corresponds to the other cases where the second type of information was provided (ii). We found that when the map size and distance to the teacher’s map information was provided (column A) the teacher only decided to attend the strongest group 18% of the times (3 out of 17 cases). On the contrary, when this information was not provided, the teacher attended the strongest group 43% of the times (9

out of 21). This confirms that showing information of each group's artefact in 'real-time' had some impact on the teacher's decision as to which group to attend next. It also validates what the teacher expressed, that looking at the number of relevant links added by each group helped her have a better idea of groups' performance.

Table 2: Analysis the groups that the teacher attended for condition (i) distance to teacher's map.

Condition: distance to teacher's map	A) Map information was provided		B) No information about the map was provided	
Total	17	Proportion %	21	Proportion %
Less achieving	5	29.41	4	19.05
Not the best groups	9	52.94	8	38.10
The best group	3	17.65	9	42.86

Condition (ii). We calculated the information provided by the visualisation *radar of physical participation* for the 38 cases when the teacher looked at the dashboard. We had the same 3 possible ranks. In this case, the strongest group was the more equilibrated in terms of participation. We measured the rank using an index of dispersion, the Gini coefficient. This is a number between zero and 1, where zero means perfect equality of students' participation. We followed the same process as the previous condition. Results are shown in Table 3. These confirm that the participation radar, at least in the way in which we presented it, did not provide information to the teacher to take decisions about which group to attend next. The teacher decided to attend low or high achieving groups almost with no difference (33%, 38% and 28% of the times). The post-tutorials interview confirmed that the teacher did not use the information about physical participation, justifying this with the argument that "not everyone was touching the tabletop but they were speaking a lot and this is good from a learning perspective". The teacher also argued that this information "would be very helpful in a bigger class". The teacher described this as follows: "I cannot observe 80 people but I can observe 20 people, I could tell who was talking. It would be fantastic to check the participation information for a bigger group".

Table 3: Analysis the groups that the teacher attended for condition (i) physical participation.

Condition: physical participation	A) No information about participation was provided		B) Participation condition	
Total	17	Proportion %	21	Proportion %
Less achieving	4	23.53	7	33.33
Not the best groups	9	52.94	8	38.09
The best group	4	23.53	6	28.57

Analysis, part 3. Finally, we investigated whether the teacher's intervention actually had an impact on students' performance immediately after. We considered as indicator of performance the number of relevant links created by each group. The teacher intervened groups a total of 74 times in the 8 classroom sessions. For each intervention, each group was ranked at the moment the teacher started the intervention from 1 to 4 (from low to high group, according to the teacher's map distance of the four groups in the class). Then, we assessed if there was an improvement (or decrease) of the map 2, 3, 4 and 5 minutes later (interventions lasted up to 2 minutes and each activity lasted from 8 to 10 minutes). For example, at minute 5:05 the teacher attends the Green group. At that exact moment, this group had the furthest map to the teacher map in the class, so their rank was 1. We divided the 74 interventions in two groups according to the 2 conditions of the information provided to the teacher. Results on the analyses of correlations between the rank of each attended group and the improvement of the teacher's map distance are shown in Table 4. For condition (i), *Distance to teacher's map*, we found significant negative correlations. This means that the groups that were lagging significantly improved their teacher's map distance after teacher's intervention. However no correlation was found in condition, (ii). We can therefore argue that the teacher's intervention had a significant impact on the groups' artefact when the information about the distance to the teacher's map was provided. This once again provides evidence that supports the benefits of showing information about the quality of students' work to the teacher in real time.

Table 4: the Impact of teacher's interventions: correlation analysis between the rank of a group among the others in the classroom and the improvement of their artefact's distance to the teacher's map.

Condition	# interventions	after 2 min	after 3 min	after 4 min	after 5 min	
Distance to teacher's map	40	-0.4, p<=0.01	-0.27, p<=0.05	-0.32, p<0.025	-0.27, p<=0.05	Significant correlations
Physical participation	34	0.08	-0.05	-0.05	-0.02	No correlation

Our analysis completed the circle of teacher's orchestration that includes: *awareness, intervention and students' action* following this intervention. We found some trends by analysing the accumulated attention and intervention by the end of the tutorials. Then, we obtained stronger evidence confirming the importance of showing indicators of quality of student's work to drive teacher's decision. Finally, we found that *informed* interventions of the teacher can lead students to improve their solutions according to teacher's perspective.

Conclusions and Future Work

We presented our enriched multi-tabletop classroom that afforded the unobtrusive data capture that makes possible to present two sets of information to the teacher in real-time. The potential of MTCClassroom can be wide, from offering simple classroom orchestration controls to awareness and reflection tools. We confirmed that the data presented to the teacher in the classroom can drive their focus of attention especially when

information about the *quality* of students work is delivered. The teacher described this as follows: “*I think the dashboard was really good, especially because it showed things about the quality of their work. If I haven’t had this information about the relevant links then I had to look at the whole diagram so it would take longer to look at each map*”. Our study also confirms that the teacher would value indicators of group work and individual participation for post-hoc analysis. The teacher expressed: “*I don’t want to see a lot of information in the dashboard, this can be distracting. But more information can be provided after the tutorials for assessment wise, like who did what, when, and the quality of the work*”. Our work in progress includes a detailed analysis of the information that should be delivered to the teacher during and after the classroom sessions; and the integration of other sources of information (e.g. verbal participation) and other analysis tools (e.g. data mining) to extract patterns of interaction that can provide more insightful indicators to the teachers in the classroom.

Acknowledgments

This research has been partially funded by Smart Services CRC, Australia. Roberto Martinez-Maldonado has been funded by CONACYT and Fundación Pablo García. Yannis Dimitriadis has been partially funded by Spanish projects (TIN2011-28308-C03-02 and PR2011-0137).

References

- Bull, S., Wasson, B., Johnson, M. D., Petters, D., & Hansen, C. (2012). *Helping Teachers Effectively Support Group Learning*. Presented at the Workshop on Intelligent Support for Learning in Groups, ITS’12.
- Cuban, L., Kirkpatrick, H., & Peck, C. (2001). High access and low use of technologies in high school classrooms: Explaining an apparent paradox. *American Educational Research Journal*, 38(4), 813-834.
- Dillenbourg, P., & Evans, M. (2011a). Interactive tabletops in education. *International Journal of Computer-Supported Collaborative Learning*, 6(4), 491-514. doi: 10.1007/s11412-011-9127-7
- Dillenbourg, P., & Jermann, P. (2010). Technology for classroom orchestration. *New science of learning*, 525-552.
- Dillenbourg, P., Zufferey, G., Alavi, H., Jermann, P., Do-Lenh, S., & Bonnard, Q. (2011b). *Classroom orchestration: The third circle of usability*. Paper presented at the International Conference on Computer Supported Collaborative Learning 2011, (pp. 510-517)
- Do-Lenh, S. (2012). *Supporting Reflection and Classroom Orchestration with Tangible Tabletops*. École Polytechnique Fédérale de Lausanne.
- Martinez-Maldonado, R., Dimitriadis, Y., Kay, J., Yacef, K., & Edbauer, M.-T. (2012a). Orchestrating a multi-tabletop classroom: from activity design to enactment and reflection. Paper presented at the Conference on Interactive Tabletops and Surfaces (ITS 2012) (pp. 119-128).
- Martinez-Maldonado, R., Kay, J., & Yacef, K. (2010). Collaborative concept mapping at the tabletop. Paper presented at the Conference on Interactive Tabletops and Surfaces (ITS 2010) (pp. 207-210).
- Martinez-Maldonado, R., Kay, J., & Yacef, K. (2011). *Visualisations for longitudinal participation, contribution and progress of a collaborative task at the tabletop*. Paper presented at the Computer supported collaborative learning, (pp. 25-32)
- Martinez-Maldonado, R., Yacef, K., Kay, J., & Schwendimann, B. (2012b). *An interactive teacher’s dashboard for monitoring multiple groups in a multi-tabletop learning environment*. Paper presented at the Conference on Intelligent Tutoring Systems, (pp. 482-492)
- Martinez-Maldonado, R., Yacef, K., Kay, J., & Schwendimann, B. (2012c). *Unpacking traces of collaboration from multimodal data of collaborative concept mapping at a tabletop*. Paper presented at the International Conference of the Learning Sciences, (pp. 241-245)
- Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning*, 2(1), 63-86. doi: 10.1007/s11412-006-9005-x
- Mercier, E., Higgins, S., Burd, E., & Joyce-Gibbons, A. (2012). *Multi-Touch Technology to Support Multiple Levels of Collaborative Learning in the Classroom*. Paper presented at the International Conference of the Learning Sciences, (pp. 187-191).
- Morgan, M., & Butler, M. (2009). *Considering multi-touch display technology for collaboration in the classroom*. Paper presented at the EDMEDIA’09, (pp. 674-683)
- Novak, J. (1995). Concept mapping to facilitate teaching and learning. [10.1007/BF02334286]. *Prospects*, 25(1), 79-86.
- Prieto, L. P., Dlab, M. H., Gutiérrez, I., Abdulwahed, M., & Balid, W. (2011). Orchestrating technology enhanced learning: a literature review and a conceptual framework. *International Journal of Technology Enhanced Learning*, 3(6), 583-598. doi: 10.1504/ijtel.2011.045449
- Roschelle, J., & Teasley, S. D. (1995). *The Construction of Shared Knowledge in Collaborative Problem Solving*.
- Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*: MIT Press

Juxtaposing Practice: Uptake as Modal Transposition

Richard Medina, Daniel Suthers, Dept. of Information and Computer Sciences, University of Hawai‘i at Mānoa,
1680 East West Road, POST 309 Honolulu, HI 96822 USA
Email: rmedina@hawaii.edu, suthers@hawaii.edu

Abstract: The analysis discussed in this paper draws attention to the interactional and inscriptional practices observed in a science laboratory setting that utilizes Group Scribbles. The critical finding is the identification of a pivotal sequence of interaction occurring in the later half of the activity in which one member of the group proposes an innovation for illuminating two light bulbs in a single circuit. The proposal and its subsequent endorsement by the other members are contingent on an immediately prior interaction in which the group appropriates another group’s circuit diagram. Together, this pair of adjacent sequential structures exposes multiple instances of uptake between participants. These uptake relations are realized through an ensemble of contingencies consisting of persistent diagrams, tabletop materials, and a locally situated interactional practice. Our analysis shows how the participants’ actions transform the setting and how these transformations are consequential for how the group proceeds in the learning activity.

Introduction

Laboratory science classrooms provide rich sources of data for understanding collaborative learning in small group settings. Laboratory activities place learners in a position to experience scientific phenomena and engage in practices of inquiry. In many cases these activities are organized in participant configurations ranging from pairs to groups of three or four learners situated at the “lab bench”. Such activities typically consist of a teacher’s instructions on procedure and access to instruments and materials needed for the experiment. In the present work we analyzed the recorded interaction of four fifth graders in a Singapore primary school as they collaborated during a thirty-minute lab activity about battery powered electrical circuits. The group is one of ten other groups situated around tables in the classroom. Each table has four tablet personal computers, one for each student, as well as a collection of batteries, small light bulbs, and wires for building circuits. The students also use the Group Scribbles (GS) software installed on each tablet computer to draw diagrams of the circuits they are experimenting with (Roschelle et al., 2007). Using the GS tools, students can share their individual diagrams with their group by dragging them to a public area of the screen. Likewise, drawings in the public area of a group’s screen can be shared with others in the classroom allowing the distribution of artifacts between groups. The inclusion of the Group Scribbles technology provides us an opportunity to assess important connections between meaning making and the appropriation of technology resources.

This paper reports on our investigation of the interaction practices of the learners as they coordinate their actions in a multimodal setting as part of an authentic laboratory science activity. Within this relatively short time frame we are able to piece together a detailed example of how the group deploys practices for handling their work at the “lab bench” as well as the ways in which persistent inscriptional artifacts permeate the group’s setting. Our analysis focuses on a three-minute portion of the video record of the group’s work in which they appropriate another group’s circuit diagram and subsequently develop an alternative solution. By way of this analysis we also demonstrate an application of the use of uptake as an analytic approach for doing CSCL research in multimodal settings (Suthers, Dwyer, Medina, & Vatrapu, 2010). In the following section we ground our analytic approach and argue that it is a valuable approach for evaluating interactive phenomena of the type we are investigating in this paper. We then move on to our analysis and findings.

Background

Investigating practices within a particular setting is a matter of careful consideration of the connection between the sequential organization of participants’ actions and the semiotic, material, and embodied aspects of that setting. Situated interaction has a mutual elaborating relationship with the environment (Suchman, 1987). The environment offers an array of external resources for action-relevant appropriation. Simultaneously, situated action modulates, redefines, or otherwise reconfigures the resources within the environment, thereby enabling and constraining subsequent acts. This dynamic ebb and flow of appropriation and modulation is sequentially organized (Garfinkel, 1994; Goodwin & Heritage, 1990). With respect to multimodal and embodied interaction Charles Goodwin writes,

The issues posed for the analysis of action in such a setting involve not simply the resources provided by different semiotic systems as self-contained wholes, but also the interactive practices required to juxtapose them so that they mutually elaborate each other in a way

relevant to the accomplishment of the actions that make up the setting. (Goodwin, 2003, p. 237)

The study of multimodal interaction requires an understanding of both the properties of these semiotic systems *and* how those systems are coordinated and appropriated in joint activity. This suggests two overlapping planes of study: recognizing the external properties of the setting in one and meaning making practices in the other.

Investigating interaction that takes place in a multimodal setting entails careful consideration of how participants' actions in the environment are made relevant for emergent, sequentially organized, and shared structures of joint activity (Sacks, Schegloff, & Jefferson, 1974). Sequential analysis techniques based on ethnomethodology and conversation analysis (EMCA) demonstrate, in a detailed manner, the contingent nature of human interaction (Çakır, Zemel, & Stahl, 2009). These techniques expose how the resources available in the setting of interaction are integrated in the very structure of communication that emerges in the activity. Taking the turn-by-turn pattern of interaction in any social activity as an analytic starting point has yielded valuable insights into the in-situ emergence of meaning making practices (Koschmann et al., 2005). Although conversation analysis was originally proposed to handle speech exchanges, numerous scholars have taken an EMCA approach, in principle, as an inroad to understanding interaction mediated by more semiotic rich settings such as online environments, scientific field work, and classrooms (Medina & Suthers, 2013; Goodwin, 1995; Roth, 2001).

The distribution of interaction across a setting has led many researchers to see the visual spatial field as a categorical entity in sequential analysis. Streeck and Kallmeyer's (2001) analysis of a rather mundane two-party business negotiation offers an example that suggests that graphic inscriptions can be taken as a form of interaction that offers a different set of opportunities for meaning making beyond conversation. The act of inscribing during interaction carries with it not only that which is being represented, its instrumental purpose, but perhaps less obviously its discursive function. Inscriptions, once recorded in a medium (paper, whiteboard, computer screen, etc.), offer structures for making arguments, substantiating claims, and indexing a range of situation relevant and epistemologically consistent communicative action. The sequential organization of inscriptional activity carries structural (e.g. rhetorical, canonical, or discursive) information that embodies taken as shared conceptions, concerns, and meanings that are relevant to the situation at hand. Inscriptional action draws upon an extended vocabulary from the visual field. They can embody forms of action such as a line intimating gesture (e.g. a line drawn around a figure may be a deictic reference to an aspect of the figure of concern in the interaction). Gesture is highly coupled with talk; however, inscriptions and instrumental acts occur independently of talk yet articulate it. This has not been studied at length especially in regard to how inscriptional action is sequentially organized. Streeck and Kallmeyer (2001) write,

Actions that can occur independently of talk, however - instrumental acts, inscriptions, and so on - have so far only rarely been studied for their possible participation in the construction of 'projectable' turns-at-talk. (Streeck and Kallmeyer, 2001, p. 469)

Inscriptions that once served as a field of calculation and measurement can be reinstated in rhetorical contexts to persuade, compare, and express ideas. Further, persistent inscriptions enable variable, situation relevant courses of action over time and setting (Latour, 1990).

Still, while arguing for analytic accountability of inscriptions and non-verbal modalities in the setting, Streeck and Kallmeyer warn against oversimplifying or fragmenting components of interaction across modal and material properties. Rather, they suggest that ongoing interaction draws upon multiple vocabularies in the making of meaning. Thibault (2011) goes further in advising against the rush to discover and extrapolate upon regularities of symbolic systems. The prudent starting place is the distributional character of communication across the senses, materiality, and symbols.

The analytic approach considered in this paper takes interaction as fundamentally multimodal and sequentially organized. Underlying theoretical assumptions are based on the notion that participants build their interaction through the moment-by-moment or otherwise sequential exchange of actions. These actions are potentially distributed across all available aspects of the setting. As Goodwin's quote makes clear, the study of interaction practices exposes the relationship between the setting and the joint activity of the participants.

Through analysis of practice we gain a rich understanding of the distributional character of action and its implications for computer supported collaborative learning and teaching in classrooms, online settings, and instances of both. Sequential multimodal interaction analysis can be used to uncover the relationship between the properties of the environment and the interactional practices that make those properties relevant and consequential for joint meaning making. More specifically, analysis of interactional practices is useful for understanding how inscriptional devices (verbal and nonverbal) are integrated in joint meaning making structures.

Uptake: A Relational Unit of Interaction

Making sense of the sequential structure of multimodal interaction presents a degree of complexity for analysis where participants' actions may be distributed across a diverse range of media. A useful strategy to begin with

might be the recognition of how any participant actions are evidenced to be relevant and consequential for the activity. How and where are actions *positioned* in the sequential unfolding of the activity and how and through what means do those actions relate to prior actions? The notion of uptake has been proposed as a useful concept for investigating precisely these questions.

Suthers, Dwyer, Medina, & Vatrappu, (2010) describe uptake as a relational construct that identifies a participant action as appropriating aspects of a prior or ongoing setting as relevant for ongoing interaction. This definition is deliberately abstract, enabling it to be purposed in a wide range of interactional settings. It is also intended to support a diverse range of theoretic and methodological approaches. Uptake provides an interpretive heuristic rather than a specific method of analysis (the authors describe it as a proto-analytic). The potential gain by interpreting interaction as uptake is that uptake does not privilege one particular communicative modality or granularity over another. A warranted interpretation of uptake only specifies that one human action is appropriating aspects of a prior or ongoing element of the setting while also transforming that setting. The value of uptake for the analysis of multimodal interaction is its provision for a more flexible consideration of sociological and environmental contingencies.

Group Scribbles Analysis

In the present work we analyzed the recorded interaction of four fifth graders in a Singapore primary school as they collaborated during a thirty-minute lab activity about battery powered electrical circuits (Chen, Looi, & Tan, 2010). The group is one of ten other groups in the classroom. Each group of four is gathered around a table. Each table has four tablet personal computers, one for each student, as well as a collection of batteries, small light bulbs, and wires for building circuits. In addition to these tabletop materials students also use the Group Scribbles (GS) software installed on each tablet computer to draw diagrams of the circuits they are experimenting with (Roschelle et al., 2007). Using the GS tools, students can share their individual diagrams with their group by dragging them to a public area of their respective screens. Likewise, drawings in the public area of a group's screen can be shared with others in the classroom allowing the distribution of GS artifacts (or drawings) between groups.

The students are instructed to begin the activity by drawing circuit diagrams using the Group Scribbles software. After these initial diagrams are complete they are instructed to evaluate their diagrams using the tabletop materials consisting of batteries, wires, and bulbs. Subsequently, over the course of the activity the students are reminded to record their exploration of circuits by diagramming different arrangements of the circuits they tested along the way.

The analysis presented in the following sections takes as a topic the multimodal interactional mechanisms demonstrated by the participants in the laboratory activity described above. The data consists of five synchronized video records of the entire thirty-minute activity. Four of the records were taken from screen capture videos of each of the four participants' tablet computer screens. The fifth video source was drawn from a camera situated adjacent to the participants' table. It captured a wide-angle view of the group's work including what they constructed on the table. The videos were imported and synchronized using the Tatiana software tool (Dyke, Lund, & Girardot, 2009). Tatiana allows simultaneous playback of all video sources. As the videos are viewed, transcriptions and analytic annotations are entered and stored in sync with the video timeline. Tatiana was useful for monitoring simultaneous events in detail. For example, at times we were able observe the drawing actions of a student by focusing on that student's screen video during the event sequence of interest. We could alternatively shift to the group video record to track how those actions related to the interaction of the group. In general, we conducted our video analysis using sequential microanalysis techniques similar in purpose to Jordan and Henderson (1995) although uptake is appropriated here because it offers a general heuristic at a suitable granularity for describing the empirical evidence. This is not to suggest that uptake is only a macro level construct. It can be utilized at any descriptive level appropriate for the analytic evidence on which it is grounded. In the present analysis, uptake is used to explicate the critical relations between participant actions that are distributed across time, media, participants, and material (classroom) artifacts.

The Group Scribbles classroom is organized like many similar learning environments that are computer supported, networked, and embodied in classroom situations. In these contexts interaction is distributed across modalities (verbal, nonverbal, textual, and visual-spatial). Three aspects of the activity were identified at the onset of the analysis. First, the students constructed a series of persistent inscriptions in the Group Scribbles environment. The production of these inscriptions, and their availability throughout the activity, suggested that inscriptions might have a role to play in ongoing group interactions. Second, the centrally located tabletop materials provided a visual spatial modality with respect to the spatial arrangement and placement of the various circuit parts such as batteries, wires, and bulbs. Third, the activity of the group is patterned. That is, the students' work occurs in a series of sequences each oriented to a unique problem or concern in the activity. The interrelation of these phenomena formed the basis for the following analytic questions.

1. How are Group Scribbles inscriptions appropriated and/or coordinated in joint action?

2. How are the elements of the (classroom) environment modulated by the situated actions of the participants and what are the consequences of this for their meaning making practice?

Observations

During the thirty-minute activity the students work through five different electricity circuit configurations, which we refer to as episodes (of interaction). For this paper we focus on the last two episodes (4 and 5) in which the group members develop an innovation for lighting two bulbs with one circuit. In the ensuing descriptions each student is referred to by his or her pseudonym, (Bruno, Serena, Agnes, and Joel). We begin with episode four.

Episode 4: Two-Battery-One-Bulb in Vertical Series

Using Group Scribbles Serena accesses another groups' diagram. She locates a diagram that has two batteries stacked upright one above the other with the light bulb contact directly on the positive end of the top battery. Two wires are arranged from the bottom of the lower battery up to the bulb shielding (see Figure 1a). Serena shares this diagram with the others by swiveling her screen around so that it faces Bruno and Joel. All four students orient to the diagram being referenced by Serena and agree to test its arrangement. The subsequent experimentation moves through two phases. In the first, Bruno and Joel have made an interpretation of the stacked diagram by proceeding to use a wire to complete the circuit connection between the bulb and the positive post of the battery. Serena publicly notes this inconsistency with respect to the diagram posted by group "SF_2". A subtle departure in the stacked diagram from the other diagrams the group has been working with is that the bulb appears to make direct contact with the positive battery post, bypassing the need for a wire (see Figure 1b). Serena points this out and grasps the bulb and places it directly on the positive battery post. The group then proceeds to successfully construct a working circuit using the stacked arrangement (diagrammed in Figure 2). The subtle yet critical diagrammatic placement of the bulb and its successful implementation appears to set up the group's immediate next experiment with two bulbs.

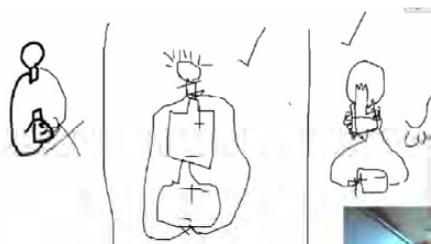


Figure 1a. Group SF_2's diagrams.



Figure 1b. Bulb contact (highlighted by the authors).



Figure 2. Agnes draws configuration.

Episode 5: Two-Battery-Two-Bulb in Vertical Series

The last experiment has the group attempting to light two bulbs. They are building on the "stacked" arrangement of the circuit they successfully tested in the immediately prior experiment (episode four). Joel initiates this experiment by spontaneously picking up two bulbs and placing them both directly on the positive post of the top battery. Joel performs this act while Bruno remains holding the batteries in the position they were in during the prior experiment. At this moment there is no direct evidence from the video record that the bulbs are actually lit. At the very least Joel's act is presenting an idea to the group. He appears to position the bulbs, then looks up to the group as if requesting their noticing (this moment is captured in the frame presented in Figure 3 inset 3 below). The group takes notice and excitedly proceeds to successfully construct the circuit with two bulbs.

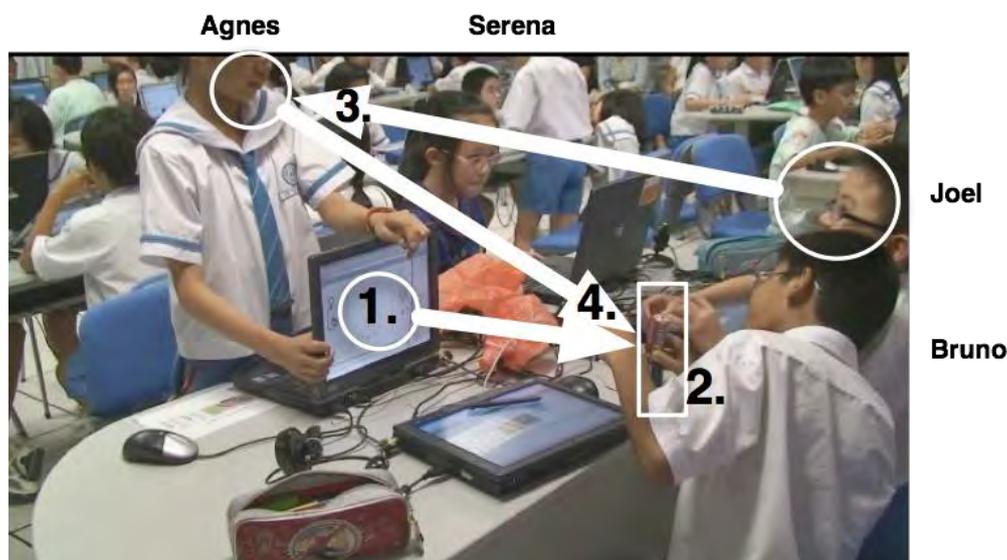


Figure 3. Proposing an innovation.

As noted above the group's approach does not require a wire for the positive contact (i.e. the bulb is placed directly on the positive post of the top battery). This innovation opens up the possibility of adding a second bulb to the positive post of the top battery. It is this opportunity that Joel seizes and demonstrates and that the others recognize. Subsequently, their competence at constructing a working circuit is evidenced in the efficiency at which they move from Joel's initial idea to a working two-bulb circuit.

Discussion

The sequence of interaction in episodes 4 and 5 described above demonstrate the learners' opportunistic appropriation of various elements of the setting to conduct their laboratory activity. Our analysis below reveals how the learners juxtapose their practices against the material and semiotic elements of the setting in order to coordinate their actions with the tabletop objects and the inscription displayed in the Group Scribbles screen. In the following we identify learners' uptakes to highlight critical aspects of the learners' coordination with the setting and to expose what we view as important implications for CSCL research and the design of media-rich technologies.

How Inscriptions Permeate Interactional Context: Representational Competence

In episode four, the group members have integrated another group's diagram into their interaction. The diagram is displayed on Agnes' screen and is directed towards the center of the table. It depicts a vertical arrangement of a circuit in which the bulb makes direct contact with the battery post (no wire is needed for that portion of the circuit), which the group had not previously considered. As Bruno and Joel work at reconstructing the diagram using the tabletop components, Serena notes that they have misinterpreted the diagram. She observes and demonstrates to the others a novel detail in the placement of the bulb with respect to the top battery (see Figure 1b). The group subsequently modifies their circuit to maintain consistency with the diagram.

Here we can observe a series of uptake relations. Bruno and Joel's reconstruction (using the tabletop materials) of the diagram is the first in the series. This uptake affirms the proposed relevance of the diagram to the group's ongoing physical experiment. Next, Serena takes up the result of this reconstruction when she observes how their arrangement is inconsistent with the diagram. Subsequently, Bruno and Joel rearrange the tabletop materials accordingly as they take up Serena's point. By definition, there are multiple uptake relations in which a participant action takes up aspects of the setting as relevant for the situation at hand and by so doing transforms the setting (Suthers et al., 2010). Bruno and Joel's uptake of both the diagram and the prior availability of the circuit materials result in a transformation of the experimental apparatus. In turn, Serena takes up this transformation and invokes the diagram that remains directed at the group (Figure 3, inset 1). Bruno and Joel take up Serena's indexical act by adjusting the experimental apparatus, once again relying on the previous setup. Figure 4 (uptake relations u1 through u6) summarizes the pattern of uptake discussed here. From an interactional exchange perspective, experiment four might appear rather straightforward and mundane. Interpreted as uptake, however, the brief exchange sheds a different light on the significance of interactions with

respect to how the participants are cooperatively and opportunistically appropriating a range of resources in the setting well beyond verbalization. Uptake relations show how a sequence of actions produces relationships between inscriptions and established practices for handling tabletop materials. Further, the very notion that Bruno and Joel can be referred to as jointly performing an action is a notable example of how tightly coordinated the participants' actions are with their laboratory practice for maintaining the experimental apparatus.

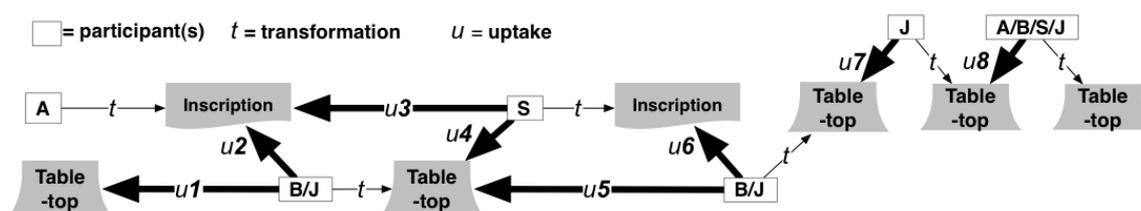


Figure 4. Uptake relations in episodes 4 and 5.

Episode four demonstrates a level of representational competence with respect to the group's ability to correlate the phenomena (electrical circuits), with its associated diagrammatic inscription (Kozma, 2003). In this group's work, there is a clear, non-abstract, relation between diagram and phenomena (bulbs look like bulbs in the diagram). The critical point, however, is that it is the group's orientation to Serena's indexical act that reveals that the others clearly understand the misinterpretation and immediately rectify the relations between diagram and phenomena. The diagram is appropriated for representation and as a resource for negotiated meanings.

Making a Proposal: Coordinating Actions

A second critical moment comes immediately after the group has completed episode four. At this moment (Figure 3) Joel makes his proposal for lighting two bulbs by building on the circuit configuration constructed in episode four. There are a number of points to note here. (1) The prior arrangement of circuit materials, originally taken up via a diagram displayed in the Group Scribbles software, remains intact on the table. Further, Bruno's hand remains on the batteries to keep them in the stacked position. These are physical contingencies for the formulation of ideas. The arrangement of batteries, wires, and bulbs require bodily coordination to be distributed across individuals and this is taken up by Joel to establish his proposal. (2) Joel places the two bulbs on the post and looks up towards Agnes as if to request a response. Agnes and Bruno take up Joel's act as demonstrating a new arrangement (two bulbs). After some excitement Serena, who has directed her attention at Bruno's and Joel's attempt to implement the proposal, reaches over the table to position the necessary wire to complete the circuit. In this instance, Serena is taking up the now established practice that requires coordinated arrangement and manipulation of the tabletop components to maintain the experimental apparatus. This practice became necessary over the course of the entire activity and is demonstrated once again here.

The proposed innovation offered by Joel and the subsequent uptake and implementation performed by the group demonstrates the coordination of multiple aspects of the setting. The proposal is initiated when the group reconstructs a circuit diagram depicted on Agnes' computer screen, which she has faced towards the group (Figure 3, inset 1; and described above). Joel's proposal and its recognition by the group members is achieved by a configuration of an indexical field consisting of the batteries held in place by Bruno along with Joel's placement of two bulbs (Figure 3, inset 2). Here, Joel is taking up a prior tabletop configuration and transforming it into another (Figure 4, u7). This is followed by his upward gaze towards the shared space above the table (Figure 3, inset 3), and Agnes' uptake of the gaze and position of the two batteries as communicating an idea (Figure 3, inset 4). Bruno's endorsement may be implicit as he maintains the physical apparatus to sustain Joel's communicative action. The group members are configuring the environment through multiple surfaces to mediate their meaning making. In this case the members of the group build on their prior locally available interactional practices for constructing experiments with electrical circuit materials. Evidence for this association is the placement of Bruno's right hand, (Figure 3), as Joel enacts his proposal for two bulbs. The configuration of body parts and orientation of the learners with respect to the table is entwined with the group's interactional context. The learners effectively take up the idea represented in the diagram by leveraging prior arrangements of bodies and hands. This arrangement is established over the course of the entire thirty-minute activity and constitutes important part of the group's interactional practice in an ongoing and changing semiotic context (Goodwin, 2007). This practice includes constructing and managing the experimental apparatus and attending to the relevant parameters of the problem. In this case, the learners have demonstrated evidence of

competence at both while also invoking the practice to make a discovery (Amerine & Bilmes, 1988; Koschmann & Zemel, 2009; Roth, McRobbie, Lucas, & Boutonné, 1997).

Confluence of Uptake

The final observation we wish to make rests on the notion of mutability. At the onset of experiment four, the group has been shown a diagram constructed by another group in the classroom. This diagram (Figure 1a) is an immutable mobile (Latour, 1990). Bruno and Joel's uptake of the diagram is to reconstruct its features using the tabletop materials. As detailed above, the next series of actions project through episodes four and five. Here we see demonstrated the transposition between modes such that the immutable object displayed on the screen is made mutable through its reconstitution in the tabletop materials. Realization of the affordance of mutability establishes the necessary relationship between the actors and the setting making it procedurally consequential (Schegloff, 1991). In other words, the uptake of the diagram transforms the setting by projecting the next set of relevant actions. These actions are to build a complete circuit, which requires the group to coordinate what they know about circuits and how to actually achieve the necessary arrangement of materials. This observation suggests that uptake relations that take the form of modal transposition (Peeters, 2010) may dramatically contribute to interactional practice because it invokes coordination with and the instantiation of the relevant aspects of the problem (Koschmann, LeBaron, Goodwin, & Feltovich, 2006). Competencies for reformulating ideas and concepts across modes has been used to assess levels of conceptual coherency in laboratory science activity settings similar to the one discussed in this paper (Lund & Gu-Robinault, 2010, 2012). This is especially relevant for analysis of interaction at the lab bench where multiple resources are assembled and invoked by learners as they move through their inquiry. This view parallels Kozma's (2003) assessment of representational competence. In the case of our participants the placement of the bulb in episode 4 becomes a problem raised by Serena but only made salient after Bruno and Joel transposed the circuit from diagram to the tabletop.

Conclusions

The directive given to the participants was to use diagrams to reason about and explore the concept of electricity flow. This analysis showed how diagrams shared as classroom artifacts provided a resource on which the participants *juxtaposed* and configured their local interactional practices and developed competencies for experimentation and meaning making in a laboratory science activity. Uptake was used to describe the interaction contingencies that were observed and to demonstrate how contingencies scale across the embodied, physical, semiotic, and temporal facets of the setting. We observe that practices are made accountable and support the development of group-level competencies. In our data we conclude (a) that persistent inscriptions provide a durable semiotic resource for making instantiations of relevant and emergent components of a problem (Alac & Hutchins, 2004; Koschmann et al., 2006). That is, salient aspects can be located or are "locatable in the setting" and are consequential to how a group proceeds and (b) that uptake across modes (modal transposition) requires mutable surfaces for any action to perform a transformation in the setting. Thus, uptake is dramatically facilitated by opportunities for distributed interaction across mutable media.

References

- Alac, M., & Hutchins, E. (2004). I see what you are saying: Action as cognition in fMRI brain mapping practice. *Journal of Cognition and Culture*, 4, 3(4), 629–661.
- Çakır, M. P., Zemel, A., & Stahl, G. (2009). The joint organization of interaction within a multimodal CSSL medium. *International Journal of Computer-Supported Collaborative Learning*, 4, 115–149.
- Garfinkel, H. (1994). *Studies in ethnomethodology*. Polity Press.
- Goodwin, C. (1995). Seeing in depth. *Social Studies of Science*, 25, 237–274.
- Goodwin, C. (2003). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where Language, Culture and Cognition Meet* (pp. 217–241). Mahwah, NJ: Lawrence Erlbaum Associates.
- Goodwin, C., & Heritage, J. (1990). Conversation analysis. *Annual Review of Anthropology*, 19(1), 283–307.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4(1), 39–103.
- Koschmann, T., LeBaron, C., Goodwin, C., & Feltovich, P. (2006). The mystery of the missing referent: Objects, procedures, and the problem of the instruction follower. In P. Hinds & D. Martin (Eds.), *CSCW '06: Computer Supported Cooperative Work* (pp. 373–382). New York, NY, USA: ACM.
- Koschmann, T., Zemel, A., Conlee-Stevens, M., Young, N. P., Robbs, J. E., & Barnhart, A. (2005). How do people learn? In R. Bromme, F. W. Hesse, & H. Spada (Eds.), *Barriers and Biases in Computer-mediated Knowledge Communication: And How They May Be Overcome* (pp. 265–294). Amsterdam: Kluwer Academic.
- Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, 13(2), 205–226.

- Latour, B. (1990). Drawing things together. In M. Lynch & S. Woolgar (Eds.), *Representation in Scientific Practice* (pp. 19–68). Cambridge, MA: MIT Press.
- Lund, K., & cu-Robinault, K. (2010). Learning physics as coherently packaging multiple sets of signs. In *Proceedings of the 9th International Conference of the Learning Sciences - Volume 1* (pp. 404–411). International Society of the Learning Sciences.
- Lund, K., & cu-Robinault, K. (2012). Conceptual change and sustainable coherency of concepts across modes of interaction. In D. D. Suthers, K. Lund, C. P. Rose, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions*. New York: Springer.
- Medina, R., & Suthers, D. (2013). Inscriptions Becoming Representations in Representational Practices. *Journal of the Learning Sciences*, 22(1), 33–69.
- Peeters, H. (2010). Multimodality and its modes in novelizations. *Image & Narrative*, 11(1), 118–129.
- Roschelle, J., Tatar, D., Chaudbury, S. R., Dimitriadis, Y., Patton, C., & DiGiano, C. (2007). Ink, improvisation, and interactive engagement: Learning with tablets. *IEEE Computer*, 40(9), 38–44.
- Roth, W.-M. (2001). Situating cognition. *Journal of the Learning Sciences*, 10(1-2), 27–61.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 696–735.
- Schegloff, E. A. (1991). Reflections on talk and social structure. *Talk and social structure: Studies in ethnomethodology and conversation analysis*, 44–70.
- Streeck, J., & Kallmeyer, W. (2001). Interaction by inscription. *Journal of Pragmatics*, 33(4), 465–490.
- Suchman, L. A. (1987). *Plans and situated actions*. Cambridge University Press.
- Suthers, D. D., Dwyer, N., Medina, R., & Vatrapu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. *International Journal of Computer-Supported Collaborative Learning*, 5, 5–42.
- Thibault, P. J. (2011). First-order languaging dynamics and second-order language: The distributed language view. *Ecological Psychology*, 23(3), 210–245.

Emotion Feedback during Computer-Mediated Collaboration: Effects on Self-Reported Emotions and Perceived Interaction

Gaëlle Molinari, Distance Learning University Switzerland & TECFA/University of Geneva, Sierre & Geneva, Switzerland, gaelle.molinari@unidistance.ch

Guillaume Chanel, Swiss Center for Affective Sciences, Geneva, Switzerland, Guillaume.Chanel@unige.ch

Mireille Bétrancourt, TECFA/University of Geneva, Geneva, Switzerland, Mireille.Betrancourt@unige.ch

Thierry Pun, Computer Science Department/University of Geneva, Switzerland, Thierry.Pun@unige.ch

Christelle Bozelle, Swiss Center for Affective Sciences, Geneva, Switzerland, Christelle.Bozelle@unige.ch

Abstract: Emotions play a crucial role in collaboration. They help to make inferences about the partner and can strongly influence task performance. Due to limitations of emotional cues in computer-mediated collaboration (CMC), the collaborative process can be impacted. In this study, we report on the effect of an Emotion Awareness Tool (EAT) designed to facilitate the sharing of emotions between partners, on the perceived emotions after collaboration and the perceived quality of the interaction. Results showed that the EAT stimulated participants to engage in a mutual modeling of emotions. In the EAT condition, the perceived amount of time spent on emotion modeling process was positively correlated to the perceived intensity of positive emotions after collaboration. The EAT increased the perceived degree of transactivity, but only for women. This study provides a first step in exploring the effect of emotion awareness in CMC tasks including a comparing approach for its gender-specific relevance.

Introduction

There is now a large consensus among researchers on the fact that emotions impact a broad range of cognitive and social processes, and that humans are able to use emotional information to regulate their activity as well as to influence others (Van Kleef, 2009). A change in individuals' emotional states can reorient their attentional focus and induce a change in the way they think, act and interact with others. In cognitive theories, emotion is defined as resulting from the appraisal of (external or internal) events, involving the synchronization of 5 interrelated organismic subsystems. These subsystems underline the 5 components of an emotion, the cognitive, neurophysiological, motivational, motor expression and subjective feeling components (Scherer, 2005).

Emotions are strongly related to people's knowledge and goals, especially in high-order activities (learning, problem-solving, decision-making...) in both individual and group settings. For instance, when learning alone from a low cohesion text, or when solving an ill-structured problem together with a partner, individuals have to cope with many difficulties in, e.g., filling cohesion gaps in the text or converging on a joint solution in the discussion with the partner. Such coping is even more complex when individuals have low prior knowledge or when the difference in knowledge between collaborators is relatively high. On the one hand, encountering those difficulties may result in (socio-)cognitive disequilibrium than can produce negative emotions like confusion, frustration or boredom in case of persistent failure. On the other hand, positive emotions like satisfaction, pride and engagement/flow may occur when learners are successful in coping with their difficulties and as a result converge to a new equilibrium state. Positive emotions were found to be positively related to individual learning (D'Mello & Graesser, 2012). Moreover, a shift towards a more positive mood can increase flexibility and creativity thinking (Davis, 2009). The relationship between negative emotions and (socio-)cognitive processes is not so obvious: Whereas negative emotional states like frustration and boredom are detrimental to learning, there is a positive correlation between confusion and learning as it can stimulate active and deep processing (Lehman et al., 2012). This positive impact of confusion could also partly explain why in some cases, collaborative learning failure can be productive (Kapur, 2009). Besides studies on decision-making also found that people in a negative mood are more inclined to seek new information, engage in analytical information processing and produce higher quality decisions (van Knippenberg et al., 2010). Stemming from the field of affective computing (Picard, 1999), recent studies have proposed and shown the feasibility to build computer interfaces able to react to learners' emotions with the goal of improving learning outcomes (Kapoor et al., 2007). The above results and technologies clearly emphasize the need to better understand the role of emotions in individual and collaborative learning.

Emotions in Social Interaction

The study of emotions in groups becomes an important research area in social psychology, and the focus is on how social processes influence the emotional feelings and expressions, and vice-versa. In an interpersonal approach of emotions, Van Kleef (2009) proposed the EASI (Emotions as Social Information) model as a framework to predict when and how the expression of emotions affects observers' behavior. In this model, there are two distinct but mutually dependent ways individuals can be influenced by their partners' emotional

expressions in social interaction settings. On one hand, partners' emotions are processed below the conscious level and such processing results in affective reactions. This is the case when individuals automatically feel the same emotions than those felt by their partners through emotional contagion (Hatfield et al., 1993). Such emotional reactions are also recognized as playing a determinant role in interpersonal liking processes. On the other hand, individuals strategically and consciously use emotions to make inferences about their partner, and adapt their behavior in consequence (e.g., providing help when the partner displays frustration). According to Van Kleef (2009), both emotion-processing paths (affective reactions and inferences) can lead to similar (e.g., feeling compassion when seeing the partner in difficulty) or opposite (e.g., keeping calm when facing an angry partner) behavior. The use of one of these paths when processing emotional expressions also depends on both observers' epistemic motivation, their cognitive resources as well as the cooperative or competitive nature of the social situation. Individuals are more likely to engage in inferences about their partner's emotions when they are motivated to build and update an accurate representation of the situation and/or when this situation is perceived as competitive. The probability to use inferential paths also decreases with fatigue and time pressure.

Research also showed that the expression of emotions depend on sets of social (and cultural) rules determined by the characteristics of the social interaction, the nature of the collective work, as well as interpersonal power processes (Ragins & Winkel, 2011). This may explain differences between men and women in the expression of emotions (Brody, 2000). It has been shown that men are less likely to display emotions they experience than women – especially with other men, while women allow themselves more easily to express their emotions to a wider range of persons, independent of their gender (Rime et al., 1991). Moreover, Ragins and Winkel (2011) argued that in work contexts, women are expected to display emotions (compassion, worry or fear) that are usually related to less interpersonal power than emotions expected of men (confidence, pride or anger). From a gender/neural perspective, McRae et al. (2008) found no difference between men and women with regard to emotional reactivity; results suggest rather a discrepancy in the emotion regulation process. Men seem to be able to regulate their emotions with less effort and greater efficiency than women. Compared to men, women tend to generate positive emotions to a greater extent when trying to down-regulate negative emotions.

Emotions in Collaborative Tasks

During the past five years, a growing body of research has focused on the role of emotions in collaborative learning (CL) situations, and more specifically on their relation to students' social-behavioral engagement (Linnenbrink-Garcia et al., 2011) and regulation processes during CL tasks (Järvenoja & Järvelä, 2009). CL situations can be viewed as being “more challenging than conventional and well-structured learning situations” (Järvenoja, 2010, p. 68), although the ultimate (shared) goal in such settings - that is, the construction of a shared understanding (Roschelle & Teasley, 1995) - can be associated with positive emotions and high motivation (Eligio et al., 2012). The CL process can be understood as an interpersonal matching process that evolves over time, that is, moment-by-moment during the course of interaction (and probably also after that). It is described as a constant adjustment of tension between interpersonal convergence (necessary for joint actions) and divergence (necessary for flexibility and creativity) in terms of perceptions, actions, knowledge and emotions. In the same vein, Andriessen et al. (2010) argued that there are two interrelated tuning processes during CL tasks, a cognitive tuning (confrontation/differentiation of ideas) and a socio-relational tuning (maintaining a collaborative working relationship). The CL experience is therefore characterized by continuous fluctuations of tensions and relaxations between learning partners. Tension may arise from the expression of divergent information (due to learners' differences in knowledge, intentions or cognitive abilities) and conceptual conflict; the greater is such a tension and “the more potential mutual gain is present in the situation” (Andriessen et al., 2010, p. 227). However, when the tension is too high and/or when the focus in the group shifts towards social comparison of competence (Darnon et al., 2006), negative emotions may emerge and as a result, learners try to protect their own competence (face-saving process and use of competitive strategies). Since negative emotions can impair learning, emotion regulation processes need to take place during interaction so as to reduce tension between partners. These emotion regulation processes are both individually and socially constructed (Järvenoja & Järvelä, 2009), and are motivated by the co-learners' need to converge towards a joint solution. According to Järvenoja and Järvelä (2009), the co-learners' efforts to overcome together socio-emotional challenges can be viewed as “critical points [...] in terms of successful learning and interaction”.

Emotion Awareness and Computer-Mediated Collaboration

Unresolved socio-relational tensions and the resulting negative emotions may have a dramatic impact on collaborative processes and outcomes. The understanding of the partners' emotions is thus necessary to trigger emotion regulation strategies that will favor successful collaboration. In this paper, we argue that emotion understanding is part of the mutual modeling process through which collaborators build a representation of what their partners know, believe and intend to do. In previous research (Molinari et al., 2009; Sangin et al., 2011), we found a positive correlation between the accuracy of mutual knowledge modeling and learning in Computer-Supported Collaborative Learning (CSCL) settings. We hypothesize here that individuals' ability to recognize

and understand their partner's emotions also plays a crucial role in the way they communicate and build a shared understanding; such ability would facilitate processes such as audience design and perspective taking.

In face-to-face (F2F) situations, people rely on a whole set of explicit and implicit mechanisms to adapt to their partners and the situation. In computer-mediated collaboration (CMC), contextual non-verbal cues – such as facial expressions, head movement, eye gazes – are missing or seriously limited. The awareness of others may therefore be impaired and this may lead to inefficient interactions. In recent years, research has been conducted to investigate the role of emotions in CMC (Derks et al., 2007). These studies showed no differences between F2F and CMC settings with respect to expression of positive emotions and even suggest that people express more freely their negative emotions in CMC. Besides emoticons and acronyms (e.g., “lol”) are regularly used in online interactions to express one's emotional states. However, men rarely use emoticons in conversation and feel less satisfied with CMC experiences than women (Lee, 2003).

The use of group awareness technologies is becoming widespread to circumvent the bottlenecks of CMC. Such technologies aim at analyzing users' characteristics and behavior and feeding that information back to the group. In CSCL contexts, group awareness tools are designed so as to improve and expand social and cognitive processes during collaborative learning (Buder, 2011), by making explicit and visible what is not directly observable like e.g., the group members' prior knowledge (Sangin et al., 2011) or their participation level during online discussions (Janssen et al., 2011). To our knowledge, there is still little research on the effects of *emotion awareness tools* designed to provide collaborators with information about their partner's affective states during online collaboration. Eligio et al. (2012) carried out experiments with the aim to investigate the relation between emotion understanding and performance in CMC. It is noteworthy to point out that only women participated in these experiments to “avoid the controversy of gender differences regarding the interpersonal understanding of emotions” (Eligio et al., 2012, p. 2049). Results showed that collaborators had difficulties to accurately assess their partner's emotions in CMC situations (Study 1). In order to overcome such difficulties, collaborators were instructed to share their self-reported emotions with their partner during specific moments of the task (Study 2). Results showed – for remote collaborators – both higher group performance and higher accuracy at estimating their partner's emotions in the emotion awareness condition. This suggests a positive impact of emotion awareness tools on collaborative processes and outcomes.

Objectives and Research Questions

The aim of our research project is to explore the impact of reciprocal emotion awareness on collaborative processes and outcomes. The overarching goals are twofold: (1) To shed light on the benefits of providing emotional feedback in CMC situations, and (2) to resort on affective computing to develop adaptive collaborative “emotionally aware” systems able to automatically provide emotional feedback when necessary. In the reported study, participants were provided with an emotion awareness tool (EAT) with which they self-reported and shared their emotions *explicitly* during a computer-mediated collaborative design task. Our goal in this paper is to investigate the effects of the EAT on the subjective perception participants had of their collaborative work experience. More precisely, we wondered to what extent sharing emotional awareness information during remote interaction influences participants' perceptions (a) of both their own- and their partner's emotional states after collaboration (Question 1), and (b) of the quality of interaction with their partner (Question 2). By perception of the quality of interaction, we meant participants' perceptions of the frequency with which they (and also their partner) defended and argued their own ideas, built up on or challenged their partner's ideas as well as processed and managed emotions during collaboration. Unlike Eligio et al. (2012), both women and men participated in our study (they were paired in same-gender dyads) and we studied how the effects of the EAT varied depending on gender. Finally, we examined the relationship between participants' perceived intensity of emotions after collaboration and their perception of the interaction with their partner, and also how the EAT can impact this relationship (Question 3).

Method

Participants and Design

Sixty participants (32 women and 28 men, mean age = 23.4 years) took part voluntarily in the study. They were randomly assigned to 30 same-gender dyads. Fifteen dyads were randomly assigned to each of the 2 following conditions: (1) experimental (EAT) condition (8 women dyads and 7 men dyads) in which the participant were provided with the Emotion Awareness Tool; (2) control condition (9 women dyads and 6 men dyads) in which they were not provided with the EAT. Group members did not know each other, and everything was done so that they did not see each other before the experiment. Each participant was remunerated 60 Swiss Francs.

CSCL Environment

DREW (Corbel et al. 2003; see also Lund & Molinari, 2007) was used as the CSCL environment in which the collaborative task had to be performed. Participants were asked to use the argument graph tool (left/blue part of

Figure 1) to construct together a joint map (see the “procedure” section to have a description of the task). In this map, boxes could be linked to each other using two types of links, “+” or “in favor” links and “-“ or “against” links. Besides, participants had the possibility to express their opinion “for” or “against” for any boxes (or links) in the map (each participant’s opinion appeared in a different color). Boxes for which two opposed opinions have been expressed, appeared in a “crushed” form. During the construction of the joint map, participants could communicate with each other through microphone headsets (their verbal interactions were recorded).

Participants of the experimental condition were asked to self-report and share their emotions to their collaborative partner using the EAT (right/red part of Figure 1). The lower part of the EAT consisted of a list of 10 positive (e.g., engaged, interested, satisfied, relaxed) and 10 negative (e.g., anxious, frustrated, unsatisfied, tired) emotions, each of them referring to a button to click. The upper part of the EAT was dedicated to the display of emotions on which participants clicked. The participants’ emotions appeared in the green area, their collaborative partner’s emotions in the blue area. Moreover, in both green and blue areas, the current emotion was displayed in a box with a lighter color background, immediately followed by the two previously inputted emotions. Participants could also enter any emotion directly in the lighter green box. Participants were instructed that they were free to self-report their emotions at any time they wanted during interaction. They were however prompted to self-report their initial emotion and a pop-up message appeared 5 minutes after the last inputted emotion to remind them to indicate the emotion they were experiencing at that moment. Participants of the control condition were not provided with the EAT (the screen part corresponding to the EAT was shaded).

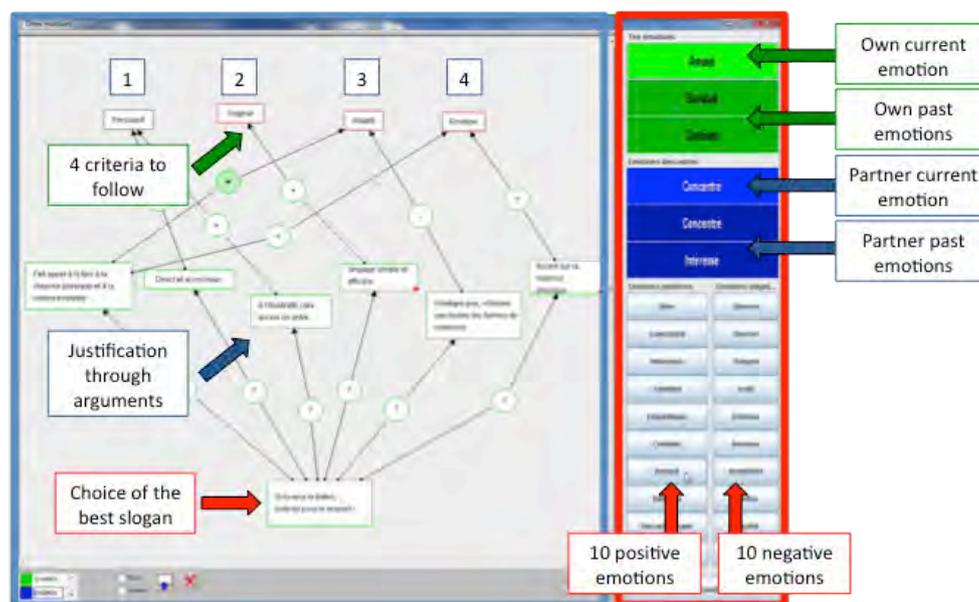


Figure 1. The DREW interface coupled with the EAT.

Procedure

The members of each dyad were separated in two different rooms. Both peers were in front of computers equipped with (a) webcams, (b) Tobii T120 eye-trackers and (c) BioSemi physiological data acquisition systems (eye-tracking and physiological data will not be the focus of this paper). They could not see each other, and could not use the webcams to communicate with each other at any point of time during the session. The experimental session lasted 140 minutes, and consisted of four phases:

(1) *Introduction* (60 min): Participants were equipped with BioSemi physiological sensors, and eye-tracker calibrations were completed. Training was given to get them familiar with DREW (and with the EAT in the experimental condition).

(2) *Collaborative design task* (40 min): Participants were asked to perform (in dyads) a brainstorming exercise so as to design together a slogan against violence in schools intended for teenagers. During the brainstorming, they drew a joint map, using the DREW argument graph tool, in 3 steps. In Step 1, both partners generated as many boxes of slogan ideas as possible in the map. In Step 2, they were asked to debate and argue slogan ideas depending on four criteria boxes given in the map (persuasive, original, adapted to audience, and emotion). Slogan ideas should be therefore linked to criteria boxes through argument boxes. After debating, peers suppressed the less relevant slogan ideas and improved those remaining. In the last step (Step 3), they were asked to negotiate and find a consensus about the best slogan; the result should appear in a new box entitled “final slogan” at the end of the brainstorming (see Figure 1).

(3) *Post-test questionnaires* (20 min): After collaboration, participants first rated the intensity of their own- and their partner's emotions. The felt intensities of 20 emotions (the same 10 positive and 10 negative emotions than those used in the EAT) were measured on 7-points items (ranging from 1 = "very low or not at all" to 7 = "very high"). They then answered a questionnaire about their perceptions of the interaction they had with their partners. This questionnaire was constructed based on that developed by Buchs et al. (2004), and consisted of 5 groups of questions. In this paper, we only reported the analysis of answers to questions of Group 4 and Group 5. These questions referred to the participants' perceptions of their own- and their partner's activities during collaboration. Seven-point scales (ranging from 1 = "little time" to 7 = "much time") measured how much time participants felt they (or their partner) had spent defending ideas and arguing about them, building up on the other's ideas, comparing their emotions to the other's emotions, communicate on emotions, etc. (see Table 1 for the complete list of items). Finally, three questionnaires were administered to assess participants' emotional characteristics, the Emotional Expressivity Scale (Kring et al., 1994), the Emotional Contagion Scale (Doherty, 1997), and the Emotion Regulation Questionnaire (Gross & John, 2003).

(4) *Debriefing* (20 min): Participants were provided with a video of their group work (including their face and the shared screen), and annotated 10 moments where they felt an emotion (and 10 without emotions).

Variables

The presence of the EAT (with vs. without the EAT) and Gender were the between-subjects factors. The dependent variables were: (a) the participants' ratings of the intensity of their own emotions and their partner's emotions (the 20 emotion items, 10 positive and 10 negative emotions), and (b) the participants' answers to two groups of questions (Groups 4 & 5) measuring their perception of their own- (15 items) and their partner's activity (15 items) during interaction (the "own-activity" and "partner-activity" items were equivalent). They were requested in the perceived interaction questions to rate the frequency with which they – and their partner – provided/imposed their own points of view, defended and argued their ideas, understood their partner's points of view, built up on their partner's ideas, as well as managed emotion during interaction (see Table 1).

Results

Q1: Effect of the EAT on Perception of Emotions after Collaboration

Concerning participants' own emotions, results showed that intensity ratings were higher for positive emotions ($M = 4.51$, $SD = 0.84$) than for negative emotions ($M = 1.76$, $SD = 0.68$), $t(1, 59) = 20.07$, $p < .001$, Cohen's $d = 3.64$. The same pattern occurred for the perception of the partner's emotions (positive: $M = 4.41$, $SD = 0.85$; negative: $M = 1.64$, $SD = 0.56$), $t(1, 59) = 21.59$, $p < .001$, Cohen's $d = 3.76$. A series of 2 (EAT) x 2 (Gender participant) ANOVAs were performed on intensity ratings for own/partner positive and negative emotions. The results did not reveal any significant main effects for the factors studied and their interactions.

Q2: Effect of the EAT on Perception of Interaction with the Partner

Thirty perceived interaction items were used in the analysis (see Table 1). They were submitted to a factorial analysis (FA) with promax rotation. The Cattell Scree test indicated the presence of three factors that accounted for 50.72% of the total variance, namely (a) F1: *to communicate on emotions and adapt to emotions* (27.51%), (b) F2: *to compare emotions and imagine reactions to emotions* (12.52%), (c) F3: *to argue and build on the other's ideas* (10.69%). The items and their factor loading are presented in Table 1. It is noteworthy that F1, F2 and F3 included both "own-activity" and "partner-activity" items; we thus decided to talk about the *activity of the dyad* when presenting the results concerning those factors.

A series of 2 (EAT) x 2 (Gender participant) ANOVAs were performed on the three factor scores. Results showed a significant effect of EAT for Factor 2. Participants reported more time spent by their dyad comparing emotions and imagining reactions to emotions in the EAT condition ($M = 0.35$, $SD = 0.96$) than in the control condition ($M = -0.36$, $SD = 0.84$), $F(1, 53) = 8.60$, $p < .01$, partial $\eta^2 = 0.14$. There was also a significant EAT by Gender interaction for Factor 3, $F(1, 53) = 6.51$, $p = .01$, partial $\eta^2 = 0.11$. Post-hoc tests revealed that women reported spending more time arguing and building on the other's ideas in the EAT condition ($M = 0.50$, $SD = 0.83$) than in the control condition ($M = -0.18$, $SD = 1.06$); this difference was marginally significant, $t(1, 28) = 1.97$, $p = .057$. The reverse pattern occurred for men (EAT: $M = -0.48$, $SD = 0.51$, control: $M = 0.05$, $SD = 1.02$), but this difference was not significant, $t(1, 25) = -1.65$, $p = .11$. Moreover, there was a significant difference between women and men in terms of perceived time spent arguing and building on the other's ideas ($\text{time}_{\text{women}} > \text{time}_{\text{men}}$) in the EAT condition, $t(1, 27) = 3.71$, $p < .001$; no significant difference occurred between men and women in the control condition, $t(1, 26) = -0.57$, $p = .57$.

Q3: Relation between Self-Reported Emotions and Perceptions of Interaction

Correlational analyses were performed to test the relation between the participants' ratings of the intensity of positive/negative emotions after collaboration, and their perceptions of the interaction with their partner (the 3

extracted factors presented in Table 1 were used here). Pearson's correlations were conducted across and within conditions (i.e., EAT and control conditions). The analysis across conditions showed that the perceived intensity of positive emotions after collaboration was positively correlated with the perceived amount of time spent (a) comparing emotions and imagining reactions to emotions (F2: $r = .41, p < .05$), (b) arguing and building on the other's ideas (F3: $r = .49, p < .05$). These positive correlations were significant only in the EAT condition (F2: $r = 0.43$; F3: $r = 0.58$). In the control condition, the perceived intensity of negative emotions after the interaction was positively correlated with the perceived amount of time spent comparing emotions and imagining reactions to emotions (F2: $r = .65, p < .05$).

Table 1. Interaction perception items and their factor loading via FA with promax rotation (pattern matrix)

		F1	F2	F3
What is the frequency with which you (own-activity items)	provided <i>your</i> own points of view			.64
	defended and argued <i>your</i> own ideas			.70
	imposed <i>your</i> own points of view			
	challenged <i>your partner's</i> ideas			
	understood <i>your partner's</i> points of view			
	built up on <i>your partner's</i> ideas			.56
	understood <i>your partner's</i> emotions		.51	
	communicated on <i>your partner's</i> emotions	.95		
	adapted <i>your</i> behavior to <i>your partner's</i> emotions	.51		
	understood <i>your</i> own emotions			
	communicated on <i>your</i> own emotions	.66		
	adapted <i>your</i> behavior to <i>your</i> own emotions			
	imagined <i>your partner's</i> reactions to <i>your</i> emotions		.85	
	compared <i>your</i> emotions to <i>your partner's</i> emotions		.94	
	appeared able to control <i>your</i> own emotions		.63	
What is the frequency with which your partner (partner-activity items)	provided <i>his/her</i> own points of view			.58
	defended and argue <i>his/her</i> own ideas			.63
	imposed <i>his/her</i> own points of view			
	challenged <i>your</i> ideas			
	understood <i>your</i> points of view			.82
	built up on <i>your</i> ideas			.61
	understood <i>your</i> emotions			
	communicated on <i>your</i> emotions	.92		
	adapted <i>his/her</i> behavior to <i>your</i> emotions	.53		
	understood <i>his/her</i> own emotions	.69		
	communicated on <i>his/her</i> own emotions	.66		
	adapted <i>his/her</i> behavior to <i>his/her</i> own emotions	.57		
	imagined <i>your</i> reaction to <i>his/her</i> own emotions		.63	
	compared <i>his/her</i> emotions to <i>your</i> emotions		.72	
	appeared able to control <i>his/her</i> own emotions			

Note: F1 = to communicate on emotions and adapt to emotions, F2 = to compare emotions and imagine reactions to emotion, F3 = to argue and build upon the other's ideas.

Discussion and Conclusion

Successful collaborative work requires group members to build a shared understanding (Roschelle & Teasley, 1995) which is contingent upon effective awareness of what their partners know, believe, feel, do and intend to do (Sangin et al., 2011). While recent research on computer-mediated collaboration and CSCL has focused on the effect of the awareness of the partner's knowledge (Dehler et al., 2011; Molinari et al., 2009; Sangin et al., 2011) or activity level (Janssen et al., 2010), little attention has been given on the role of the awareness of the partner's emotions (Eligio et al., 2012). On the one hand, emotions are now recognized as being strongly related to individual and group processes (D'Mello & Graesser, 2012; Van Kleef, 2009). On the other hand, collaboration is challenging at the socio-emotional level (Järvenoja, 2010) since it consists of continuous fluctuations of tension and relaxation between group members who alternate between performing the task and maintaining a good working relationship (Andriessen et al., 2011; Baker et al., 2009). One may therefore expect

that the awareness of the partner's emotions can play a determinant role in collaborative work/learning situations as it contributes to adaptive and regulatory social processes.

In this study, an emotion awareness tool (EAT) was developed which provided awareness about the emotions felt by the partner during a remote (computer-mediated) collaborative design task. Our objective was to investigate to which extent the explicit emotional feedback could impact participants' subjective perceptions of their collaborative work experience. With respect to our first question, results showed no effect of the EAT on participants' perceptions of their own emotions and their partner's emotions after collaboration. In other terms, raising emotion awareness between collaborative partners by giving them the possibility to explicitly share their emotions does not seem to influence their perception of emotional states after collaboration. Our second question concerned the effect of the EAT on participants' perceptions of the quality of their interaction with their partner. We also examined the relationship between participants' perception of emotional states after collaboration and their perception of the quality of interaction, and also how the EAT can impact this relation (Question 3). On the one hand, self-reported measures indicated that the EAT stimulated group members to engage in a process of mutual modeling of emotions during interaction – which is a process by which they compared their respective emotions and also anticipated their respective reactions to emotions. Moreover, the perceived amount of time spent on emotion modeling process was positively correlated with the perceived intensity of positive emotions after collaboration in the EAT condition. This is consistent with results found in the study conducted by Eligio et al. (2012) where participants reported experiencing more positive affect after sharing their emotions during computer-mediated collaboration. In contrast, in the condition where participants did not explicitly share their emotions (control condition), the social comparison of emotions was related to negative emotional states after collaboration. On the other hand, it seems that the effect of the EAT on the perceived quality of interaction varied depending on gender. The EAT had a positive effect on the perceived amount of time spent by the dyad confronting and arguing points of view, building up on the other's ideas – or in other words, on the perceived *degree of transactivity* (which has been found to be positively correlated with collaboration outcomes; Weinberger et al. 2007) – but only for women. There was no significant difference in terms of perceived degree of transactivity between women and men in the control condition. Therefore, these results do not suggest any significant effect of emotion awareness on men's perceived degree of transactivity. Finally, we found that the perceived degree of transactivity was significantly correlated with positive emotional states after collaboration, but only in the EAT condition. In the condition where participants explicitly shared their emotions with their partner, the more they reported focusing and contributing to the ideas proposed by their partner, the more they reported positive emotional states after collaboration.

In our study, results support the hypothesis of a beneficial effect of an emotion awareness tool on (perceived) collaboration – but only for dyads of women. This was also the case in the study of Eligio et al. (2012) who found a positive effect of sharing emotions on group performance, but only with women (since men were excluded from their study to avoid the controversy of gender differences). In our study, men who shared their emotions with their partner during interaction reported being engaged in less partner-focused interactions. Previous research (Brody, 2000) showed that compared to women, men are much less likely to display emotions. This can be explained by socio-cultural beliefs that the display of emotions (in particular, negative emotions) is related to less interpersonal power (Ragins & Winkel, 2011). One may thus expect that forcing men to display and share their emotions during a collaborative work could inhibit them from performing the task with their collaborative partner.

To conclude, the present study has potential to enrich our understanding of how emotion awareness influences collaboration processes, and how to design group awareness tools as means of improving emotion awareness between coworkers in CSCL settings. It also provides a first step in exploring the effect of emotion awareness in collaborative tasks including a comparing approach for its gender-specific relevance. As a further step in the analysis, we will examine verbal interaction data (captured by the microphone headsets) so as to give a more reliable insight into the quality of the interaction, and we will also assess the effect of the EAT on group outcomes (e.g., the number of intermediate slogans created by the groups). Additional analyses on eye-tracking and physiological data are also in progress to better understand the effect of the EAT. Finally, one main limitation of the study is the difficulty in disentangling the effect of reflecting upon one's own emotions from the effect of sharing one's emotions with the partner. In order to overcome this limitation, a new experiment is planned that will use an additional control group in which participants will state their own emotions through the EAT but will not have any access to their partner's emotion statements.

References

- Andriessen, J., Baker, M. & Van der Puil, C. (2010). Socio-cognitive tension in collaborative working relations. In Ludvigsen, S. Lund, A., Rasmussen, I. & Säljö, R. (Eds.). *Learning across sites: New tools, infrastructures and practices*. London Routledge. The Learning series.
- Brody, L. R. (2000). The socialization of emotion: Display rules, infant temperament, and differentiation. In A. Fischer, (Ed.), *Gender and Emotion* (pp. 24-47). Cambridge: Cambridge University Press.

- Buder, J. (2011). Group awareness tools for learning: Current and future directions. *Computers in Human Behavior*, 27(3), 1114-1117.
- Darnon, C., Muller, D., Schragger, S., Pannuzzo, N., & Butera, F. (2006). Mastery and performance goals predict epistemic and relational conflict regulation. *Journal of Educational Psychology*, 98(4), 766-776.
- Davis, M. A. (2009). Understanding the relationship between mood and creativity: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 108(1), 25-38.
- Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior*, 27(3), 1068-1078.
- Derks, D., Fischer, A. H., & Bos, A. E. R. (2008). The role of emotion in computer-mediated communication: A review. *Computers in Human Behavior*, 24, 766-785.
- D'Mello, S. K. & Graesser, A. C. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, 22, 145-157.
- Eagly, A. H., & Wood, W. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin*, 17, 306-315.
- Eligio, U. X., Ainsworth, S. F., & Crook, C. K. (2012). Emotion understanding and performance during computer-supported collaboration. *Computers in Human Behavior*, 28, 2046-2054.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). *Emotional contagion*. New York: Cambridge University Press.
- Järvenoja, H. (2010). *Socially shared regulation of motivation and emotions in collaborative learning* (Doctoral Dissertation). University of Oulu: Acta Universitatis Ouluensis.
- 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(3), 463-481.
- Kapoor, A, Burleson, W, & Picard, R. W. (2007). Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65, 724-736.
- Kapur, M. (2009). Productive Failure in mathematical problem solving. *Instructional Science*, 38(6), 523-550.
- Lee, E. J. (2003). Effects of “gender” of the computer on informational social influence: the moderating role of task type. *International Journal of Human-Computer Studies*, 58(4), 347-362.
- Lehman, B., D'Mello, S. K., & Graesser, A. C. (2012). Confusion and Complex Learning during Interactions with Computer Learning Environments. *The Internet and Higher Education*, 15(3), 184-194.
- Linnenbrink-Garcia, L., Rogat, T. K., & Koskey, K. L. K. (2011). Affect and engagement during small group instruction. *Contemporary Educational Psychology*, 36, 13-24.
- Janssen, J., Erkens, G., & Kirschner, P. A. (2010). Group awareness tools: It's what you do with it that matters. *Computers in Human Behavior*. Elsevier.
- MacLeod, C. M., Koster, E. H., & Fox, E. (2009). Whither cognitive bias modification research? Commentary on the special section articles. *Journal of Abnormal Psychology*, 118, 89-99.
- McRae, K., Ochsner, K.N., Mauss, I., Gabrieli, J.D.E., Gross, J.J. (2008). Gender differences in emotion regulation: An fMRI study of cognitive reappraisal. *Group Processes and Intergroup Relations*, 11, 143-162.
- Molinari, G., Sangin, M., Dillenbourg, P., & Nüssli, M-A. (2009). Knowledge interdependence with the partner, accuracy of mutual knowledge model and computer-supported collaborative learning. *European Journal of Psychology of Education*, 24(2), 129-144.
- Picard, R. W. (1999). *Affective Computing for HCI*. Proceedings HCI. Munich: Germany.
- Ragins, B. R., & Winkel, D. E. 2011. Navigating the emotional battlefield: Gender, emotion and power in work relationships. *Human Resources Management Review*, 21, 377-393.
- Rime, B., Mesquita, B., Philippot, P., & Boca, S. (1991). Beyond the emotional event: six studies on the social sharing of emotions. *Cognition and Emotion*, 5, 435-465.
- Roschelle, J., & Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In O'Malley, C.E., (Ed.), *Computer Supported Collaborative Learning* (pp. 69-97). Springer-Verlag, Heidelberg.
- Sangin, M., Molinari, G., Dillenbourg, P., & Nüssli, M-A. (2011). Facilitating peer knowledge modeling: effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in Human Behavior*, 27(3), 1059-1067.
- Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44(4), 693-727.
- Van Kleef, G. A. (2009). How emotions regulate social life: The emotions as social information (EASI) model. *Current Directions in Psychological Science*, 18, 184-188.
- van Knippenberg, D., Kooij-de Bode, J.M. & van Ginkel, W.P. (2010). The interactive effects of mood and trait negative affect in group decision making. *Organization Science*, 21(3), 731-744.
- Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning & Instruction*, 17, 416-426.

Knowledge Organization with Multiple External Representations in an Argumentation Based Computer Supported Collaborative Learning Environment

Bahadir Namdar, University of Georgia, Athens, GA, USA, baha@uga.edu
Ji Shen, University of Miami, Coral Gables, FL, USA, j.shen@miami.edu

Abstract: Collaboration is one of the core practices in science education (NRC, 1996; NRC, 2012) that has been built into many technology-enhanced learning environments to promote deep understanding (Manlove, Lazonder, & De Jong, 2009). Typically, these environments provide multiple external representations (MERs) for students to understand and communicate scientific knowledge. However, little is known about how students organize knowledge in MERs when they engage in collaborative argumentation. In this study, we designed an argumentation based science unit in a computer supported collaborative learning environment. We investigated how this learning environment affected students' knowledge organization and argumentation on the socio-scientific issue of nuclear energy. We found that the students used all the available representational modes in the environment to make sound arguments and it appeared that the textual representation knowledge entries were the most linked nodes in the knowledge web the students produced as a group.

Introduction

Students nowadays access information in the forms of multiple external representations (MERs) such as computer models, dynamic and static pictures, and texts through information communication technologies (ICT). Research suggests that students benefit from MERs when learning complex scientific phenomena and processes (Ainsworth, 2006; Kozma, 2003). MERs can be used in the classrooms as pedagogical tools to promote scientific argumentation. The quality of MERs "becomes the focal point of the discussion in the classroom as students evaluate and critique methods, explanations, evidence, and reasoning" (Sampson & Clark, 2009, p.450). However, it is still unclear how students organize MERs to make coherent scientific arguments (Erduran, 2012).

In this study, we designed and implemented an argumentation based computer supported collaborative learning (CSCL) unit on socio-scientific issues (SSI) in order to study the connection between students' knowledge organization through MERs and their scientific argumentation processes. Through the study, we addressed the following research questions:

1. How does an argumentation-based CSCL environment affect learners' collaborative knowledge organization with MERs?
2. How does students' collaborative knowledge organization with MERs affect their argumentation on a given socio-scientific issue in a CSCL environment?
 - a) What is the most prominent type of MERs students mostly rely on when they collaboratively organize their knowledge on a given SSI in an argumentation based CSCL environment?
 - b) What is the process of using MERs when they argue on a given SSI in an argumentation based CSCL environment?

Theoretical Framework

Knowledge Organization

In today's technology mediated society, information is easily generated, distributed and accessed through ICT. Since there is almost no limit on accessing information, education should help students develop competencies in searching for, archiving, using, and generating relevant information and organizing it in ways that make sense for current or future use. Our conceptualization of knowledge organization stems from the literature in learning sciences including MERs, knowledge integration (Linn, 2006) and knowledge building (Scardamalia & Bereiter, 2006).

People understand, communicate, and organize information in a variety of modalities including actions, verbal explanations, written texts, physical experiments, computer models, and static and motion pictures and diagrams. The use of MERs can help capture learners' interest (Ainsworth, 1999), and enhance their understanding of science concepts (e.g., Chandrasegaran, Treagust, & Mocerino, 2011; Waldrup, Prain, & Carolan, 2010).

Linn and colleagues developed the knowledge integration framework that emphasizes students' abilities in establishing connections among scientific ideas (Linn & Eylon, 2011; Linn, 2006). The framework

promotes coherent understanding by encouraging students to elicit, add, distinguish, and sort out ideas. The framework has been used to develop a rich library of technology enhanced curricula supporting student learning in science (Linn, Lee, Tinker, Husic, & Chiu, 2006; Liu, Lee, Hofstetter, & Linn, 2008; Slotta & Linn, 2009).

Knowledge building advocates for creating communities of learners who build knowledge together (Scardamalia & Bereiter, 2006). Members of a knowledge community use epistemic artifacts (Sterelny, 2005) such as abstract or concrete models to reflect their understanding. Since these epistemic artifacts are fundamental components of the knowledge building process, organizing them in a coherent way is the key to developing scientific understanding for both individual and groups of learners.

We define knowledge organization as the individual or collaborative processes of searching, sorting out, archiving, and externalize knowledge in a systematic way to achieve a better understanding of the world and to prepare for future learning. It differs from knowledge integration (Linn, 2006) in the sense that it focuses on tagging and organizing knowledge that is externally represented, but not necessarily conceptually integrated in the mind of the learner at the moment. Knowledge organization can be personnel; therefore, the organization structure may not be aimed for public use, which is essential in knowledge building (Scardamalia & Bereiter, 1994). A knowledge organization process of an individual or a group could be intertwined with and complement to knowledge integration and knowledge building.

Collaboration

Collaboration has been advocated as one of the core practices of science education in national science education policy documents over the past two decades (NRC, 1996; NRC, 2012). Collaboration is also a core component of inquiry activities (Simons & Clark, 2005) as it engages students to knowledge construction and delve into their own understanding of scientific phenomena (Komis, Ergazaki, & Zogza, 2007). Additionally, when students involve in collaborative learning, they encounter vast amount of distinct ideas and views, which urges them to organize and integrate those ideas (Linn, 2006). Research indicates that students achieve higher learning goals when they collaborate comparing to individual learning (see, for example, Cohen & Scardamalia, 1998; Lou, Abrami, & D'Apollonia, 2001).

Collaboration has been built into many technology-enhanced learning environments to promote deep understanding (Manlove et al., 2009). As a rising field, computer supported collaborative learning (CSCL) studies how people learn together with the help of computers (Stahl, Koschmann, & Suthers, 2006). It is different from traditional types of learning as it is "concerned with collaborative meaning making processes that go beyond information sharing among multiple people" and "highlights the potential impact of social community through computers as vehicles for transforming activity procedures" (Yoon & Brice, 2011, p.251).

Scientific Argumentation

There are several reasons for practicing argumentation in science classrooms, especially on SSI related topics (e.g., Sadler & Donnelly, 2006; Zohar & Nemet, 2002). As Duschl and Osborne (2002) argued "situating argumentation as a critical element in the design of science learning environments both engages learner with the conceptual and epistemic goals and, for the purposes of the practice of formative assessment by teachers, can help make science thinking visible" (p. 44). Practicing argumentation allows students to use available data and evidence to construct knowledge, clarify meanings, and reflect on their own thinking (Duschl & Osborne, 2002). In addition to making thinking visible, Jimenez-Aleixandre and Erduran (2007) propose five advantages of using argumentation in science classrooms: (1) the access to students' cognitive and metacognitive processes, (2) the development of discourse practices and thus critical thinking, (3) increased scientific literacy, (4) enculturation into scientific culture and the development of epistemic criteria, and (5) developing reasoning and rational criteria. CSCL may incorporate scientific argumentation. Noroozi, Weinberger, Biemans, Mulder, and Chizari (2012) did an extensive review of literature on argumentation based CSCL and found that these environments foster in-depth discussions (Andriessen, Baker & Suthers, 2003), and help learners achieve deeper understanding and productive arguments (Buckingham-Shum, 2003). SSIs are particularly useful for promoting interest in learning science (Sadler & Zeidler 2004; Zeidler, Walker, Ackett, & Simmons, 2002). Research indicates that facilitating argumentation through the use of SSI can increase students' use of scientific knowledge in constructing arguments (Zohar & Nemet, 2002).

In sum, in an active and collaborative learning environment, individual learners work with each other to learn and co-construct scientific knowledge. Their understanding will evolve through processes including re-organization of information, representation of knowledge in MERs, and cogeneration of argumentative discourse. So will the knowledge artifacts produced by the group of learners. With this vision in mind, we build and test a new learning environment, the innovative knowledge organization system (iKOS), which will be described in the following.

Methods

consented to participate in the study. We only collected data from the students who consented. The class met once a week (2.5hrs/wk), and the unit was carried out in three consecutive weeks (only the second week was entirely devoted to this unit). In week one, it took the class about 45 minutes to go through the iKOS system. They were then assigned as homework to create individual entries, one per mode. In week two, the students were divided into four small groups (2 with four students each and 2 with five students each) and carried out the bulk of the activity (~150min). Also, two groups were able to present their work during week two. In week three, the two remaining groups presented their findings and we carried out a whole class discussion (~35min).

Data Collection and Analysis

Our data collection included participant observation (Suzuki, Ahluwalia, Arora, & Mattis, 2007), video recording, and iKOS entries and descriptive statistics generated through the iKOS system. We were conscious that video recording might affect the students' responses and behavior in the learning environment so we tried to move the camera as little as possible to keep the distraction at a minimum level.

To investigate our first research question, we focused on examining the patterns of the iKOS entry networks created by the students. We considered each individual iKOS entry as a node, and the links between the nodes were categorized into two types: direct links and indirect links. A direct link is defined as a connection between two entries when the title of one entry is a keyword of the other; an indirect link is defined when two entries share one or more keywords.

Applying the social network analysis techniques, we calculated the degree centrality for each iKOS entry which measures the extent to which one node is connected with the rest of the knowledge web (Knoke & Yang, 2007). In the iKOS system, we summed all the links associated with one entry; to get the normalized degree centrality, we divided this number by the possible number of links this entry has in the knowledge web. We then calculated the mean normalized degree centralities associated with the particular representational modes and the whole knowledge web.

In order to investigate our second research question, we transcribed the videos verbatim and analyzed students' final presentations. Our unit of analysis was a coherent statement made by a student that delivers a stand-alone meaning. We coded each relevant statement in terms of its direct relevance to the type of iKOS entry for each group and we summed up these numbers and reported the percentages. In addition to the quantified results, we also reported a representative case in order to provide a more vivid scenario for the reader to make sense of the learning processes (Creswell, Hanson, Clark Plano, & Morales, 2007).

Results

Network of Students' iKOS Entries

The students created 20 event 19 wiki and 24 concept map entries for the homework assignment and created additional 4 event, 4 wiki, and 4 concept map entries during the class activity in the second week. Table 1 shows the total number of links for the entries in each of the three iKOS modes. Apparently, indirect links are much more pervasive than direct links between entries. In terms of total links, the wiki mode has the highest number.

Table 1: Degree Centralities of iKOS Entry Modes

Types of Links	Event	Wiki	Concept Map
Direct	11	20	29
Indirect	402	467	389
Total	413	487	418

Table 2 shows the mean normalized degree centrality for the iKOS knowledge web, in which the entries were linked either directly or indirectly. Our results indicated that the wiki mode had the highest mean degree centrality, for both the directly and indirectly linked knowledge webs. This suggests that the most "centralized" entry mode was the wiki mode in this learning unit. This was not expected because students had to manually type in keywords for the wiki mode, while in other modes the computer will automatically register the keywords. This result suggests that the students were mindful of and good at generating their wiki entry keywords.

Table 2: Mean Normalized Degree Centrality (NDC) for iKOS Entries in terms of Direct and Indirect Links

Types of Links	Mode	Mean NDC	SD
Direct	Event	0.009	0.015
	Wiki	0.013	0.031

	Concept Map	0.010	0.030
	Overall	0.009	0.015
Indirect	Event	0.245	0.218
	Wiki	0.291	0.179
	Concept Map	0.261	0.201
	Overall	0.265	0.199

Knowledge Organization and Argumentation

We analyzed students' arguments on the final presentation sessions. Table 3 shows the percentages of each groups' argumentation relevant to their MERs. Apparently, different groups relied on different types of representations in their arguments. Overall, there was a significant relationship for groups and representational modes, meaning that categorical variables were not independent and groups relied on to certain modes more on their arguments (Chi square test for independence, $\chi^2 = 27.6$, $df=6$, $p < 0.001$).

Table 3: Distribution of Knowledge Organization Entries in Collaborative Argumentation

Representation Mode	Relevance of Statements to Representation Types			
	Group 1	Group 2	Group 3	Group 4
Wiki	37%	42%	30%	11%
Event	55%	21%	13%	28%
Concept Map	8%	37%	57%	61%

Although the groups referred to the representation modes quite differently in their argumentation, through the analysis of group discussions and presentations, we found that students incorporated MERs in a holistic manner to present a sound argument on the given SSI. In the following we present a case to illustrate this point.

Group 1 focused on the environmental impact of the nuclear energy in general and the group was against the construction of the nuclear power plant and relying on nuclear energy as an energy source. A total of 55% of their statements were based on the event entry they created. However, they initiated their argument based on the wiki entry they created. In their wiki entry they stated that they will focus on the "environmental impacts associated with the disposal of effluent wastewater produced by a secure and functioning nuclear power facility" and described the effluent water waste in the process of cooling nuclear rods (Group1, Wiki Entry). They initiated the discourse based on their definitions and knowledge organization about the effluent water waste:

Derrick: We are looking actually at the water which is most of the time not retaining and injected right back into environment. Ok? This water typically starts as sewage. They take sewage water, remove contaminants, cool the nuclear rods, and then ship it right back at ocean, river or lakes. So what are the impacts associated with that? Number one we super-heated the water. And so... we also have environmental impact there as far as changing ecosystems.

These statements made by Derrick also reflected his group's collaborative knowledge organization in the Concept map mode. Because in their concept map, they illustrated that nuclear waste flows to water and the water temperature rises. Ray, on the other hand, changes the direction of the conversation using the event entry they created.

Ray: This [shows their event entry to the class] is just a representation that you [Derrick] are talking about how the water is either evaporated and this contaminants get into the atmosphere or they are directly injected into the water system. And the end result is it affects humans directly and the atmosphere and plant life animal life affected negatively as well.

Derrick: It also represented in the bio magnification. Because when you start to impact what we eat that is magnified to us tenfold twentyfold. So there is a lot of contamination out there that we are not looking at when we just think about nuclear power.

Here Ray showed the part of their pictorial representation where contaminants from the water go through the irrigation process to grass and then to the cows and then humans. This argument was also reflected in their concept map, which showed that nuclear waste can leak to underwater repository and directly harms people's or animals' lives.

The discourse above illustrates that although this group relied mostly on the event mode, they also incorporated all the other representation modes in order to drive their argument to the point that they stated possibilities of how effluent water waste can impact the environment. Derrick was a representative when they were presenting their argument. Based on their knowledge organization practices with using MERs Derrick concluded that:

We don't think there are enough standards or safe guards in place based on what just happened in Japan, which is supposed to be the technology leaders. As far as natural disasters protecting yourself against natural disaster we do not feel there is enough safeguards in place yet put one in our own backyard.

This case shows how students' knowledge organization process affects their discourse and arguments. In this case, although the students' argumentation benefited from one particular type of representational mode, further analysis of their argumentation indicated that the students used all the three representational modes in the proper places of the discourse to arrive at their conclusion and make sound arguments with the help of data they put in MERs.

Discussion and Conclusion

Recently, science education field has been witnessed the growing body of research advocating for using those representations to capture learners' interest (Ainsworth, 1999), and capture and enhance students' understanding of science concepts (e.g., Chandrasegaran, Treagust, & Mocerino, 2011; Waldrip et al., 2010). Based on the positive effects of using MERs in science learning, we were interested in students' collaborative knowledge organization practices with MERs on socio scientific issues. Our results indicated that degree centrality of students collaborative knowledge entries were highest in their wiki entries. We were hypothesizing that the concept maps would be most prominent representations in terms of degree centralities for indirect links since the computer automatically generates key words and links the entries. However, our results showed that students made their knowledge organization process more interconnected in the wiki mode, because students generated better key words to connect their entries with other entries. Also, our interpretation is that even though students can represent their knowledge in the forms of concept maps and pictorial representations, they may still feel the need for verbally explaining their findings to make their knowledge organization and understanding visible for their audience.

We were also interested in seeing how students' knowledge organization relied on the use of MERs when students argue on a given SSI. Our results indicated that students used different modes of representations in their arguments. We want to point out that even though this result might be interpreted as students with different learning abilities use different representation in their arguments, our analysis of students' discourse illustrated that the students used all three representational modes in their arguments to produce a sound argument based on their collaborative knowledge organization with MERs. Therefore, we conclude that students can create MERs in our system when they organize their knowledge and make their arguments based on multiple data sources they put in those MERs.

Although our learning unit is promising from several aspects, it also has several limitations. First, our intervention is short for making general judgments on students' collective learning using iKOS. Secondly, we did not analyze the student-generated arguments in terms of their structural, content, and justification qualities. In the future we will analyze students' collaborative arguments with a larger sample size to make judgments about the effect of our learning unit. It is our hope that teachers and students will be able to use this tool in order to make quality arguments, enhance deep understanding of the scientific concepts, and organize scientific knowledge in a more coherent way to learn complex scientific phenomena in a more holistic manner, and collaborate and assess their peers' entries in the system.

References

- Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, 33(2-3), 131–152. doi:10.1016/S0360-1315(99)00029-9
- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198. doi:10.1016/j.learninstruc.2006.03.001
- Andriessen, J., Baker, M., & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn:*

- Confronting cognitions in computer-supported collaborative learning environments* (pp. 1–25). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Buckingham-Shum, S. (2003). *The roots of computer supported argument visualization, Visualizing argumentation: Software tools for collaborative and educational sense-making*. London: Springer-Verlag.
- Chandrasegaran, B. A. L., Treagust, D. F., & Mocerino, M. (2011). Facilitating high school students' use of multiple representations to describe and explain simple chemical reactions. *Teaching Science*, 57(4), 13–21.
- Cohen, A., & Scardamalia, M. (1998). Discourse about ideas: Monitoring and regulation in face-to-face and computer-mediated environments. *Interactive Learning Environments*, 6, 93–113.
- Creswell, J. W., Hanson, W. E., Clark Plano, V. L., & Morales, a. (2007). Qualitative Research Designs: Selection and Implementation. *The Counseling Psychologist*, 35(2), 236–264. doi:10.1177/0011000006287390
- Duschl, R. A., & Osborne, J. (2002). Supporting and Promoting Argumentation Discourse in Science Education. *Studies in Science Education*, 38(1), 39–72. doi:10.1080/03057260208560187
- Erduran, S. (2012, October). Debunking folk myths on food: Incorporating argumentation in the context of food science education. The University of Georgia, Athens GA, USA.
- Jimenez-Aleixandre, M. P., & Erduran, S. (2007). Argumentation in science education: An overview. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 3–27). Dordrecht, Netherlands: Springer.
- Knoke, D., & Yang, S. (2007). *Social Network Analysis (Quantitative Applications in the Social Sciences)*. (T. Liao, Ed.) (p. 133). Sage Publications, Inc.
- Komis, V., Ergazaki, M., & Zogza, V. (2007). Comparing computer-supported dynamic modeling and “paper & pencil” concept mapping technique in students' collaborative activity. *Computers & Education*, 49(4), 991–1017. doi:10.1016/j.compedu.2005.12.007
- Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, 13(2), 205–226. doi:10.1016/S0959-4752(02)00021-X
- Linn, M. C. (2006). The Knowledge Integration Perspective on Learning and Instruction. In R. Sawyer, Keith (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 243–264). New York, NY: Cambridge University Press.
- Linn, M. C., & Eylon, B.-S. (2011). *Science learning and instruction: Taking advantage of technology to promote knowledge integration* (p. 340). Florence, KY: Routledge, Taylor & Francis Group.
- Linn, M. C., Lee, H., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and Assessing Knowledge Integration in Science. *Science*, 313(5790), 1049–1050. doi:10.1126/science.1131408
- Liu, O. L., Lee, H.-S., Hofstetter, C., & Linn, M. (2008). Assessing Knowledge Integration in Science: Construct, Measures, and Evidence. *Educational Assessment*, 13(1), 33–55. doi:10.1080/10627190801968224
- Lou, Y., Abrami, P. C., & D'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, 71, 449–521.
- Manlove, S., Lazonder, A. W., & De Jong, T. (2009). Collaborative versus individual use of regulative software scaffolds during scientific inquiry learning. *Interactive Learning Environments*, 17(2), 105–117. doi:10.1080/10494820701706437
- Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2012). Argumentation-Based Computer Supported Collaborative Learning (ABCSSL): A synthesis of 15 years of research. *Educational Research Review*, 7(2), 79–106. doi:10.1016/j.edurev.2011.11.006
- Novak, J. D. & A. J. Cañas, (2008). The theory underlying concept maps and how to construct them, technical report IHMC CmapTools 2006-01 Rev 01-2008, Florida Institute for Human and Machine Cognition, available at: <http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf>
- Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*, 41(5), 513–536. doi:10.1002/tea.20009
- Sadler, T. D., & Donnelly, L. A. (2006). Socioscientific argumentation: The effects of content knowledge and morality. *International Journal of Science Education*, 28(12), 1463–1488. doi:10.1080/09500690600708717
- Sadler, T. D., & Zeidler, D. L. (2004). The morality of socioscientific issues: Construal and resolution of genetic engineering dilemmas. *Science Education*, 88, 4–27. doi:10.1002/sci.10101.
- Sampson, V., & Clark, D. (2009). The impact of collaboration on the outcomes of scientific argumentation. *Science Education*, 93(3), 448–484. doi:10.1002/sci.20306
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer, R. (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 97–115). New York, NY.

- Simons, K., & Clark, D. (2005). Supporting Inquiry in Science Classrooms with the Web. *Computers in the Schools*, 21(3-4), 23–36. doi:10.1300/J025v21n03_04
- Slotta, J., & Linn, M. C. (2009). *WISE Science: Web-Based Inquiry in the Science classroom*. New York, NY: Teachers College Press.
- Stahl, G., Koschmann, T., & Suthers, D. D. (2006). Computer-supported collaborative learning. In R. Sawyer, K. (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 409–426). New York, NY: Cambridge University Press.
- Sterelny, K. (2005). Externalism, epistemic artefacts and the extended mind. In R. Schantz (Ed.), *The Externalist Challenge: New Studies on Cognition and Intentionality*. Berlin: de Gruyter.
- Suzuki, L. A., Ahluwalia, M. K., Arora, A. K., & Mattis, J. S. (2007). The Pond You Fish In Determines the Fish You Catch: Exploring Strategies for Qualitative Data Collection. *The Counseling Psychologist*, 35(2), 295–327. doi:10.1177/0011000006290983
- Waldrip, B., Prain, V., & Carolan, J. (2010). Using Multi-Modal Representations to Improve Learning in Junior Secondary Science. *Research in Science Education*, 40(1), 65–80. doi:10.1007/s11165-009-9157-6
- Yoon, J., & Brice, L. (2011). Water Project: Computer-Supported Collaborative E-Learning Model for Integrating Science and Social Studies. *Contemporary Educational Technology*, 2(3), 250–263.
- Zeidler, D. L., & Nichols, B. H. (2009). Socioscientific issues: Theory and practice. *Journal of Elementary Science Education*, 21(2), 49–58. doi:10.1007/BF03173684
- Zeidler, D. L., Walker, K. a., Ackett, W. a., & Simmons, M. L. (2002). Tangled up in views: Beliefs in the nature of science and responses to socioscientific dilemmas. *Science Education*, 86(3), 343–367. doi:10.1002/sci.10025
- Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39(1), 35–62. doi:10.1002/tea.10008

Acknowledgements

This material is based upon work supported by the Office of STEM Education at The University of Georgia. Any opinions, findings and conclusions expressed in this poster are those of the authors and do not necessarily reflect the views of the Office.

Epistemic Trajectories: Mentoring in a Game Design Practicum

Padraig Nash, University of California-Berkeley, pnash@berkeley.edu
David Williamson Shaffer, University of Wisconsin-Madison, dws@education.wisc.edu

Abstract: We use two constructs to examine mentoring in this practicum. *Epistemic frames*—the configurations of the skills, knowledge, identities, values, and epistemologies that professionals use to think in innovative ways—provide a model for looking at professional expertise (Shaffer, 2006). Building on epistemic frames is *epistemic network analysis* (Shaffer, et al., 2009), a method for quantifying changes in epistemic frames (Shaffer, 2010). Our claim here is that the mentor, using Schön’s “Follow me!” coaching model (1987), leads the team on a path that illuminates the nature of learning to think professionally and on the function of a mentor in that process.

Introduction

There is growing concern that the 20th century mode of education, with its focus on problems with standardized answers, is leaving increasing numbers of young people unprepared for the future (Collins & Halverson, 2009). The professions—and more specifically, the processes of creating new professionals—offer an alternative model for educational activities better aligned with the abilities required to navigate a complex and changing future (Gee & Shaffer, 2010; Shaffer, 2006). Innovative professionals, who face nonstandard problems that come up in practice, often learn their particular way of solving problems in simulations of professional practice, such as apprenticeships and practica (Goodwin, 1994; Schön, 1987; Sullivan, 1995).

Recent work has described the pedagogical and developmental benefits of apprenticeship-based learning models in which young people are guided by mentors (Halpern, 2010; Rose, 2004). Other work looking at the processes by which mentors instill professional ways of thinking is also informed by an interest in how the training practices of professionals can serve as models for the development of technology-supported learning environments for K–12 students (Nash & Shaffer, 2010; Shaffer, 2006)

In this paper, we examine the learning relationship between a mentor and team of college students through an ethnographic study of a game design practicum at a European arts school. We use a novel ethnographic technique, called *Epistemic Network Analysis* (ENA), to analyze the process of mentoring in this professional setting, and argue that such an analysis is useful for further studies of professional education, as well as for studies of apprenticeship-based programs for youth.

Theory

Professional Problem-solving

Professional problem-solving is itself problematic. Problems in real-world practice are ill-formed, with potentially unlimited relevant facts and features (Schön, 1987). Thus, the problem for professionals is not only how to solve problems but how to identify them. Different professions approach ill-defined situations differently. Charles Goodwin calls the shared way that professionals see and categorize their domain “professional vision” (Goodwin, 1994). Professional vision, according to Goodwin, is employed by a community of practitioners who expect from each other a common way of organizing the world that is consistent with the values and methods of the profession. In other words, each profession is a “community of practice” (Lave & Wenger, 1991): a group of people who share similar ways of seeing and solving problems. Professional communities of practice rely on systemically organized professional values, preferences, and norms that inform their ways of seeing and solving problems (Schön, 1987). Thus, to learn to be a professional, one must be initiated into a professional community of practice.

Practica Mentoring

Many professionals join their community of practice via a practicum. In practica, novices “take on real-world projects under close supervision” (Schön, 1987, p. 37). While Schön offers some insight into the nature of the “close supervision” that mentors in practica use to help novices develop professional ways of thinking, there has been relatively little study into the ways that professional mentors do the actual work of mentoring. Schön argues that students cannot be taught, but they can be coached (Schön, 1987, p. 17). At the beginning of a practicum, learners lack both the vocabulary to talk about the work and the experience that would give that language any meaning. As they begin to do the work and talk about it with the mentor, novices imitate the mentor’s talk and actions until it is internalized. Characterized by Schön as a “Follow me!” model of coaching, this way of mentoring is fundamental to the practicum experience (Schön, 1987).

Implicit in the “Follow Me!” coaching model is that the learner ultimately goes to where the mentor is: to the professional vantage where expert decisions are made. Less clear, however, is where the mentor needs to take them within the practicum. If it is the mentor’s job to arrange, as Schön puts it, the “right” kinds of experiences for the learners, by what logic is an experience “right”? Schön is clear that professional expertise must be grasped as a whole in order to be grasped at all. It cannot be learned in a molecular way. If the path to professional expertise is not linear what is the path that is taken? And, what would be the logic of such a path?

Epistemic Frame Theory and Epistemic Network Analysis

In this study, we use two theoretical constructs for examining and measuring the modeling and development of professional thinking: epistemic frame theory and epistemic network analysis. The professional way of seeing the world that Schön describes as “the competence by which practitioners actually handle indeterminate zones of practice” (1987, p. 13) has been further explicated by Goodwin (1994), who describes the ways practitioners highlight and elide things that are important or not according to their professional perspective, and Sullivan (1995), who explains how professionals employ an ensemble that includes intellectual, practical, and ethical components. Shaffer extends these insights by pointing out that it is “the combination—linked and interrelated—of values, knowledge, skills, epistemology, and identity” that characterizes the professional ensemble. This combination, what he calls the *epistemic frame*, emphasize the ways that ways seeing and solving problems are linked in practice (Shaffer, 2010). For example, a reporter may write a certain way because she views her job as serving the societal function of being a community watchdog. In this epistemic frame, a particular journalistic skill is informed by a specific journalistic value and a sense of professional identity.

ENA is a technique for quantifying and analyzing an epistemic frame (Shaffer, et al., 2009). It adapts a social network analysis framework—for mapping social elements—to instead map the constituent elements of complex thinking. In social network analysis, the objects of interest are not the individual actors but the structure of the relationships among those actors. Similarly, epistemic frame theory suggests that complex thinking is not characterized merely by a collection of values, knowledge, skills, epistemology, and identity, but rather by a particular structure of relationships or connections between these components of expert practice. Thus, analyzing thinking in terms of simple counts of frame elements in an (aspiring) professional’s frame is not sufficient to account for the development of expertise. Instead, ENA uses co-occurrence of frame elements in discourse to model the pattern (or patterns) of association characteristic of a particular professional community. By quantifying the relationships between epistemic frame elements, ENA describes and quantifies the structure of an epistemic frame. Where one frame may have one set of important relationships between its constituent elements, another frame may emphasize different relationships. ENA provides a method for examining when and how often frame elements are linked, and can show trends in how epistemic frames change over time, between individuals, or across different interactional contexts (Nash & Shaffer, 2010). ENA is thus a potentially useful tool for comparing the epistemic frames of a team of learners and a mentor in a practicum.

Many of the studies that examine the learning that takes place in a practica rely on surveys (Ryan, Toohey, & Hughes, 1996), which do not measure learning in situ and rely on self-report data, or are qualitative ethnographies (see, for example, Hutchins, 1995). While both of these methods shed light on the processes of learning and of mentoring in a practicum, ENA is a potentially useful addition to the scholarly toolkit because it can quantify thinking in the context of action. In the case of a practicum, the conversations between mentor and learner provide the occasion for the mentor to model professional thinking and for the learner to imitate that modeling (Schön, 1987). Using ENA to examine the reflective conversations in the practicum, we can examine whether (and, more important, how) the team of learners “follows” the mentor, as suggested by Schön’s model.

We operationalize such a question by comparing the “distance” between a mentor’s frame and the learners’ collective frame when they have meetings together across the time of the practicum. By distance, we mean the projection of the structure of the epistemic frame, as quantified by ENA, in to a high dimensional space. Using multi-dimensional scaling, we can visualize epistemic trajectories: models that show the distances between team and mentor’s frames throughout the practicum change over time. Finally, constructs borrowed from social network analysis can help us interpret these trajectories. For example, the relative centrality of a frame element quantifies the extent to which it is connected to other elements in the professional discourse, and therefore can help us interpret the mathematical concept of “distance” by suggesting in what ways two different frames are similar or different.

This study: Game design practicum

The starting point for this study is the idea that in a practicum a mentor leads learners through the authentic work of a professional practice and that though his coaching the learners develop a particular epistemic frame. As in all practica, the students in a game design practicum do the work that professional game designers do. They seek to create games that feature *meaningful play*, which game design experts describe as occurring “when the relationships between actions and outcomes in a game are both discernable and integrated into the larger context of the game” (Salen & Zimmerman, 2004, p. 34). In other words, a critical job for game designers is to

make a game in which players can perceive the immediate outcome of the actions they take—the relationship between action and outcome in games is often called *gameplay* (or *game mechanic*)—and that those outcomes are consistent with the game as a whole. The game as a whole, also known as the game concept, can be simple (e.g. Angry Birds) or complex (e.g. World of Warcraft), and encompasses the vision of the overall player experience of the game. It often includes narrative elements, and in the case of educational or “serious” games, learning goals (Abt, 1970; Gee, 2003; Squire & Jenkins, 2004). Both gameplay and game concept simulate some phenomena that is usually, but not exclusively, real-world phenomena (Salen & Zimmerman, 2004, p. 457). Thus, one important activity in game design is researching the phenomena to be simulated, which are referred to as the *content domain*. Learning to link these three elements of game design—*gameplay*, *game concept*, and *concept domain*—is a fundamental task of building a game.

This study looks at the role of mentoring on the learning trajectory of the team of students as they develop these and other elements of game design’s epistemic frame. Although the study is, of course, situated in an examination of one particular practicum, our hope is to show how this analytical method can reveal learning processes in practicum and similar settings. We ask three questions:

1. Does the team “follow” the mentor in the sense that the team and mentor’s frames during meetings become more similar over time?
2. Where does the team follow the mentor? Specifically, do the team and mentor’s frames converge in linear trajectories?
3. Whether the convergence is linear or not, why does it have the observed form?

To answer these questions, we first conduct a qualitative (ethnographic) analysis of the practicum. Then, we triangulate this analysis in quantitative terms using ENA.

Methods

Setting and participants

The game design practicum took place in an undergraduate level arts school. The semester-long practicum was organized around the production of a single game. Student teams were assigned a client, who provided the team with their assignment. In this case, the team learned that they needed to create a game to encourage consumers to choose sustainable fish to eat. The team was comprised of seven students. The mentor assigned to work with the team is a professional designer and developer of educational software.

Data Collection

We observed every team meeting until the midterm review. Meeting data were collected in digital audio recordings. Recordings were transcribed to provide a detailed record of interactions. We interviewed the team’s mentor twice during the two months, asking general questions about game design, as well as specific questions about the progress of the team and his ideas about his role as a mentor. We conducted the same type of interview with two other mentors assigned to different teams.

Data Analysis: Coding Scheme

From interviews conducted with three mentors at the school, we used a grounded theory approach (Strauss & Corbin, 1998) to generate a set of 32 qualitative codes representing the epistemic frame of game design. Three frame elements were of particular interest in this study: the knowledge of game mechanics, the knowledge of the content domain, and the skill of developing a game concept.

Data Analysis: Segmentation and coding

This study is based on three team meetings at which the mentor was present. We segmented the three meetings into interactive units (“stanzas”) which were defined as sequences of utterances with a consistent topical focus. For example, if the team started discussing their strategy for an upcoming meeting with a client and then switched to discussing the profile of their target users, the switch in discourse topic would indicate two separate interactive units. There were 14 stanzas in the first meeting, 17 in the second, and 24 in the third. For the three team meetings, we coded each stanza for the contributions of both the mentor and the team for articulations of the elements of the game design frame (as generated from the interviews with the three mentors). If discourse in the stanza was determined to meet the criteria of a code’s definition, that stanza was coded with a ‘1’ for that code. If none of the discourse in that stanza met the criteria of the code’s definition, that stanza was coded for a ‘0’ for that code. The team’s contributions were coded collectively: if no team member discussed a given frame element, that stanza would be coded with a 0 for that element, but if at least one team member talked about a given frame element, that stanza would be coded with a 1 for that element.

Data Analysis: Measuring frames

Adjacency Matrices

We created an adjacency matrix for each coded data segment (each stanza of each meeting). Adjacency matrices record the links between individual frame elements. When a pair of elements co-occur within a stanza, they are considered conceptually linked (Shaffer, 2010). We summed each meeting's constituent stanzas' adjacency matrices to construct final adjacency matrices for the team and mentor for each meeting.

Relative Centrality

For the mentor and the team in each meeting, we calculated the relative centrality of each of the frame elements that constitute their epistemic frames. Relative centrality is a measure of how often each element is connected to all of the other elements in discourse; in other words, it is a measure of the relative weight of an epistemic frame's constituent elements (Shaffer, et al., 2009). To compute the *centrality* of a frame element, the square root of the sum of squares of its associations with its neighbors is calculated. To compute the *relative centrality* of an individual frame element, its weight is divided by the frame element with the greatest weight in the network (Shaffer, et al., 2009).

Data Analysis: Comparing frames

Frame Similarity Index (FSI)

The Frame Similarity Index provides a testable measurement of the similarity between two epistemic frames. The Euclidean distance between two frames is calculated by finding the root mean square of the differences of the relative centrality of each frame element between the two frames. Thus, identical frames would have an FSI score of zero. The maximum theoretical distance between two frames is the Euclidean distance between a frame where every element has a relative centrality of 100 (all possible connections) and a frame where every frame element has a relative centrality of 0 (no connections). The maximum theoretical distance between two frames with 32 constituent frame elements would be 565.68. Dividing the distance between the two frames by the maximum possible distance between those two frames provides the distance expressed as a percentage of the maximum possible distance.

To estimate confidence intervals for the difference between two frames, a jackknifing method can be used, in which the relative centralities are systematically recomputed leaving out one stanza at a time. The standard deviation and confidence intervals of the FSI statistic are then estimated from the variability within the calculated subsamples.

Data Analysis: Epistemic frame trajectories

To see all of the team and mentor's meeting frames in relation to each other, and to more accurately compare those frames, we require a more sophisticated model: epistemic frame trajectories. First, the distances between all points of interest are calculated using FSI. In the case of this study, those points of interest are the mentor and team's frames in each of the three meetings. These distances are organized in a symmetric distance matrix, made up of 6 distance vectors. A classical multidimensional scaling (MDS) algorithm is applied to the distance matrix in order to identify the dimensions that capture the most variance in the data. Because MDS does not preserve directionality, specific dimensions in the low dimensional projection are not interpretable (Bartholomew, Steele, Moustaki, & Galbraith, 2008). However, relative position in space is still meaningful in that points closer together in the high-dimensional space have more similar patterns of co-occurrence than points farther apart.

Using these tools, we operationalize our research questions in the following ways:

Research Question 1

Does the team follow the mentor? In other words, do the team and mentor's frames during meetings become more similar over time? To answer this question, we used FSI to measure the distance between the team and mentor's frames in each of the three meetings. We looked to see whether the distance between the team's frame and the mentor's frame was reduced with each successive meeting.

Research Question 2

Where does the team follow the mentor? In other words, do the team and mentor's frames converge in linear trajectories? To answer this question, we constructed epistemic frame trajectories to show the distances between the team and mentor's frames across the meetings in a two dimensional projection of a multidimensional space. We used multi-dimensional scaling (specifically principal coordinates analysis) to create two dimensional trajectories from the six dimensional space of the team and mentor's meeting frames. We used the second and third of the resulting six dimensions to display the two dimensional trajectories. Although the first dimension

captured the most variance, and thus could be argued to be the most important, it appeared to be highly correlated with time. Since we already knew that the meetings were separated by time, we considered the second and third dimensions more revealing.

Research Question 3

Why do the trajectories have the observed form? To explain the logic behind the trajectories, we examined the change in relative centralities of the frame elements across the three meetings for both the team and mentor. Since the positions in the epistemic trajectories are determined by FSI scores, which in turn are calculated by the similarity of the relative centrality of the frame's constituent elements, changes in relative centrality would influence the position of the team and mentor's epistemic frames with each meeting.

Results

We describe our observations of the game design practicum in four parts below. The first part is a qualitative look at the three meetings, and the final three parts quantify what was observed in that qualitative investigation.

Qualitative results

The team had met before their first meeting with the mentor and had brought a list of brainstormed game concepts. The mentor ran the first meeting, and was critical of the team's work. He felt that not enough work had been done, that the team was not adequately organized, and that they were prioritizing the wrong things. When one student observed that they had "mostly just talked a lot," the mentor agreed, telling them to "quit the endless brainstorm and start working." None of the team's game concepts were grounded in either a vision of gameplay or rigorous research; as one student admitted, "There isn't any gameplay yet." In response, the mentor reinforced the link between gameplay and game concept, telling the team that "the challenge of your game has to be in line with your educational challenge."

The mentor was similarly explicit about the link between game concept and the content domain. Noting that the team's research into the content domain was "okay", the mentor urged the team to "do the next step of research... if you want to do the fish tycoon concept that means you have to do a lot of extra research on fish management and business." The main problem with the team's work in this stage was that their game concepts were not connected to game play or a content domain. The mentor sent the team away with instructions "to fill it in. A lot of concepts seem nice, but then you [need to] fill it in."

The second meeting was run by the team. They had, as the mentor had instructed, done "the next step of research." They spent most of the second meeting reporting their research on the content domain. The mentor prompted the team to be explicit about the utility of their new information, at one point asking, "So, what is the main question you need to know for your game with respect to the storing of the fish?" Much of the activity in the meeting was precisely this kind of decision making. For example, a student remarked how "it's very important to be more specific about what fish we're going to use." When the mentor asked whether the team was deciding to abandon fish farms as part of their game concept to focus more specifically on fish marts, one student responded, "we have to map the game completely." Mapping the game completely required the team to describe the concept and game play in terms of the content that they had researched. Although the second meeting focused on content, the mentor still reinforced the link between game mechanic and game concept; as the team was figuring out what to do next, the mentor reinforced again the three elements together, explaining, "it's important to first get your idea of gameplay clear.... to rev up the research.... [and to] elaborate on the design idea" Although all three are important, he suggested that the team now turn their focus on game play.

The third meeting was also run by the team. They were planning on conducting play-tests of some of their game prototypes, in order to get feedback on the game mechanics, and also preparing for a mid-term presentation, at which they needed to present their game concept and reports of the research they had done to support their design. The team discussed a number of prototypes for mini-games within their game concept. The mentor continued to give advice about the three key frame elements. In a discussion about one of the mini-games the team was creating for their game, the mentor talked about the "story" of the game:

The story has to be matching the reality. If you say there's always bad-catch, you can only say that if that's the reality. Or you could say, this is just ordinary fishing.... Then later in the story maybe you say something about environmental, dolphin-friendly tuna. Or you play this game again and the score is really [about how you] find and get no dolphins or throw the dolphins out again.

The game story is a way of talking about the game concept: the sequence of challenges the player must face in order to complete the game and, since this is a "serious game," to learn the intended lesson. Both the sequence and the individual challenges must make sense internally (in terms of the game mechanic) and externally (in terms of the content domain). One student demonstrated these connections in their discourse by using content domain language to describe how the gameplay must be persistent in their concept:

One game is like fishing in the Baltic Sea and avoiding environmentalists. The other is scaring off locals with your bulldozer. The other one we haven't defined yet....in the end, all those things have to be present again.

Result 1: Following the mentor?

The frame similarity index (FSI) was calculated to compare the team's epistemic frame and the mentor's epistemic frame in each of the three meetings. In the first meeting the distance between the team and mentor's frame is .208 (95% CI [.212, .214]), in the second meeting the distance is .166 (95% CI [.162, .169]), and in the third meeting is .128 (95% CI [.127, .129]). With each successive meeting, the team's game design frame was closer to the mentor's in the same meeting. This result suggests that the team increasingly mirrored the mentor's game design discourse in the meetings.

Result 2: Following the mentor where?

To examine the team and mentor's paths, we created epistemic trajectories of the mentor and team's frames across the three meetings. The trajectories show the convergence reported above, with the team and mentor frames becoming successively similar with each of the three meetings (Figure 1).

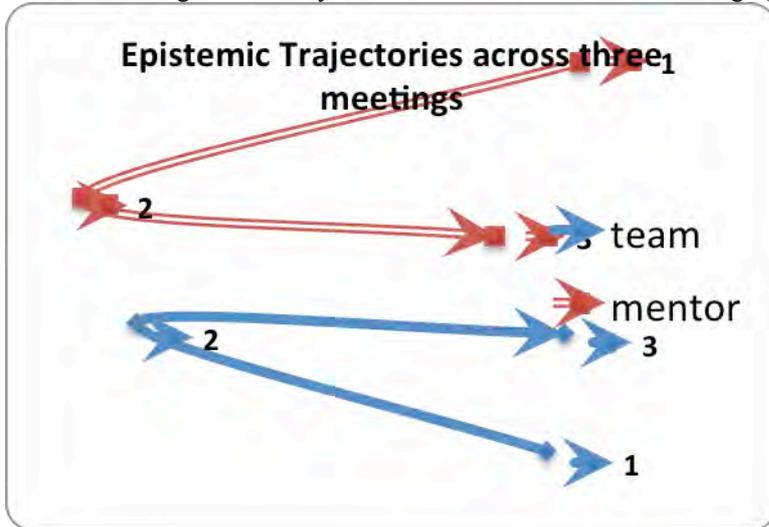


Figure 1: Team and Mentor Epistemic Frame Trajectories

Although the distance between the team and mentor's frame was reduced with each meeting, both the mentor and team's frames in the second meeting were much further from the frames in either the first or third meetings. In other words, the team's third meeting and first meeting frames were more similar than their frame in the second, and the same is true for the mentor. These trajectories suggest that the team's epistemic frame development did not proceed uniformly from the first to the final meeting.

Result 3: Why that path?

To see what might have caused the nonlinear trajectories, we looked at the change in the relative centralities of three key frame elements across the three meetings (Figure 2). This figure shows the relative centrality of the skill of *concept development*, the knowledge of *game mechanic*, and the knowledge of the *content domain* in the mentor and team's frame in each of the three meetings.

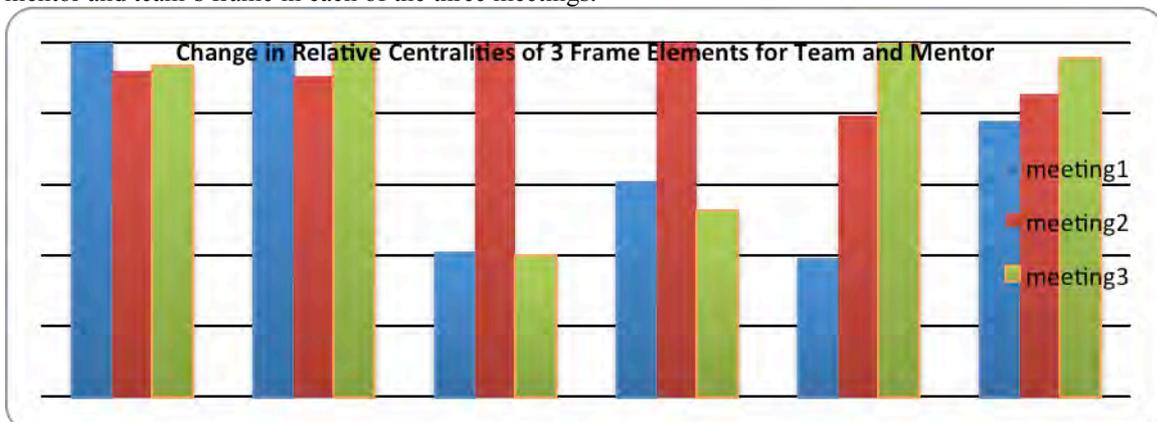


Figure 2: The change in importance of frame elements.

The development of a concept is important for both the team and mentor in all three meetings. In the second meeting, and only the second meeting, content research becomes the most central concern. Finally, we see that the game mechanic becomes increasingly important across the three meetings. The spike in the relative centrality of K/Content Domain would appear to explain the large distance between the frames in the second meeting and the other meetings.

Discussion

The results presented here describe how a team of novices in a practicum come to think like professionals under the guidance of a mentor. By the time the practicum was half over, which was the end point of this study, the team needed to have a game concept and game mechanic that were tightly bound together and that simulated a well-researched content domain. This study is about the path and mechanism by which the team arrived there.

First, the team imitated the discourse of the mentor. The team's epistemic frame grew more similar to the mentor's frame with each meeting. A feature of the "Follow Me!" coaching style is the mentor's demand that the learners imitate (Schön, 1987). It is possible that the increasing similarity of discourse in the later meetings is due to differences in the degree of freedom in the conversational topics: perhaps earlier in the practicum the conversations were simply more "open-ended." We argue, however, that the convergence of professional discourse is by design, not happenstance. The conversations the team had in their meetings were informed by the particular intentions and direction of their mentor as well as the general construction of the practicum. Regardless of whether any particular topic of conversation was due to one or the other, the conceptual development of the game designer frame requires this convergence. The "Follow me!" coaching model serves to guide learners who are doing work without necessarily knowing how to do it or why they are doing it.

The epistemic frame trajectories show the path down which the mentor led the team by offering an additional perspective on how the team and mentor's frame's converged. Mentors and students did not take a "direct route" from where they started to where they each ended. Instead, their second meeting, distant from the first and third meetings, indicated a developmental "detour." The change in the importance of three key frame elements—similar for the mentor and team across the three meetings—helps explain indirect path. The development of a concept was important in all three meetings. In the second meeting, the content domain became the most central concern. The game mechanic became increasingly important across the three meetings. Although the game ultimately must have a mutually reinforcing concept and mechanic, the mentor first led the team to other connections: to the connections between the content domain that the game is simulating and both the game concept and mechanic. That is, the development of a concept and game mechanic rely on the development of an understanding of the content domain. The students progressed from where they "mostly just talked" without "any gameplay yet" to where they used content domain language to link their game mechanics to their game concept. In other words, the mentor concentrated on one part of the frame in order to scaffold another part of it, which suggests that the detour in the epistemic trajectories was not a detour at all. Rather, the shortest distance from novice to professional thinking may not be to simply model best expert practice.

That the team has this indirect trajectory is perhaps unremarkable. After all, sometimes learners take steps backward on their trajectory forward. That both the team and the mentor share this type of trajectory, on the other hand, implies a learning experience quite different from the way traditional curricula, instruction, and assessment are organized. Most school subjects, for example, are organized to be taught in a strictly atomized and sequential manner. If learning to think like a professional requires rather than just accommodates indirect learning trajectories, as these results suggest, then the type of coaching by which mentors scaffold different connections within an epistemic frame is a type of learning relationship that deserves more attention.

Another valuable finding of this study is that ENA was shown to be a useful way to quantify the development of epistemic frames, as well as the relationship between the students' and mentor's frames. Other methods of discourse analysis may offer similar or additional results to those found in this study. However, the promise of ENA is that it is driven by frequencies of the co-occurrence of qualitative codes, and thus captures how practitioners connect the aspects of professional vision. In particular, projecting the distances between interactive units—whether they be meetings, activities in within meetings, turns of discourse within activities, or who the unit is associated with—by creating epistemic trajectories is a promising way to explore the nature of developing epistemic frames and complex ways of thinking in general.

The results presented here have several limitations. The ethnographic nature of this study means that any conclusions are limited to what one particular group did in the context of one particular practicum. Further, this study focuses on a limited amount of data. Team activity without the mentor may have added information about aspects of the epistemic frame the team internalized. Similarly, examining epistemic trajectories within meetings would allow us to map how frames developed in relation to the activity and discourse in the meetings. In addition, when looking at the relationships between the mentor and student frames, this study treated the students as a team, and so does not show individual development. The "Follow me!" coaching strategy might work differently for different students, differently during different activities or at different stages of the

practicum. Finally, ENA is a new method for understanding the development of an epistemic frame. As such, we expect it to develop in ways that allow us to better test significant events in frame development. Despite these limitations, the results here suggest that focusing on how mentors coach learners to develop epistemic frames should be useful for further studies of professional education or of apprenticeship-based programs for youth, and that epistemic network analysis is a useful tool for uncovering these learning processes.

References

- Abt, C. C. (1970). *Serious Games*. New York City, NY: The Viking Press, Inc.
- Bartholomew, D. J., Steele, F., Moustaki, I., & Galbraith, J. I. (2008). *Analysis of Multivariate Social Science Data*. Boca Raton: Chapman & Hall/CRC
- Collins, A., & Halverson, R. (2009). *Rethinking Education in the Age of Technology*. New York, NY: Teachers College Press.
- Gee, J. P. (2003). *What Video Games Have to Teach Us About Learning and Literacy*. New York: Palgrave/Macmillan.
- Gee, J. P., & Shaffer, D. W. (2010). Looking where the light is bad: Video games and the future of assessment. *Phi Delta Kappa International EDge*, 6(1).
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96(3), 606-633.
- Halpern, R. (2010). *The Means to Grow Up: Reinventing Apprenticeship as a Developmental Support in Adolescence*. New York: Routledge.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, Mass: MIT press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Nash, P., & Shaffer, D. W. (2010). Mentor modeling: the internalization of modeled professional thinking in an epistemic game. *Journal of Computer Assisted Learning*, 27(2), 173-189.
- Rose, M. (2004). *The Mind at Work: Valuing the Intelligence of the American Worker*. New York: Penguin Books.
- Ryan, G., Toohey, S., & Hughes, C. (1996). The Purpose, Value, and Structure of the Practicum in High Education: A Literature Review. *Higher Education*, 31(3), 355-377.
- Salen, K., & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. Cambridge MA: The MIT Press.
- Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*: Jossey-Bass San Francisco.
- Shaffer, D. W. (2006). *How computer games help children learn*. New York: Palgrave Macmillan.
- Shaffer, D. W. (2010). *The Bicycle Helmets of "Amsterdam": Computer games and the problem of transfer*. Madison: University of Wisconsin-Madison.
- Shaffer, D. W., Hatfield, D. L., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E. A. S., et al. (2009). Epistemic Network Analysis: A prototype for 21st Century assessment of learning. *The International Journal of Learning and Media*, 1(2), 33-53.
- Squire, K. D., & Jenkins, H. (2004). Harnessing the power of games in education. *Insight*, 3(1), 5-33.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research* (3rd ed.). Thousand Oaks: Sage Publications, Inc.
- Sullivan, W. M. (1995). *Work and integrity: The crisis and promise of professionalism in America*. New York: Harpercollins.

Acknowledgments

This work was funded in part by the MacArthur Foundation and by the National Science Foundation through grants REC-034700, DUE-0919347, DRL-0918409, DRL-0946372, and EEC-0938517. The opinions, findings and conclusions do not reflect the views of the funding agencies, cooperating institutions, or other individuals.

Gameplay as Assessment: Analyzing Event-Stream Player Data and Learning Using GBA (a Game-Based Assessment Model)

V. Elizabeth Owen, University of Wisconsin-Madison (*GLS Center*), vowen@wisc.edu
R. Benjamin Shapiro, Tufts University, ben@cs.tufts.edu
Richard Halverson, University of Wisconsin-Madison (*GLS Center*), halverson@education.wisc.edu

Abstract: This extended study (building on pilot research) presents a Game-Based Assessment model (GBA) designed to capture relevant information on play and test whether it can constitute reliable evidence of learning. A central challenge for videogames research in education is to demonstrate evidence of player learning. Assessment designers need to attend to the ways in which gameplay itself can provide a powerful new form of assessment. The GBA model has two key layers which build on content-based educational game design: a semantic template that determines which click-stream data events could be indicators of learning; and learning telemetry that captures data for analysis. This study highlights how the GBA was implemented in a stem-cell science learning game, and shows how the GBA demonstrates a relationship among success, kinds of failure, and learning in the game.

Objectives and Theoretical Framework

A central challenge for videogames in education is to demonstrate evidence of player learning. A typical approach to assess learning in games is to measure the quality of player learning in terms of independent, pre-post instruments. This process can compare game-based learning against other kinds of interventions, but, in treating the game itself as a black box, we lose the unique characteristics of the games as a learning tool. James Gee has suggested that games themselves provide excellent models for designing the next generation of learning assessments. Well-designed games reward players for mastering content and strategies, scaffold player activities toward greater complexity, engage players in social interaction toward shared goals, and provide feedback (through gameplay) that allows players to monitor their own progress (Gee, 2005). Rather than ignore the motivating and information-rich features of games in capturing learning, assessment designers need to attend to the ways in which gameplay itself can provide a powerful new form of assessment. This requires learning researchers to think of games as both intervention *and* assessment; and to develop methods for using the internal structures of games as paths to generate evidence of learning.

This study's framework is the Game-Based Assessment model (GBA), designed to capture data on player learning in the midst of gameplay. It's an extension of GBA pilot research, which introduced the model and preliminary findings around gameplay patterns and learning (Owen et al., 2012). The GBA framework has been developed by the Games, Learning and Society (GLS) research group as a process for capturing relevant information on play and testing whether it can constitute reliable evidence of learning. The GBA model draws on concepts and tools from evidence-centered design (e.g., Mislevy & Haertel, 2006), stealth assessment (Shute, 2011) and educational data mining (e.g. Baker & Yacef, 2009) to describe a strategy for building assessment tools into game design from the ground up in order to use game play itself as the barometer of player learning.

GBA Model and Methods

The Game-Based Assessment model is grounded in the content model and game-flow design of the game development process, and emphasizes two key layers: the *semantic template* and *learning telemetry*. Below, we describe each feature of the model in context of *Progenitor X*, a GLS game about regenerative medicine.

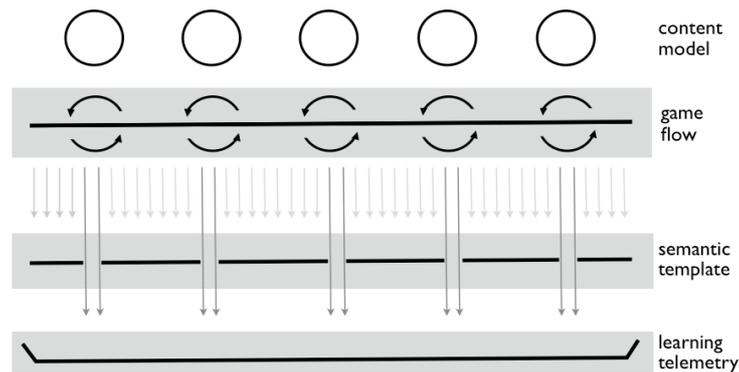


Figure 1: Game-Based Assessment model

The GBA model is designed to draw significant gameplay moves from the game-context. The model is integrated into an overall four-layer GLS game design strategy: the *content-model*; the *game flow* design; the *semantic template*; and the *learning telemetry* (Figure 1). The first two layers, the content model and the game flow design, constitute the game design process. The content model outlines the learning goals for the game. The game flow design builds player interaction opportunities around these learning goals to create a gaming experience. The final two layers, the semantic template and the learning telemetry, form the assessment process. The semantic template selects relevant data from the click-stream generated by gameplay; the learning telemetry layer collects and organizes the resulting data-record into player-profiles. Here we provide a brief overview of how these layers, using the game *Progenitor X* as an example, comprise a generic blueprint for our approach to assessment-driven game design.

Content Model

The content model for a GLS game consists of several content chunks that string together a series of core concepts along a process that represents current thinking in a domain. Because the resulting medium for interaction is a game, rather than a simulation, the design team is concerned with creating motivating conditions of play as well as the representational accuracy of the content model. *Progenitor X* provides an example.

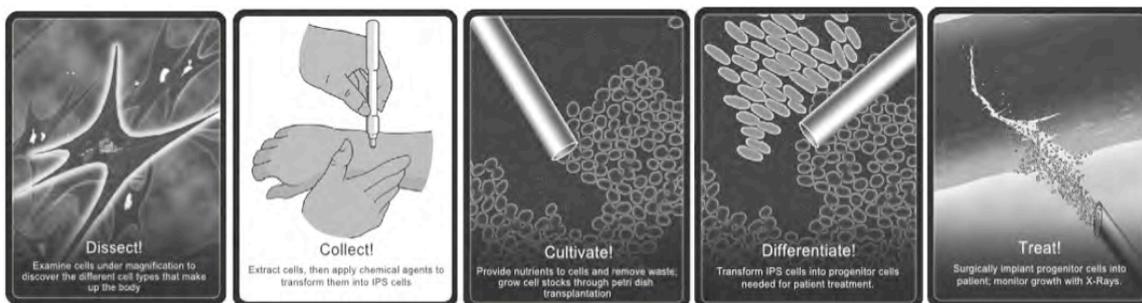


Figure 2: *Progenitor X* content model

Progenitor X invites players to dissect, collect, cultivate, differentiate and treat diseased tissues via adult epithelial stem cells (Figure 2). Each verb in the content model provides an occasion for interaction. A process derived from professional practice provides a simplistic but coherent account of real scientific procedures, designed for accessibility to the study demographic of secondary school students.

Game-flow Design

The game is designed to motivate player interaction and learning. Through the iterative design process, the content model is embedded in a world that allows players to interact with the core ideas. The verbs of the content model are translated into key moments in interactive gameplay. *Progenitor X* embodies this process, taking the verbs of the content model and creating a turn-based puzzle game in which players assume the role of a regenerative biologist to prevent a zombie apocalypse. Given a series of content-based objectives, *Progenitor* players perform three main actions in game-flow: cultivate (or *start* a cycle of) cells, *treat* them, and then *collect* the resulting target material.

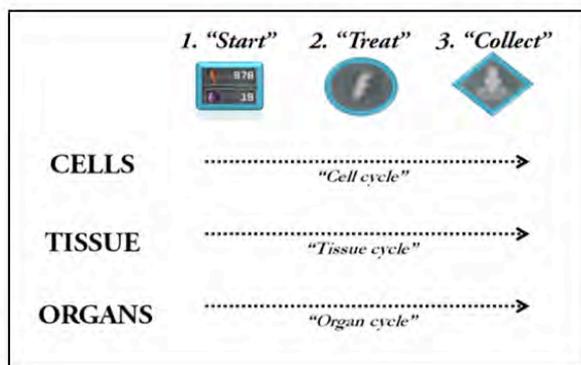


Figure 3: *Progenitor X* game-flow design

Semantic Template

The semantic template defines conceptual windows of interest in the game that represent key moments of learning. It is designed around the intersection of the content model with the game-flow design. The key question for semantic template design is: of all the clicks that players make in the game, which ones indicate learning? The semantic template represents a hypothesis about which in-game actions can generate interesting evidence of learning.

In *Progenitor X*, the semantic template revolves around the *start*, *treat*, and *collect* verbs of the content model. The first sequence of player action is the **cell cycle**, in which players *start*, *treat*, and *collect* a group of vital cells. These new cells are used to create tissue in the next cycle (i.e. **tissue cycle**), where players use the same action sequence. Then comes an **organ cycle**, where the player uses the newly collected tissue to *start*, *treat*, and *collect* their way to a whole, healthy organ. Two views of the semantic template in *Progenitor* are shown below.



	A	B	C	D	F	G
In-Game Sequence	Mission	Cycle**	Cell population type (Stage 1)	Treatment tool type (Stage 2)	Collection cell type (Stage 3)	
1						
2	1	1 A. ips	fibroblasts	electroporate	ips	
3	2	1 B. meso i	ips	growth factor	meso	
4	3	1 A. ips ii / C. ecto i	fibroblasts	electroporate	ips	
5	4	1 C. ecto ii	ips	growth factor	ecto	
6	5	2 E. tissue i	meso	N/A	meso / tissue	
7	6	2 E. tissue ii	meso	N/A	meso / tissue	
8	7	2 A. ips iii / D. endo i	fibroblasts	electroporate	ips	
9	8	2 D. endo ii	ips	growth factor	endo	
10	9	2 E. tissue iii	endo	N/A	endo / tissue	
11	10	3 F. organ i	N/A	scan	necrotic tissue	
12	11	3 E. tissue iv	meso	N/A	meso / tissue	
13	12	3 F. organ iii	N/A	scan	necrotic tissue	
14	13	3 A. ips iv / B. meso ii	fibroblasts	electroporate	ips	
15	14	3 B. meso iii	ips	growth factor	meso	
16	15	2 F. tissue v	meso	N/A	meso / tissue	

Figure 4: A) Basic and B) Detailed semantic templates of *Progenitor X*

Learning Telemetry

The learning telemetry layer collects the data specified by the semantic template and organizes it for analysis. It is a mechanism of the game environment that coordinates the components of the game world into a sequential data-stream that enables analysts to track player paths across the game-world.

In *Progenitor*, capturing telemetry started with identifying gameplay moments within the semantic template on an event-stream level. Significant click-stream events (over 400) around the action sequence (*start*, *treat*, and *collect*) were documented and flagged for recording. Then, search parameters were constructed, allowing reconstruction of interface cues as context for player actions. Lastly, a query schema was developed to pull the specified event-stream data from the massive database. Ultimately, through synchronizing GBA’s semantic template and learning telemetry, we were able to identify and collect three kinds of telemetric action-sequence data: cycle-specific, cumulative, and individual.

_id	4fbd2a48cf86304117000004	Objectid
_id	4fbd2a48cf86304117000004	Objectid
created_at	2012-05-23 13:19:52 -0500	Date
gameName	ProgenitorX	String
key	ToolSelectedData	String
schema	4-18-2012	String
session_token	2012-05-23_115607	String
timestamp	854	Int
toolName	Move	String
updated_at	2012-05-23 13:19:52 -0500	Date
user_id	235	Int
wasFlashing	YES	Bool
_id	4fbd2a48cf86304117000005	Objectid

Figure 5: Learning telemetry - raw click-stream data output (*information from one click*)

	A	B	C	D	E
1	cycle type	IP51			
2	destroys	Shock	Move		
3	0	1	12		
4	cycle type	Meso			
5	destroys	Collect	GrowthFar	Move	
6	0	1	1	16	
7	cycle type	IP51			
8	destroys	Collect	Shock	Move	
9	0	1	2	12	
10	cycle type	Ecto			
11	destroys	Collect	GrowthFar	Move	
12	0	1	3	16	
13	cycle type	Tissue1			
14	destroys	Collect	Move		
15	0	1	3		
16	cycle type	Tissue2			

Individual

	A	I	L	M
1	ID	Populate	Total Collect	Grid Destro
2	c_10@stu.de.nt	58	50	6
3	c_11@stu.de.nt	31	20	8
4	c_2@stu.de.nt	16	9	2
5	c_3@stu.de.nt	33	28	3
6	c_4@stu.de.nt	20	5	13
7	c_5@stu.de.nt	29	19	6
8	c_6@stu.de.nt	15	12	1
9	c_7@stu.de.nt	28	12	14
10	c_8@stu.de.nt	31	21	6
11	c_9@stu.de.nt	44	35	7
12	d_1@stu.de.nt	17	10	5
13	d_2@stu.de.nt	26	18	6
14	d_3@stu.de.nt	18	13	2
15	d_4@stu.de.nt	29	17	8
16	d_5@stu.de.nt	42	27	5
17	e_1@stu.de.nt	20	10	5

Cumulative

	A	F	G	K	P	Q	R	T	U	V
	email	cell started	cell destroys	cell success	tissue started	tissue destroy	tissue success	organ started	organ destroy	organ success
1										
2	d_5@stu.de.nt	8	0	9	8	0	8	9	0	2
3	f_10@stu.de.nt	6	0	4	8	2	6	5	3	3
4	f_9@stu.de.nt	5	0	6	8	2	5	4	0	1
5	c_5@stu.de.nt	6	0	6	10	1	5	4	0	1
6	f_13@stu.de.nt	4	0	4	5	0	5	5	0	1

Cycle-Specific

Figure 6: Learning telemetry - processed click-stream data output

Data Sources and Evidence

Data analysis required synthesizing learning telemetry data output with additional assessment. Specifically, we added two additional data sources to the core telemetric corpus: an adapted measure of success in gameplay, and data from an isomorphic pre- and post-test.

In order to sort the player data into meaningful patterns, we developed an *efficiency ratio* that measured the number of successful cycle completions by a player over the number of times the cycle was tried. (A “success” means the player has collected the right material at the end of the cycle.) For example, if a player successfully collected the required number of cells in a cycle 2 times, and tried to complete the cycle five times, the player’s efficiency ratio would be 40%. (The higher the percentage, the more efficient the play.)

$$\text{Efficiency Ratio} = \# \text{ of successes} / \# \text{ of tries}$$

We also aggregated results from the pre- and post- content assessment, which included a series of questions about the stem-cell content model based on consultation with regenerative biologists Dr. James Thomson, Dr. Rupa Shevde, and Dr. Gary Lyons. Here, we specifically looked at change in player performance on content questions as measured before and after gameplay.

Results

Along with the telemetry data, players’ efficiency ratio and the change in performance on the pre-post content questions became key data features. In this study, these features were analyzed within the aggregate data set of n=110 with nonparametric statistical methods, given the non-normal distribution of pre-post percentage change

and event-stream variables. The tests were directional (one-tailed) and conducted at a baseline alpha of .05, guided by our main hypothesis that on-task gameplay would result in increased learning outcomes (as measured by pre-post performance). Specifically, we used a paired-sample Wilcoxon rank test (Table 1) and Spearman’s correlation test throughout the analysis. Multiple pairwise contrasts were conducted using the Holm procedure for assigned alpha, centering around constructs of play efficiency (see “Efficiency Ratio” above) and temporal game progression. The data was collected in the summer of 2012 from 110 randomly-selected middle-school students who played *Progenitor X* as part of a summer school curriculum (either in their Dane County school classroom, or on-site at the Wisconsin Institute for Discovery).

Aggregate results revealed intriguing reasons to look further into the “black box” of the game. First, with a 19.5% average increase in pre-post content scores, the game seemed to be a noteworthy learning vehicle. A first look at player progress through the game revealed a significant positive relationship between successive completion of game objectives and learning ($r = +.272, p = .002$). Success as well as game progression mattered; the number of successful cycles in gameplay was positively correlated with learning outcomes ($r = +.216, p = .012$). Thus, it seemed that player performance during specific points throughout the game held a key to deeper understanding of the relationship between gameplay and learning outcomes. This led us to investigate what was going on with players on a micro level within each given cycle.

Table 1: Aggregate results summary

	Pre-Post Gains
Total Gameplay	19.5% average increase
Objectives Added	Significant positive correlation ($r = +.279, p = .002$)
Objectives Completed	Significant positive correlation ($r = +.272, p = .002$)
# of Successful Cycles	Significant positive correlation ($r = +.216, p = .012$)

In order to examine player interaction, we mapped all possible cycle outcomes. Within a cycle, players populate (*start*) an initial grid with the right kinds of cells (green check, Figure 7), and then transform those cells (*treat*) into a target cell/tissue to *collect*. After initial population with the right cell, the cycle can end in three ways: collecting the right cell (success), collecting the wrong cell (failure), or over-manipulating/treating the cells so that the Ph becomes toxic (failure).

Additionally, a player could have also initially populated the grid with the wrong cell (red X, Figure 7). In this case, there are two options for ending the cycle: collecting the wrong cell, or over-manipulating the cells until the Ph levels (health) becomes toxic.

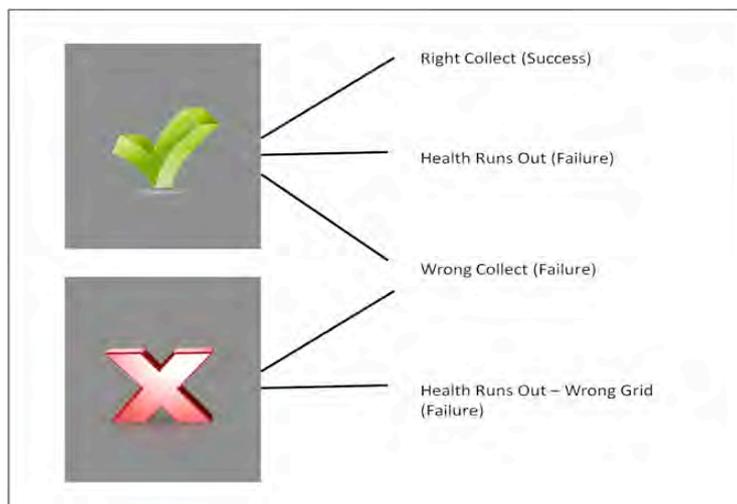


Figure 7: Progenitor gamespace - incorrect AND correct initial grid population

The possible outcomes imply varying degrees of player compliance with multiple in-game cues (e.g. flashing buttons & in-game narration). To explore this idea, we clustered the types of failures into “near” and “far failure” (Figure 8). We grouped three possible player outcomes: correct collection (successful); correct set-up but health runs out (“near failure”); incorrect setup and/or incorrect collection (“far failure”).

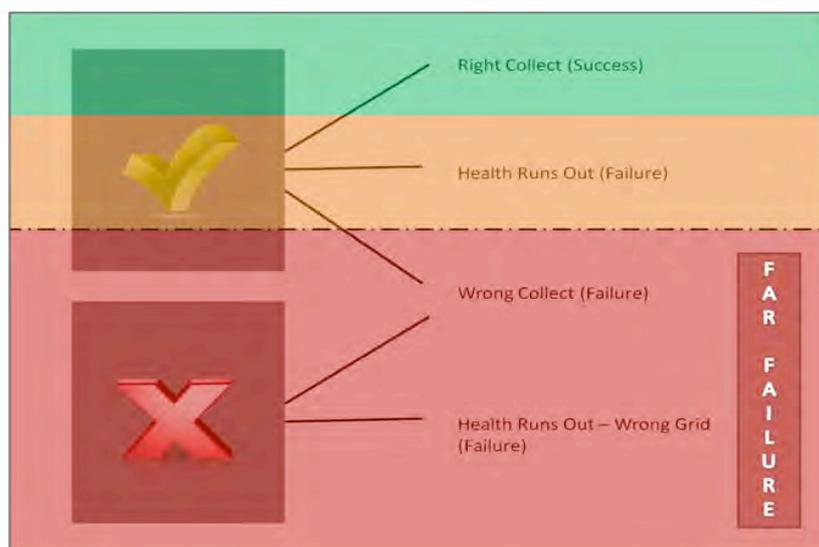


Figure 8: “Far failure” in *Progenitor* gamespace

The analysis of far failure gave considerable insight, especially after parsing game data by level sequence (from the beginning Objective 1 to the final Objective 8). Far Failure and success in the first and last objectives in the game proved vital, with implications for early scaffolding and “boss-level” assessment. Objective 1, the game level immediately after the tutorial, seemed to be a critical filter point for those trying to “game the system” by not attending to instructional cues (expressed in far failure numbers). Within Objective 1, far failure had a significant negative correlation with learning outcomes (Table 2). In fact, for players with extreme numbers of Objective 1 far failure (over two standard deviations from the mean), the average increase on the pre-post biology assessment was only 3.6% (as opposed to a 19.5% in the aggregate group). Conversely, success in Objective 8 (the “boss” level) had significant positive correlations with pre-post gains, both in terms of raw number of successes ($r = +.272$) and efficiency ratio ($r = +.193$). (The final stage, this Objective 8 “boss” level, required a cumulative performance of all lab skills learned in the game.) Thus, far failure and success in these bookend levels seemed to hold critical significance for learning outcomes.

Table 2: Objective-level aggregate results summary

	Pre-Post Gains
Objective 1 Far Failure	Significant negative correlation ($r = -.167, p = .04$)
Upper Extreme: Players w/ Obj 1 Far Failure	3.9% average increase (<i>15.6% lower than aggregate</i>)
Objective 8 Successes	Significant positive correlation ($r = +.272, p = .002$)
Objective 8 Efficiency Ratio	Significant positive correlation ($r = +.193, p = .022$)

To explore these phases of learning connection more deeply, we divided players into quartiles according to pre-post change. The upper quartile (33 students) had the largest gains in stem-cell content question scores, while the lower quartile (41 students) had the smallest. Interestingly, the patterns of play in each quartile supported the trends from the aggregate analysis: overall game progression and success was positively connected with pre-post gains, as was performance on the boss level, while far failure in critical early game cycles was negatively associated with learning outcomes.

Overall quartile trends revealed key differences in play progression, success, and far failure. During the same duration of gameplay, Objective progression was significantly different for the two groups ($p = .019$). The upper quartile, on average, got to Objective 7 (out of eight total), while the lower quartile made it to Objective 6 (Table 3). Like total playtime, the number of total successful cycles between these quartiles was not significantly different. However, upper quartile successes had positive correlation with learning, while the lower quartile’s had none. Proportionally, the lower quartile also had twice as many off-task failures – per objective, low performers had two far failures, while the upper quartile averaged one. Within similar duration and success counts, the upper quartile got further in the game, had fewer far failures, and had contextual success that supported learning gains. Thus, each group seemed to be using their time very differently, prompting further level-specific investigation.

Table 3: Overall quartile results summary

	Upper Quartile	Lower Quartile
Timestamp	No significant difference.	
# of Objectives Added	7	6
# of Objectives Completed	6	5
# of Far Failures per Objective	1	2
Total # of Successful Cycles	No significant difference.	
Successful Cycles vs. Pre-Post Gains	Significant positive correlation ($r = +.377, p = .015$)	NO correlation

Objective-level play data exposed telling differences between the quartiles. The first trend was that early far failure was associated with learning losses. Compared to the upper quartile, the lower quartile had twice as many far failures (on average) in Objective 1 of the game. These Objective 1 far failures had a significant negative correlation with learning for the lower quartile ($r = -.277$). Since each quartile had similar numbers of total failures, these far failure proportions are a stark contrast. Essentially, the lower quartile had more frequent early far failures, which then associated with poor learning outcomes. Conversely, in the second trend, players with the greatest learning gains performed very well in the final objective. The upper quartile had significantly higher Objective 8 successes than the lower quartile ($p = .023$). Cell cycle success, a skill specifically taught in early levels, was also significantly higher in the final level for the upper quartile (twice as many as the lower quartile; $p = .023$). These endgame contrasts imply that the top learners' on-task performance in early levels provided the gameplay mastery necessary to excel at the boss level, and demonstrate knowledge of the baseline biology lab practices that underlie core game mechanics.

Table 4: Objective-level quartile results summary

	Upper Quartile	Lower Quartile
# of Obj 1 Far Failures	2	4
Obj 1 Far Failure vs. Pre-Post Gains	NO correlation	Significant negative correlation ($r = -.277, p = .040$)
# of Total Failures	No significant difference.	
# of Obj 8 Successful Cycles	3	2
# of Obj 8 Successful Cell Cycles	2	1

Overall, four major trends emerged in the data. In the aggregate set, general gameplay progression as well as total gameplay success had a positive relationship to learning. Far failure in tutorial levels of the game was negatively correlated with learning outcomes, and boss-level performance was positively associated with pre-post gains. Quartile analysis reflected these trends, reinforcing that far failure and success in the bookends of the game were crucial differentiation points for learning.

Important implications arise from these data. First, gameplay progression and success seem to effectively support content learning. Secondly, early-game correlations suggest that certain types of failure in context - not failure itself - inform learning. Far failure in the crucially-scaffolded Objective 1 could be a critical indicator of players losing learning due to "gaming the system" (signaled by lack of attention to instructional input). Thirdly, boss-level performance seems to be an effective gauge of overall content knowledge gains. These bear major impact for generalized game design on two counts. In early levels, far-failure-based adaptive feedback may be key in changing off-task players' learning trajectories. Finally, the boss level, designed as an effective indicator of learning content via game mastery, may render pre-post content tests needless – ultimately making *gameplay itself* the only necessary assessment.

Conclusion

The GBA model allowed us to move beyond a simple pre-post comparison of game play to learning outcomes by providing data on how players interacted with the game environment. The design of the semantic template allowed us to collect data at key moments in gameplay; the learning telemetry allowed us to tag and assemble these click-stream data points into play profiles we could use for analysis. The resulting data allowed insight into the role of success and failure in *Progenitor X* game play. As we have seen, games allow players to experiment with failure without real-world consequences. However, the kinds of failures players experience matter. Productive failure (Kapur, 2008) suggests that effective learning environments encourage students to activate prior knowledge as a condition for direct instruction. *Progenitor X* introduces players into an unfamiliar subject matter context (regenerative medicine), but in a familiar game-genre context (puzzle-based videogames). Familiarity with the game-conventions invites players to interact with a system in order to learn programmed

relationships between cells, tissues, tools and cultures. One interpretation of our analysis is that productive failure and success happen when players bridge game-mechanic knowledge to content-model knowledge through gameplay; non-productive failure happens when players ignore the content model and treat *Progenitor X* solely as a colorful puzzle game with zombies. Specific junctures of connection between content learning and in-game performance, shown in this study, have major implications for educational game design. In early gameplay, differentiating non-productive “far” failure in vital tutorial levels may be key in guiding off-task players towards better learning outcomes. In final stages, these data highlight the potential of a well-designed boss level to be a comprehensive, naturalistic, summative assessment of content knowledge. In future game design as well as research, the richness of the data generated by the GBA will allow us to further explore the relations between player interaction and learning.

References

- Baker, R. & Yacef, K. (2009). The state of educational data mining: A review and future visions. *Journal of Education Data Mining*, 1(1), 3-17.
- Gee, J.P. (2005). Learning by design: good video games as learning machines, *E-learning*, 2(1), pp. 5-16.
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379-424.
- Mislevy, R. J., & Haertel, G. D. (2006). Implications of Evidence-Centered Design for Educational Testing. Principled Assessment Designs for Inquiry Technical Report 17. SRI International Center for Technology in Learning. Accessed May 31, 2012 at http://padi.sri.com/downloads/TR17_EMIP.pdf
- Owen, V.E., Wills, N., & Halverson, R. (2012). *CyberSTEM: A Game-Based Evidence Model*. Presented at Games+Learning+Society, Madison, WI, June 13-15.
- Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–523). Charlotte, NC: Information Age.

Acknowledgments

This work was made possible by a grant from the National Science Foundation (DRL-1119383), although the views expressed herein are those of the authors’ and do not necessarily represent the funding agency. We would also like to thank the *GLS Center* team, including Kurt Squire, Mike Beall, Ted Lauterbach, Allison Salmon, Kevin Alford, Meagan Rothschild, Shannon Harris, Keri Bell-Gawne, Greg Vaughan, and Nate Wills.

The Role of Identities in the Process of Knowledge Construction in CSCL settings

Murat Oztok, OISE / University of Toronto, 252 Bloor Street West, Toronto, Canada,
murat.oztok@utoronto.ca

Abstract: The learning scientists have argued that knowledge construction is a process of collective thinking within a learning community. Thus, knowledge construction is simultaneously an individual and social process that requires group cognition and situated meanings. However, while the CSCL researchers have investigated the situated knowledge in the process of collective thinking, little work has been done to fully understand how different identity categories play a role in sense-making and knowledge construction. This research, therefore, explored in detail how individuals utilize their different identity categories to make situated meanings when they collaborate with each other in the process of knowledge construction in online learning environments. Results demonstrated that individuals do not experience online learning through only one aspect of their identity but rather that learning experiences evoke different elements of their identities that are used continuously and simultaneously when they collaborate with each other at the every phase of knowledge construction.

Introduction

The learning sciences literature has long argued that learning is simultaneously an individual, social, and cultural process that involves collaboration and active participation in learning communities (Brown, Collins, & Duguid, 1989). Exploring active participation in online learning communities, CSCL literature suggests that there is a close relationship between collaboration and identity development (Ke, Chávez, Causarano, & Causarano, 2011). Socio-cultural learning theories explains this link by arguing that learning is about practices and activities in cultural worlds (Holland, Lachicotte Jr, Skinner, & Cain, 2001), where identities are central for participation (Gutiérrez & Rogoff, 2003). Analyzing relationships among practice, identity, and learning, scholars have conceptualized learning as an aspect of practice-based identity and defined identity as a result of learning through practice (Lave & Wenger, 1991). Such an understanding is particularly important since it “reconceptualizes learning from an in-the-head phenomenon to a matter of engagement, participation, and membership in a community” (Nasir & Cooks, 2009 p. 42). Thus, building upon socio-cultural learning theories, learning scientists have argued that learning is tied to the context (Cole, 1996) and that identity is one concept through which individuals make sense of their context (Wenger, 1998).

However, while CSCL research has deemed identity as an important concept in understanding how students engage with each other (Oztok, 2012), little work has been done to fully understand how different identity categories (e.g., race, gender, class, profession, or ethnicity) play a role in sense-making and knowledge construction. Previous research has understudied how students make sense of the subject-matter in relation to how they perceive themselves and their peers as they involve in the process of knowledge construction. Specifically, understanding the role of identities in the process of knowledge construction is still a major challenge for the learning sciences (Nasir & Hand, 2008) and CSCL research (Ke et al., 2011). This research, therefore, explores how individuals utilize their different identity categories to make situated meanings when they collaborate with each other in the process of knowledge construction in online learning environments.

Background and Rationale

Since knowledge construction is not a mere exchange of information but requires coherence and convergence among participants (Suthers, 2006), individuals need to make situated meanings in collaborative learning practices (Stahl, 2010). Considering the importance of sense-making for collaboration (Scardamalia & Bereiter, 1994) and knowledge construction (Stahl & Hesse, 2009), I argue that by exploring how identities are manifested in threads and how students utilize their identities to cultivate, share, discuss, and negotiate meanings, CSCL research can understand the processes by which individuals make situated meanings and construct knowledge in online learning environments. Indeed, employing the concept of identity for exploring how individuals make sense of the context invites a discussion about the meaning of the concept and its appropriateness as a theoretical framework to explore knowledge construction.

The concept of identity has always been at the center for many political, philosophical, economic, or academic debates. Academically, it has been deemed vital by many disciplines; yet, identity means different things to different scholars from different disciplines. Indeed, notions of identity are as diverse as the bodies of literature that have taken up the concept. Fields as diverse as psychology, sociology, humanities, and philosophy offer discipline-specific conceptualizations and definitions of identity. While other definitions exist, the field of

education is mostly influenced by psychological and sociological conceptualizations and much of the debate around identity in educational research derives from the tensions between these two perspectives (Buckingham, 2008). Psychological perspectives are built upon the idea that identity is a single state that one achieves over time and development (Erikson, 1968). According to this perspective, individuals have a coherent and authentic self that is internally consistent and inexorable. Currently, socio-culturally informed learning scientists have begun to move from this normative perspective and have suggested that identity is a complex and continuously shifting phenomenon and that it is context-based and linked to the learning practices (Esmonde, 2009).

Socio-cultural learning theories conceptualize identities as enactments within figured worlds (Holland et al., 2001), where individuals' practices are constrained or enabled through sets of norms (Nasir & Cooks, 2009): "[w]ithin these figured worlds, identity is constructed as individuals both act with agency in authoring themselves and are acted upon by social others as they are positioned ..." (p. 41). In this sense, identities are subject positions readily available for individuals and these identities are performed or enacted as individuals engage with each other. That is, identity is something people perform or practice in collaborative learning situations, as opposed to something people have. However, identity enactments are not neutral or straightforward; rather, they are guided through social, cultural, political, or historical symbols (Jenkins, 2008): "[identity do] not, and cannot, make people do anything; it is, rather, people who make and do identity, for their own reasons and purposes" (p.9). Thus, identity is conceptualized as simultaneously an individual and a social practice.

Employing the concept of identity as a theoretical lens to analyze threaded discussions can provide means for understanding how individuals perceive themselves in relation to others when they engage in collaborative learning practices. Particularly, depending on the context in which individuals collaborate, they choose to saliently use different identities (Wenger, 1998), through which they analyze their previous experiences (Holland et al., 2001) while they make sense of the present subject-matter (Nasir & Cooks, 2009). Individuals' identities, in this sense, reflect sets of meanings derived from negotiations, agreements, or disagreements that occur in the process of collaboration (Gutiérrez & Rogoff, 2003). Indeed, the CSCL research has already shown that students engage in shared knowledge-building discussions when they share their experiences with one other and build on each others' thoughts to interpret the learning materials (Arvaja, 2012). Therefore, by analyzing how individuals enact their identities in threaded discussions, one can understand the cognitive processes by which collaboration enables knowledge construction, meaning-making (Stahl, 2010) and shared understanding (Suthers, 2006). Thus, CSCL research should explore how identities play a role in cultivation, distribution, and construction of knowledge.

Current Research

This study explores how individuals manifest and utilize different aspects of their identities (i.e. gender, profession, or ethnicity) in the process of knowledge construction through multiple case studies (Creswell, 2006). In order to purposefully select cases, we (the author and his colleagues) analyzed the participants' biography pages (or profile pages, which allow students to create their identities and their online existence by introducing themselves with their own words along with their picture or avatar) and identified two individuals who can maximize the exploration of the phenomenon. For deciding the cases, we paid considerable attention to choosing individuals who utilize a number of identities. For each of the two case studies, we created an online persona by analyzing their profile pages and examined how these online personas are enacted when individuals engage with each other in a collaborative knowledge construction process.

Data is collected from a fully online graduate education course (N=13) offered at a large North-American research university that took place in Winter 2012. The course comprised twelve modules, each corresponding to one week, in which students discussed weekly readings. Each week, one or two students acted as moderators. They facilitated discussion throughout the week, kept discussions on track, and finally offered a summary of the week's issues, providing opportunities for sustained discourse, increased interaction, and rich discussions.

Analyzing Knowledge Construction and Identity Manifestations

In order to understand how individuals make sense of themselves, of their peers, and of the subject matter through their identities, we (the author and his colleagues) examined the online discussions through "interaction analysis model" (inter-rater consistency is .82), identified threads with knowledge construction, and analyzed identity manifestations in those threads in relation to knowledge construction. For this research, threads are analyzed semantically since semantic analysis "more accurately represents each groups' development of ideas over time" (Wise & Chiu, 2011, p. 458).

Interaction analysis model (Gunawardena, Lowe, & Anderson, 1997) is employed for examining the process of knowledge construction. It is based on the socio-cultural learning theories, theoretically and empirically grounded, specifically developed for analyzing asynchronous threaded discussions, and have already been employed by CSCL researchers (e.g., Ke et al., 2011; Wise & Chiu, 2011). The model

conceptualizes knowledge construction as a process of negotiation in which meanings, perspectives, and perceptions play roles. While not strictly sequential, interaction analysis model suggests five phases for knowledge construction to occur: 1) sharing and comparing of information, 2) discovery and exploration of dissonance or inconsistency among participants, 3) negotiation of meaning of knowledge co-construction, 4) testing and modification, and 5) phrasing of agreement and applications of newly constructed meaning. According to this model, interactions begin by sharing and elaborating ideas (phase 1), leading individuals to identify potential conflicts among each other (phase 2). Individuals build on these conflicts by negotiating meanings and perspectives (phase 3); then, they revise their ideas and perceptions (phase 4), allowing individuals apply their new knowledge (phase 5).

Since individuals do not exist as physical beings but enact their identities through language-in-use in online learning environments, discourse analysis is employed to explore how language-in-use mediates between identities, meanings, and practices. Discourse analysis reveals how identities regulate particular forms of meanings and social experiences by deconstructing the relationships among saying, doing, and being in the language-in-use (Gee, 2011):

If I say anything to you, you cannot really understand it fully if you do not know what I am trying to do and who I am trying to be by saying it. To understand anything fully, you need to know who is saying it and what the person saying it is trying to do. (p. 2)

According to this perspective, language-in-use not only gets its meaning from the context in which it is used but also it creates, sustains, or transforms meanings, negotiations, and practices in the context. Thus, it is an essential tool to critically analyze the otherwise hidden intersections between identity enactments and participants' pedagogical practices in technologically-mediated environments.

In order to understand how identities are manifested in language-in-use, I employ seven interrelated building tasks to analyze the discourse (Gee, 2011): (1) Significance, (2) Practices, (3) Identities, (4) Relationships, (5) Politics, (6) Connections, and (7) Sign systems and Knowledge. Specifically, we semantically analyzed each note in a thread and interpreted the meanings and identities in that particular note through building tasks in order to explain how individuals choose to enact particular identities and how such identities affect their engagements with each other. Indeed, while all building tasks may not be readily available in all text and while some building tasks may be more salient than others, each building task can provide means for understanding how individuals move through and within the online learning environment. Therefore we used the building tasks in relation to each other and triangulated between them.

Furthermore, it is important to note that I do not conceptualize identity manifestations in a Cartesian sense (that identities are either present or absent in a particular note) but rather argue that identities exist in various forms and are almost always so well blended into practices and meanings. That is, language-in-use not only conveys academic knowledge but also communicates bits of identity manifestations; either explicitly (i.e. As an artist, I think ...) or implicitly (i.e. I don't agree with you because I had problems with my Grade 7 class, ...)

Results

We identified 16 threads in which knowledge construction happened while students collaborated with each other. Of those 16 threads, we identified 9 threads in which our both case studies engaged with each other. Here I present 3 instances (due to space limitations) of how the two cases utilize their identities in those knowledge-construction threads.

Meet Michelle and Xiaomei. Michelle is a part-time PhD student. As a White-Canadian, she lives in Beijing with her daughter, where she works as an English lecturer at a university. She defines herself as an activist and hopes to employ critical pedagogy in her dissertation. Xiaomei is from China. She is a full-time PhD student and teaching assistant at her university. She got her master's degree from an English university, where she taught English as a Second Language courses. Her research interest is teaching English with digital media.

In week 2, class discussed the pedagogical potentials of web 2.0 and digital media. While Michelle has many identities to choose from, she enacted her "political-activist" identity and drew attention to political issues as she engaged with subject-matter. Michelle deconstructed the social and political aspects of using digital media in schools. This is, indeed, what Gunawardena et al. (1997) identify as "phase 1: sharing and comparing of information" in their interaction analysis model. Then, Michelle challenged the perspectives offered in weekly readings by articulating her concerns about the tyranny that social media creates (phase 3) and, in later notes, invited her peers to consider the motives behind the knowledge produced in social media (phase 4). Her peers built on these ideas by noting the importance of social, political, and historical structures and power relations in the reproduction of knowledge (phase 5). In response to her peers, Michelle further analyzed how mainstream newspapers influence both public opinion and public policies regarding schooling (phase 5). In the same threaded discussion, however, Xiaomei enacted her "teacher" identity and drew from her experiences as a teacher to make sense of the weekly readings (phase 1). Specifically, Xiaomei embedded her disagreement with

her peers (that cell phone use should be restricted for students) in her teaching experience (phase 3). She explained how she lets her students use a dictionary application on their cell phones for writing courses, suggesting such technologies are useful in class (phase 4). Xiaomei further posited that although a few students surfed on the internet instead of studying, her experience also suggests that mobile phones have a pedagogical value (phase 3). She concluded that, “as teachers, we need to be not only innovative but also motivate our students to use technology for teaching and learning” (phase 5). In this threaded discussion, while Michelle enacted her identity as an activist, Xiaomei enacted her teacher identity. However, either an activist or teacher, they both interpreted the subject-matter according to their own perspectives or backgrounds and in return their peers perceived them as how they enacted their identities in that particular thread. As this example demonstrates, identities can provide situated-meanings by which individuals can make sense of their learning practices.

The topic in week 4 was teaching and learning in Computer-Mediated-Communication (CMC) and social-networking sites. In a thread about communication types, Michelle drew from her teaching and living experience in China for more than a decade and suggested that asynchronous communication suits her students best (phase 1). In another note, Michelle disagreed with a claim in the readings (that since Chinese students are shy, they prefer not to participate in discussions; thus, they underachieve compared to their Western counterparts) (phase 2). She further argued that while her Chinese students are shy, online courses suit them because they are comfortable expressing their thoughts and opinions (phase 3). She concluded that, contrary to the readings, her Chinese students are successful in online courses (phase 3). In a later note, one of her peers asked Michelle whether she would think differently if her students were asked to use wikis (since wiki-based web applications allow individuals to edit each others work) instead of asynchronous threaded discussions (phase 4). Responding to this note, Michelle reappraised her thoughts (phase 3) and considered the different pedagogical values that different asynchronous communication types have (phase 5). In the same threaded discussions, Xiaomei directly engaged with Michelle enacting her Chinese and teacher identity while drawing from her experience in England and Canada. She articulated her perspective about communication types (phase 1) and particularly provided her insights about editing someone else’s work in a wiki-based application (phase 2). Xiaomei indicated that the idea of changing another’s work without permission is intimidating since she believes that “it is like saying that you think that what you have to say is more important or more valid than what someone else has to say” (phase 1). She suggested that when she was teaching in China no one edited anyone’s work to maintain the group harmony; thus, she saw no pedagogical value in collaborative wiki-based web applications (phase 2). Therefore, based on her experience in China, Xiaomei disagreed with her peers as well as with the weekly readings (phase 3). However, in a later note, she elaborated on her teaching experience in England and suggested that wiki-based applications has certain pedagogical value since her students in England were able to work productively and comfortably in wiki-based assignments (phase 4). As it is exemplified in this thread, Michelle and Xiaomei used their different identity traits in a single thread: they both enacted their teacher identity in interpreting the weekly readings; however, while Michelle further developed her knowledge as a Western person living in China, Xiaomei further enhanced her understanding building on her experience as a Chinese person living in a Western country.

In week 8, the class discussed the social and cultural issues of teaching and learning in CMC settings. In a thread about authenticity of learning context, Michelle enacted her maternal identity. Even though such an identity was not salient in Michelle’s other notes, being a mother was one of the identity traits to shape her experience in this particular collaborative process. Challenging one of her peers’ perspective, Michelle explained how she actively volunteers for her kids’ school (and according to her, it is not common practice in China) and argued that a authentic learning context requires collaboration between parents and teachers (phase 3). Another class-mate (also enacting maternal identity) built on Michelle’s perspective by further elaborating her experience with her kid’s school and suggested that the weekly readings offer an idealized understanding of what authenticity is and unfortunately do not reflect real-life situations (phase 4). Summarizing the weekly discussion and affirming her classmates (phase 5), Michelle synthesized weekly readings and noted that “most of us have the best of intentions as teachers and parents, but as you put it so well, life... happens!”. In the same thread, Xiaomei enacted her artist identity to make sense of the weekly readings. She agreed with her peers and further underscored the importance of social and cultural issues in creating an authentic learning context by providing examples from her learning art experience (phase 5). Specifically, she analyzed the role of culturally-relevant materials in learning local arts and explained how such materials helped her as an artist (phase 4). According to Xiaomei, as an artist, “authenticity is about individuals themselves rather than the material itself or [its] geographical situation”. Furthermore, to some extent, Xiaomei enacted her student identity (though not a unique identity trait, it was saliently used in this particular thread). Summarizing her learning experience in China, Xiaomei described she felt disconnected when her arts teacher used Western paintings as course material and disengaged because the material was not meaningful enough (phase 4). She also reflected on her current learning experience in Canada and suggested that it did not differ from China in terms of authenticity (phase 1). As this instance illustrates, while Michelle and Xiaomei had similar perspectives on the importance of

authenticity for learning, they enacted two different identity traits: Michelle is a mother and Xiaomei is an artist. Furthermore, this instance shows that individuals not only utilize their basic identity categories (i.e. ethnicity or profession) but also draw from their broader out-of-classroom identities when they collaborate with each other in the process of knowledge construction.

Discussion

Three instances provided in this paper represent three different and unique ways that identities play a role in collaboration and knowledge construction. In the first example, Michelle and Xiaomei utilized their identities to make situated meanings in their interpretations of the weekly readings. Furthermore, when Michelle and Xiaomei enacted certain identities, their peers accepted them as such and engaged with them accordingly. In the second example, both Michelle and Xiaomei enacted multiple identities at once; that is, while Michelle was a teacher and a Westerner living in China, Xiaomei was a teacher and a Chinese person living in Western world. In the third example, Michelle and Xiaomei had similar perspectives and agreed with each other's conceptualizations, they enacted different identities; thus, they explained their perspectives from different point of views.

Taken together, these three instances can provide initial understanding for the role that identities play in collaborative learning activities. For example, this research shows that individuals bring their various identities into the collaborative learning processes and utilize their different identities under different circumstances for different reasons. That is, individuals do not experience online learning through only one aspect of their identity but rather that learning experiences evoke different elements of their identities (Buckingham, 2008) that are used continuously and simultaneously (Gee, 2000) as they collaborate with each other. However, despite the variety of identity traits being used, in all these situations, identities play a canonical role: they provide situated meanings for individuals to draw from their experiences in order to make sense of their learning experiences. Depending on the context in which they participate, they choose to saliently use different identities, through which they analyze their previous experiences while they make sense of the present subject matter. Thus, they have different learning experiences and outcomes since learning is an aspect of practice-based identity (Nasir & Cooks, 2009). Identities, in this sense, become socio-cultural, historical, and cognitive artifacts by which individuals legitimize their learning experiences (Lave & Wenger, 1991) within their communities of practice (Wenger, 1998). This finding is align with the learning sciences literature in suggesting that:

identities allows a way to understand the intrapersonal dimensions of learning and to capture the ways that learning settings can support or fail to support not just the acquisition of skills and knowledge but a deep sense of connection with participants. This connection is more than just membership or belonging. In this way, participation in learning settings extends beyond learning (though learning is certainly critical) to the very definition of who one is and who one is in the process of becoming through participation. (Nasir & Hand, 2008, p. 176)

While this research affirms the existing literature, it further explains how identities create, support, and sustain the "interpersonal dimension of learning" by providing examples of how individuals make sense of each other and of their learning in relation to their identities.

This research also provides initial understanding for the role of identities in the process of knowledge construction. The current CSCL research (e.g., Gunawardena et al., 1997; Wise & Chiu, 2011) suggests that knowledge construction begins with basic interactions that facilitate the sharing of individuals' experiences. Once the foundation for common ground is established, meaningful dialogue and collective reflection takes place. Through the process of negotiations, individuals provide detailed analysis or criticism, drawing from their experiences to construct new knowledge. When new knowledge is constructed, individuals enhance their insights by developing an understanding that enables them to reconsider their understandings. "This reflects the cohesive conception of collaborative learning according to which learning through discussions can be conceptualized as developing, challenging, and re-conceptualizing ideas" (Arvaja, 2012, p. 99). While this research affirms the current literature, it further explains that identities are manifested at every phase of knowledge construction; however, they play a unique role in each different phase. For example, while identities can provide basic information about an individual in phase 1, they can they provide further detailed information about individuals and their perceptions in phase 2. In phase 3, individuals rely on their identities to challenge current perspectives offered by their peers or by weekly readings. While individuals analyze the learning material or the subject-matter in relation to their identities in phase 4, they explain what they learned from that particular discussion in relation to their experiences phase 5. In other words, individuals simply use their identities to articulate what their prior thoughts are in the early stages of knowledge construction. Then, they use their experiences to further develop or challenge the existing perspectives in middle stages of knowledge construction. Finally, in later stages, they find a common ground and reconsider their thoughts and further explain what they learned in relation to their identities.

Conclusion

This research is built upon the idea that learning is simultaneously an individual and social process that involves collaborative practices within learning communities (Brown et al., 1989). One approach to understand how social practices mediate cognitive activity (Cole, 1996) is to understand the role of identities in the process of knowledge construction. Indeed, identities provide opportunities for individuals to make situated meanings (Gutiérrez & Rogoff, 2003; Lave & Wenger, 1991) and incorporate aspects of themselves into the learning practice (Nasir & Cooks, 2009; Wenger, 1998). Analyzing in detail and demonstrating the role identities play in collaborative learning activities, this research suggests that identities provide more dialogic and reflective interaction; therefore, identities are not tangential to the collaborative learning practice but rather a central part of it.

The learning scientists have long argued that collaborative learning is a process of collective thinking (Stahl, 2006) and that it is manifested in and by dialogue (Sawyer, 2006). The CSCL researchers, therefore, have focused on the situated knowledge in the process of collective thinking and contended that the development of well-articulated identities in online learning situations can actually build a greater sociology of learning (Ke et al., 2011). Indeed, using individuals' own experiences as a source of learning "can support student agency... by giving opportunities to make personal sense through personal lives" to learning activities (Arvaja, 2012, p. 86). While this research affirms the current perspectives, it further explains in detail how identities play a role in each and every stage of collaborative knowledge construction processes.

Along with its epistemological contributions, this research also provides conceptual understandings for the CSCL research. Studying knowledge construction as a dialog distributed among individuals and exploring the collective discourse through individuals' identities go beyond focusing on cognitive artifacts and capture the socio-cultural and historical nature of situated meaning-making. By conceptualizing knowledge construction as a socio-cultural and historical process can provide opportunities for understanding knowledge construction not as limited in temporal time-scales (e.g., Arvaja, 2012; Mercer, 2008) but as a long-term identity based collaborative process.

The findings should be considered in relation to the limitations of the study. First, while this study provides important guidance toward understanding the role of identities in the process of knowledge construction, it must be noted that this study is just one preliminary investigation. Results are based on one course and two case studies. Thus, more research is needed in order to make stronger claims about such relationships between identity and knowledge construction. Second, identities in this study are not analyzed in relation to the concept of power. Indeed, including the concept of power and broader societal structures in analyzing identity can provide better insights about the constraints of collaborative work. Furthermore, in this study, identities are stripped from their social, political, and historical meanings since the aim of the research was not to provide hidden curriculum of collaborative learning practices but rather was to analyze the ways in which individuals utilize their identities for their learning. Thus, a research with more critical agenda should consider such meanings in its analysis.

Overall, this research suggests that identities are in the center of collaborative knowledge construction. As the theme of CSCL 2013 argues, this research supports the current understanding but further explores the phenomenon at different levels. However, more research is needed to confirm or challenge the findings of this research in order to provide better and stronger understating of the role of identities in the process of knowledge construction.

References

- Arvaja, M. (2012). Personal and shared experiences as resources for meaning making in a philosophy of science course. *International Journal of Computer-Supported Collaborative Learning*, 7(1), 85–108. doi:10.1007/s11412-011-9137-5
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, 18(1), 32–42.
- Buckingham, D. (2008). Introducing Identity. In D. Buckingham (Ed.), *Youth, Identity, and Digital Media* (pp. 1–22). Cambridge, MA: MIT Press Journals.
- Cole, M. (1996). *Cultural psychology: a once and future discipline*. Cambridge, MA: Harvard University Press.
- Creswell, J. W. (2006). *Qualitative Inquiry and Research Design: Choosing among Five Approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Erikson, E. H. (1968). *Identity: Youth and Crisis*. New York, NY: Norton.
- Esmonde, I. (2009). Mathematics Learning in Groups: Analyzing Equity in Two Cooperative Activity Structures. *Journal of the Learning Sciences*, 18(2), 247–284. doi:10.1080/10508400902797958
- Gee, J. P. (2000). Chapter 3: Identity as an Analytic Lens for Research in Education. *Review of Research in Education*, 25(1), 99–125. doi:10.3102/0091732X025001099
- Gee, J. P. (2011). *An Introduction to Discourse Analysis: Theory and Method* (3rd ed.). New York, NY: Routledge.

- 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.
- Gutiérrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher*, 32(5), 19–25.
- Holland, D., Jr Lachicotte, W., Skinner, D., & Cain, C. (2001). *Identity and Agency in Cultural Worlds*. Cambridge, MA: Harvard University Press.
- Jenkins, R. (2008). *Social Identity* (3rd ed.). New York, NY: Taylor & Francis.
- Ke, F., Chávez, A. F., Causarano, P.-N. L., & Causarano, A. (2011). Identity presence and knowledge building: Joint emergence in online learning environments? *International Journal of Computer-Supported Collaborative Learning*, 6(3), 349–370. doi:10.1007/s11412-011-9114-z
- Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. New York, NY: Cambridge University Press.
- Mercer, N. (2008). The Seeds of Time: Why Classroom Dialogue Needs a Temporal Analysis. *Journal of the Learning Sciences*, 17(1), 33–59. doi:10.1080/10508400701793182
- Nasir, N. S., & Cooks, J. (2009). Becoming a Hurdler: How Learning Settings Afford Identities. *Anthropology & Education Quarterly*, 40(1), 41–61. doi:10.1111/j.1548-1492.2009.01027.x
- Nasir, N. S., & Hand, V. (2008). From the Court to the Classroom: Opportunities for Engagement, Learning, and Identity in Basketball and Classroom Mathematics. *Journal of the Learning Sciences*, 17(2), 143–179. doi:10.1080/10508400801986108
- Oztok, M. (2012). Tacit knowledge in online learning: community, identity, and social capital. *Technology, Pedagogy and Education*, 1–16. doi:10.1080/1475939X.2012.720414
- Sawyer, R. K. (Ed.). (2006). *The Cambridge handbook of the learning sciences*. New York: Cambridge University Press.
- Scardamalia, M., & Bereiter, C. (1994). Computer Support for Knowledge-Building Communities. *Journal of the Learning Sciences*, 3(3), 265–283. doi:10.1207/s15327809jls0303_3
- Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge, MA: The MIT Press.
- Stahl, G. (2010). Guiding group cognition in CSCL. *International Journal of Computer-Supported Collaborative Learning*, 5(3), 255–258. doi:10.1007/s11412-010-9091-7
- Stahl, G., & Hesse, F. (2009). Paradigms of shared knowledge. *International Journal of Computer-Supported Collaborative Learning*, 4(4), 365–369. doi:10.1007/s11412-009-9075-7
- Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, 1(3), 315–337. doi:10.1007/s11412-006-9660-y
- Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, UK: Cambridge University Press.
- Wise, A., & Chiu, M. (2011). Analyzing temporal patterns of knowledge construction in a role-based online discussion. *International Journal of Computer-Supported Collaborative Learning*, 6(3), 445–470. doi:10.1007/s11412-011-9120-1

Effects of an Interculturally Enriched CSCL Script on Students' Attitudes and Performance

Vitaliy Popov, Harm J.A. Biemans, Martin Mulder

Chair Group of Education and Competence Studies, Wageningen University, the Netherlands
(Email: vitaliy.popov@wur.nl, harm.biemans@wur.nl, martin.mulder@wur.nl)

Abstract: In an attempt to foster collaboration in general as well as to bridge intercultural differences in culturally heterogeneous groups engaged in CSCL, this study introduces an interculturally enriched collaboration script (IECS). A randomized, two-group pretest-posttest research design was used to compare the effects of the IECS with a general collaboration script (CS) on students' collaborative learning in culturally heterogeneous groups in a CSCL environment. In this study, Master students from a university in the Netherlands (74 subjects, representing 22 countries) worked in dyads on an environmental problem about biodiversity collapse in tropical forest protected areas. The results showed that the IECS positively affected students' attitude towards online collaboration and satisfaction with learning processes and outcomes. Student groups in both conditions achieved comparable task performance and exhibited a comparable level of willingness to collaborate online.

Cultural issues in computer-supported collaborative learning

Individuals from diverse and distinct cultural backgrounds are brought together by the need to collaborate for professional, personal and academic reasons. This comes from the idea that more can be achieved through cultural synergy. Using collaborative technologies in intercultural educational environments creates both potential benefits – by promoting collaborative learning, and sharing culturally divergent knowledge – and challenges, in terms of equitably supporting learners, specifically with different cultural backgrounds. Many culture-related factors have a considerable impact on the group interaction process (Lim & Liu, 2006) and should be taken into consideration when designing and implementing collaborative learning environments (Cox, Lobel, & McLeod, 1991; Weinberger, Clark, Hakkinen, Tamura, and Fischer, 2007).

The present research investigates collaborative learning from a social-constructivist learning perspective. Social constructivist scholars consider the collaborative learning environment as a place where learners may work together and support each other as they use a variety of tools and information resources in their guided pursuit of learning goals and problem-solving activities (Zhu, 2009). A particular emphasis in social constructivism is put on the importance of the background and culture of learner, since the learner's cultural context of cognition influences the way he or she attains knowledge in the learning process.

There is a number of learning aspects that can be studied in this context. Among these are:

Attitude towards online collaboration (students' perspectives on online collaborations which may influence their learning behaviors). While accomplishing a task collaboratively students from various cultures differ in terms of their perspectives on group work, procedural knowledge of how to collaborate and learn together (Weinberger et al., 2007; Cox, et al., 1991). These differences may lead to conflicts because of the mismatch of students' perspectives and approaches to online collaboration (Zhao & McDougall, 2008). Previous research suggests that students' attitude towards the learning environment is one of the determining factors for assessing the quality of educational interventions, as well as predicting both learning processes and outcomes (Zhu, 2009);

Willingness to collaborate online (a psychological mechanism that drives individual's behavior in terms of both acceptance of new learning information and adaptation and use of new communication technologies in the future). Given the increased emphasis on the importance of participation in online collaborative learning environments, it is essential that students are willing to collaborate by these means. It has been suggested that willingness can be seen as a dependent variable (Davis, 1989). In case of intercultural collaboration, culturally diverse students have to overcome the extra layer of complexity owing to culture-related differences;

Satisfaction with online collaborative learning processes and outcomes (students' opinions or feelings as experienced in computer-mediated collaboration). A large and growing body of literature has investigated satisfaction related to the learning process and the other with learning outcomes. Satisfaction can be used as used as one of the criteria to assess the quality of online training. A number of previous research studies reported positive relationships between learning satisfaction and instructional approaches, social presence, task characteristics, group dynamic (e.g. LaPointe & Gunawardena, 2004); and

Online collaborative task performance (the extent to which the learner achieved certain learning goals which can be reflected in the form of the final group products). The task performance is the extent to which the

learner achieved certain learning goals, which can be reflected in the form of the final group products. Early studies in the field mainly focused on the quality of collaborative learning products or individual learning results, but often overlooked the fact that the outcome is mediated by the quality of group learning processes (Lim & Liu, 2006). Meanwhile, many social and cultural factors impact significantly on the interaction process and are yet to be taken into account when studying CSCL (Weinberger et al., 2007; Cox et al., 1991). However, there is a dearth of research on various learning outcomes between students from different cultures in the CSCL context. In this regard, this study aimed to gain insights into relationship between the dynamics of the group interactions and learning performances in culturally diverse groups in CSCL.

In an attempt to foster collaboration in general as well as to bridge intercultural differences in culturally heterogeneous groups engaged in CSCL, this study introduces a scaffolding technique, in this case the IECS. The basic premise of this script is that collaborating students from different cultural backgrounds may particularly benefit from following a shared external collaboration script that scaffolds their interaction, coordinates their learning activity and incorporates culture-related differences. For this purpose, the researchers developed three external collaboration scripts and implemented them in an authentic learning environment: one conventional CS without intercultural elements, and two collaboration scripts with intercultural enriched ingredients tailored specifically for two groups of culturally distinct backgrounds based on previous research (Rummel & Spada, 2005; Nisbett, 2003; Hofstede, 1991; House, Hanges, Javidan, Dorfman, & Gupta, 2004). The objective was to compare the effects of the IECS with the CS, both scripts were embedded in the CSCL environment, on students' collaborative learning in culturally heterogeneous groups in the CSCL environment. This was done by using a two-group pre-test-post-test research design, and applying quantitative and qualitative measurements.

Phenomenon of cultural and behavioral collaboration scripts

Over the last twenty years, a range of techniques and approaches have been developed that are conducive to successful and productive online group work (Dillenbourg, & Jermann, 2007). However, only few studies provide guidance to facilitate online collaborative interaction among culturally diverse students. These studies focus mostly on intercultural foreign language education through CMC, internet-mediated approaches that are used to raise learners' intercultural awareness, and the use of various e-tools to support brainstorming in order to benefit from cultural diversity in knowledge and perspective (Wang & Fussell, 2010). But, knowledge is still lacking in what instructional support may help culturally diverse collaborative learners obtain the maximum benefit from shared experiences in CSCL.

Specifically in response to this need, this experimental study introduces a scaffolding technique, in this case an external IECS intended to both promote collaboration in general and bridge the cultural gap between collaborative learners with different cultural backgrounds. There is a well-documented body of research focused on the application of collaboration scripts in computer-mediated systems and their effects, which have proved to be particularly influential in supporting collaboration (e.g. Koschmann, 1999; Rummel & Spada, 2005). The basic premise of the collaboration script in this study is that collaborating students may particularly benefit from following a shared external collaboration script that scaffolds their interaction and coordinates their learning activity. Our assumption was that culturally diverse collaborating students with the help of an external IECS can overcome differences between them, minimize the amount of effort required to coordinate their activities and generate collaborations that would lead to more positive experiences and higher performance in such groups.

The design of the general CS in this study is based on an approach developed by Rummel and Spada (2005). By integrating empirical findings from several research studies on communication and computer-mediated collaboration, Rummel and Spada (2005) introduced three levels merging in a good collaboration. These include macro (I) and micro (II) levels, plus domain-specific requirements (III) depending on the learning task. Each of these three levels appoint to certain collaborative behavior in online student interactions (i.e. social, cognitive and coordinating behaviors) and all of them are necessary for successful online group collaboration.

At the core of the CS is a sequence on how to pursue a goal of collaboration process, with the precise prescriptions of learning activities in how students may engage in collaborative discourse. The collaborative work consisted of three phases: initial phase, main phase, and final phase. All three levels necessary for successful online group collaboration (macro, micro, and domain-specific) were incorporated throughout all three phases. In this study, the collaboration process and script instructions were tailored to the collaborative learning activities required to analyze the problem of biodiversity collapse with the application of the Driver-Pressure-State-Impact-Response (DPSIR) model (i.e. DPSIR is a framework that helps to identify and describe processes and interactions in human-environmental systems).

The initial phase covered the technical and social introduction. The main phase included discussion of the background literature, analysis of the problem by constructing the DPSI-part of the DPSIR Diagram, and identifying possible responses to avert biodiversity collapse in tropical forest protected areas. The final phase focused on prioritization of the responses.

The design of the IECS was based on the CS including exactly the same collaboration steps and instructions to them, plus additional cultural enrichment elements developed specifically for each of the two culturally distinct groups of students. These enrichments were derived from cross-cultural psychology findings (Nisbett, 2003; Hofstede, 1991; House et al., 2004). They provide a foundation to both develop an adequate understanding of these cultural differences and to design socio-technical support for collaborative learning that could involve, overcome and bridge the differences between cultures. This support was based on “the assumption that the creation of shared practices on the micro level would allow to bridge cultural gaps on higher level social aggregates” (Hinds, Zhao, Wulf, Thomas, Fussell, & Zhang, 2010, p.609). Thus, we build on by playing on learners’ social/cognitive diversity and knowledge interdependency so as to foster the different mechanisms needed for productive collaboration and maximize the benefits of culturally divergent knowledge.

For instance, different patterns might emerge at the early stage of collaboration in terms of how students get introduced to one another, and in the ways they start to accomplish the learning task. For example, Westerners tend to focus their communication on the task rather than on maintaining relationships at the early stage of working together (Hofstede, 1991; Nisbett, 2003). Some group members from Western countries may unintentionally offend other members from non-Western countries because they are so focused on their tasks that they omit some socialization protocols. A potential solution in this phase (i.e. get to know each other and start building group dynamics) could be to encourage collaborative partners to discover individual and cultural similarities and differences within the group. This can be realized by completing and exchanging students’ personal profiles, which may enable them to introduce themselves to other group members and discuss personal concerns and interests with other members. In this context, we asked students to complete personal profiles, which contained questions, both about content-related experience related to the task and about personal background. The only difference was that students from non-Western countries were instructed to construct a concrete idea of the content-related experience of themselves and their group members, whereas students from Western countries were requested to get to know their group members in terms of their personal backgrounds and try to build a relationship of trust in order to solve the task together. It was predicted that this script element might help to fulfill specific needs of two culturally distinct groups of students at the initial stage of collaboration.

Research questions

This study is aimed at answering the following central research question:

What are the effects of the IECS compared to the CS on students’ collaborative learning in culturally heterogeneous groups in a CSCL environment in higher education?

In order to answer the central research question, this research has been divided into four sub-questions.

To what extent does the IECS compared to the general CS affect students’:

RQs1....attitude towards online collaboration?

RQs2....willingness to collaborate online?

RQs3....satisfaction with online collaborative learning processes and outcomes?

RQs4....online collaborative task performance?

Method

Participants

Participants in this study were first year Master program students enrolled in educational program in the field of life and environmental sciences in university in the Netherlands. Our sample of 74 students comprised 18 Dutch and 56 international students; 52.7% of whom are women. Of the international students, 18 come from Europe (outside the Netherlands), 6 from Africa, 25 from Asia, 6 from South America and 1 from North America. The total number of countries represented in our study was twenty two. The age group of the respondents ranged from 19 to 37 years, with a mean age of 24.04 ($SD=3.17$), and 95.9% of respondents were below the age of thirty. 98.3 % of international students had been staying in the Netherlands on average two to three weeks by the time when this study was conducted. Almost all students, regardless of their cultural background, had some short-term previous travel experiences, internships, traveling for work purpose outside of their home countries for both academic and non-academic purposes. Well over half (64.9%) of the participants stated that they had much prior experience working in student group work and only 32.2% had prior experience working in multicultural student group work. 74.3% of the respondents had hardly any or not at all experience working online with students from the other country. All study participants must have proven English language proficiency by passing an oral and written exam while enrolled at the university where the given research was conducted.

Research setting

This study was conducted as a part of the course named Principles of Environmental Sciences (PES). This course is particularly designed for Master program students in Environmental Sciences. The PES course offers students the opportunity of updating and extending their knowledge of the basic concepts of environmental sciences. In a case study, which was used in this research study, students analyzed an environmental problem about biodiversity collapse in tropical forest protected areas. More specifically, Laurance, Carolina Useche et al. (2012) published a scientific paper on the 26th of July in Nature. After publication the article received a lot of media attention globally. The issue addressed in the paper forms the basis of the case study. Students were required to successfully complete this online collaboration assignment in order to pass the PES course. The quality of the students' group work was evaluated by a system of quantitative criteria developed by experts in the area of environmental sciences.

Procedure

All students were asked to collaboratively (in pairs) analyze an environmental problem related to the biodiversity collapse in tropical forest protected areas. All groups used the same online VCRI learning environment (after reading the same pre-study instructions), the same software and the same course manual. The same questionnaires were administered to all study participants. The only difference was that one half of the participants received only CS and the other half received the same collaboration guidance, but with extended intercultural instructions, the so-called the IECS. The participants were randomly assigned to the two conditions (the IECS and the CS), although it was ensured that each pair included one student of non-Western cultural background and one of Western cultural background. The IECS was tailored specifically for students with distinct cultural backgrounds. Collaborating students with the IECS were expected to follow a three-level instruction for all subtasks: (1) a general instruction on WHAT you needed to do; (2) an instruction on HOW you needed to do a certain subtask; (3) an explanation of WHY the subtask instruction is important. Collaborating students with the CS were provided only with the general instruction on WHAT to do. This design allowed us to empirically investigate the effects of the IECS on students' collaborative learning in culturally heterogeneous groups in the CSCL environment.

In total, there were one introductory and two online group work sessions during three consecutive days. Overall time required for completion of the assignment, including filling in pre-posttest questionnaires plus task introduction and debriefing, was about 10 hours. During online group work sessions students were seated at individual computers. The students interacted with the study personnel and with each other in English. Prior to the actual study, the participants were asked to fill in a number of questionnaires focusing on: (a) the demographic information and (b) prior experience with working collaboratively, specifically in intercultural setting, (c) technical (computer) skills, and (d) pre-test of attitude towards online collaborative learning.

During the initial phase, the students were introduced to the VCRI groupware program, to the assignment and to the procedures on collaboration. The participants were specifically asked to follow the guidance instructions. Before continuing with the main phase of collaboration, students were requested individually answer questions about the driving forces, pressures, states, and impacts related to the biodiversity collapse. The main phase of the collaboration consists of two subtasks: (1) analysis of the problem by constructing the DPSIR-part of the DPSIR Diagram, and (2) identification of possible responses to the problem of biodiversity loss. During both subtasks, a three-step pattern (a.b.c) is followed. Step (a) consists of individual work, which allows collaborative partners to bring in their own disciplinary knowledge and ideas. Following this, (b) the individual ideas should be discussed, ensuring the exchange of unshared information. After the discussion, (c) the individual proposals have to be integrated into a group joint solution, and reported usually in the COWRITER window. During the final phase, students were expected to conduct the prioritization of the different responses that they identified. This means that collaborative partners needed to identify which responses they would address or implement first, second, third (etc.), if they would be in the position to make such a decision. Also, students were instructed to provide supporting argumentation for the prioritization. At the conclusion of the final collaboration phase, each participant completed a number of post collaboration questionnaires (students' attitude, willingness and satisfaction towards online collaboration).

Learning platform

A Virtual Collaborative Research Institute (VCRI) groupware program was used in this study as a CSCL learning environment (Figure 1.). The VCRI incorporates a number of features designed to facilitate various collaborative activities online. More specifically, a Chat-tool allows a student to communicate with his/her collaborative partner(s) by exchanging instant messages. A Sources-tool includes all necessary information related to group task itself (e.g. assignment description, a literature source) and instructions to perform the task. Collaboration scripts used in this study was embedded in the Source-tool. All information available in this tool

can be opened and read from the screen. The VCRI program has a Cowriter-tool as a shared word processor, where students can simultaneously work together on their texts. Group members use a Diagrammer-tool to collaboratively make representations of their ideas by constructing various sorts of diagrams or flowcharts. To write down some ideas in a personal space, students can use a Notes-tool.

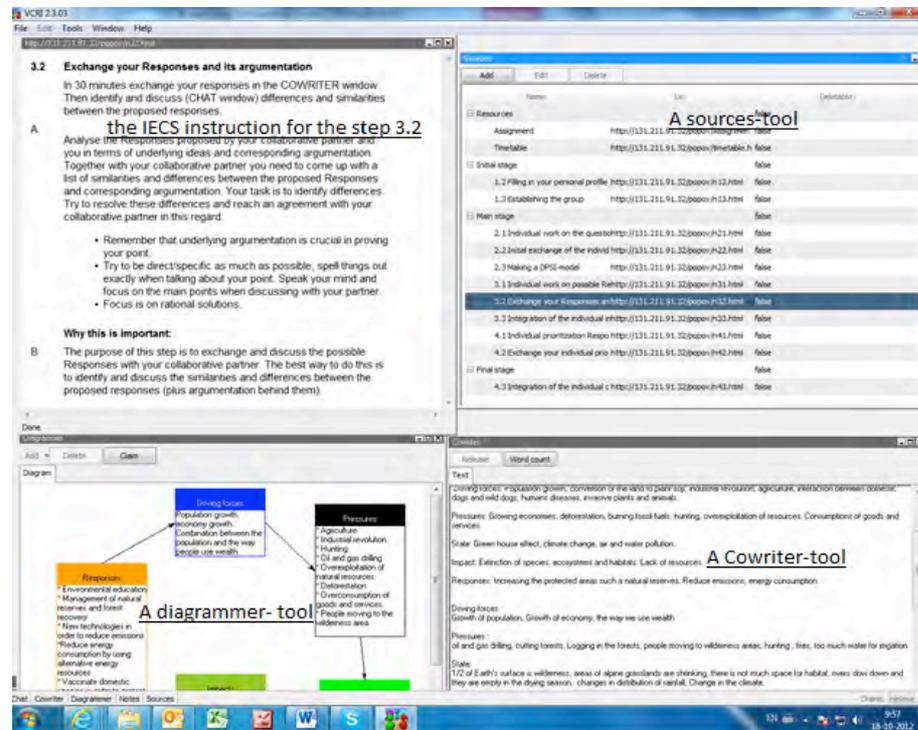


Figure 1. Screenshot of the VCR platform.

Instruments

Table 1 below presents a summary of the various empirical study constructs, their respective measuring instruments and data sources.

Table 1. Variables, instruments, and data sources.

Study construct	Instruments	Data source
Prior to collaboration		
(a) Demographic information, (b) prior experience working collaboratively, specifically in intercultural setting.	Self-made prior to collaboration survey.	Questionnaire
(c) Technical (computer) skills.	Self-made five multiple-choice questions	Questionnaire
(d) Pre-test of attitude toward online collaboration.	Survey developed by Thompson and Ku (2006).	Questionnaire
Post collaboration		
Post-test of students' attitude towards online collaboration.	Post-collaboration survey by Thompson and Ku (2006)	Questionnaire
Students' willingness to collaborate online.	Modified version of the survey developed by Chen, et al., 2006.	Questionnaire
Students' satisfaction with online collaborative learning processes and outcomes	Modified version of the 'perceived usefulness of the online course' developed by Giesbers et al., 2009.	Questionnaire
Students' task performance	Self-made system of quantitative criteria to assess the online group work (i.e. outcomes scores based on the competing the DPSIR assignment).	Log files of the constructed DPSI model and joint solutions (data obtained from Diagrammer and Cowriter tools of VCRI)
Control of the script use by the participants	Click count and survey	Log files; Questionnaire

Data analyses

To show the impact of the IECS on the students attitudes towards online collaboration, a repeated measures ANOVA was conducted with condition (IECS and CS) as the between-subjects variable and students' attitude scores (pre- and post study) as the within subjects (repeated measures) variable.

Separate univariate ANOVAs were used to determine whether there are any differences between experimental condition (collaboration with the IECS) and control condition (collaboration with only CS) for each variable (i.e. willingness to collaborate online, satisfaction with online collaborative learning processes and outcomes, and task performance).

Results

Pre-test control measures in two conditions

None of the participants had prior knowledge or experience working with the DPSIR framework. No significant differences were observed between the students in the IECS and the CS conditions with respect to age, $F(1, 72) = .30, p = .58$, gender $F(1, 72) = .20, p = .65$, the mean scores of technical (computer) skills, $F(1, 72) = .16, p = .33$, and prior group work experiences, $F(1, 72) = 1.25, p = .26$. These results indicated that there were no significant differences between participants in the two conditions.

Attitude towards online collaboration

A repeated-measures analysis of variance enabled us to assess the effectiveness of the introduced IECS by examining differences in changes in students' attitudes across time (pre-post study) between the groups. The results indicated a significant interaction effect between the scripted condition and attitude change before and after the study $F(1, 72)=4.97, p<0.05$, partial $\epsilon^2=.065$. The main effect of the attitude change over time was also significant $F(1, 72)=31.68, p<0.001$, partial $\epsilon^2=.306$. Figure 2 shows a profile plot indicating attitude towards online collaboration before and after the study for both conditions. The students in the IECS condition tend to adopt a more positive attitude toward online collaboration in culturally heterogeneous groups (before the study $M=3.71, SD=.51$, and after the study $M=4.07, SD=.40$) compared to those in the CS condition (before the study $M=3.90, SD=.38$, and after the study $M=4.05, SD=.39$).

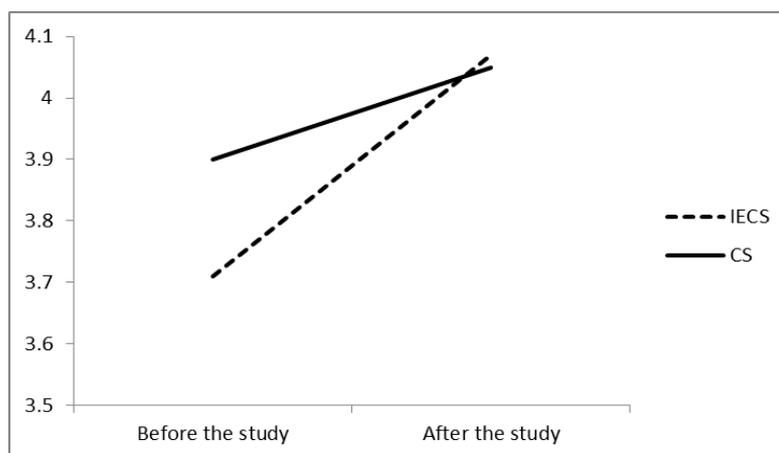


Figure 2. A profile plot indicating attitude towards online collaboration before and after the study in the conditions with the IECS and the CS.

Willingness to collaborate online

Non-significant effect was found on willingness to collaborate online between two conditions ($M=4.02, SD=.59$ for the IECS, and $M=3.77, SD=.75$ for the CS; $F=2.61; p=0.11$). Culturally heterogeneous group members led by the IECS and the CS showed a comparable level of willingness to collaborate online (see Table 2).

Satisfaction with online collaborative learning processes and outcomes

The table 2 shows that the IECS condition ($M=4.19, SD=.48$) has a statistically significant effect on satisfaction compared to the CS condition ($M=3.87, SD=.80$), $F(1,73)=4.41; p<0.05$. Student dyads that were led by the IECS showed higher satisfaction with online collaborative learning processes and outcomes in comparison with those in the CS condition.

Task performance

No significant differences in learning performance were found between the IECS condition ($M=3.17$, $SD=.64$) and the CS condition ($M=3.20$, $SD=.26$), $F=0.02$; $p=0.87$). Student dyads in both condition achieved comparable performance (Table 2).

Table 2. Means and Standard Deviations and Univariate Tests of Significance for willingness, satisfaction and task performance in two conditions the IECS and CS.

Factor	IECS		CS		F	Sig.
	Mean	SD	Mean	SD		
willingness	4.02	.59	3.77	.75	2.61	.11
satisfaction	4.19	.48	3.87	.80	4.41*	.039
task performance	3.17	.64	3.20	.26	.02	.87

* Significant at $p < .05$; ** Significant at $p < .01$.

Concluding remarks

An experimental study with a pre-test/post-test two-group design was conducted to examine the effects of the introduced IECS on the students' collaborative learning in culturally heterogeneous groups in the CSCL environment. When the collected data were analyzed, the following significant findings emerged: (1) the introduced IECS tends to positively affect students' attitude towards online collaboration compared to the condition with CS; (2) students in both conditions exhibited a comparable level of willingness to collaborate online; (3) students in the IECS condition displayed higher levels of online collaboration satisfaction than those students who collaborated in the CS condition; and (4) student groups in both condition achieved comparable performance.

Students in the scripted condition IECS tended to have more positive attitudes toward online collaborative learning in culturally heterogeneous groups in comparison with the students in the CS condition. It seems possible the nature of the instruction offered in the IECS may have contributed to the identified differences. Previous research (e.g. Zhu, 2009; Zhao & McDougall, 2008) suggests that students' collaborative learning dynamic might affect their attitudes towards collaborative learning. By supporting students by means of the IECS instructions from the very beginning of the collaboration (i.e. creating and exchanging personal profiles in order to establish the group) and throughout the whole collaboration process on how to approach a certain subtask, it may be possible that collaborating students did not experience the same challenges or barriers as the students encountered in the condition where only CS was introduced. Thus, this extra facilitation may have led students reporting more positively about their CSCL experiences in such groups.

The second research question addressed the influence of the IECS on students' willingness to collaborate online. In this study, it was assumed that having students follow the script instructions, tailored specifically to two culturally dissimilar group of students, to proceed in a collaborative task, would positively affect their willingness to be engaged in similar collaborative activities in the future. However, the results showed that all student groups regardless of the script condition expressed equally moderate levels of willingness to collaborate online. It could be possibly explained, based on the follow-up survey results, that most students would still prefer a face-to-face format to a computer-mediated communication environment or a combination of both, at least in the early stages of group collaboration.

To investigate the effectiveness of the IECS on learning performance, the scores students received for their joint post-collaboration products were calculated. The results showed that student groups in both scripting conditions achieved similar level of task performance. Apparently, in this study the structure of the task in both conditions did not limit students' creative and critical thinking. The further analysis of the behavioral and part-task data is needed to help us to shed the light on the task performance results.

This socio-technical support was based on combination of previous conceptual contributions about cultural differences and a number of techniques and approaches conducive to successful and productive collaborative problem-solving in the CSCL research field. Taken together, the results of this study suggest that the instructional support of the IECS has proved useful in the context of intercultural online collaborative learning. The theoretical value and practicality of the research discussed in this article rests predominantly on the fact that its methodology is transferable to other educational settings with western-eastern partnerships.

A number of important limitations need to be considered. First, though culture type is proved to be a very important issue, it should not be overestimated as a defining factor and absolute research parameter. Although it is organically accepted that there are individual differences, there have to be and will be certain generalizations, despite which the feasibility and viability of the results is still believed to be in place. Second,

another limitation reflects the brief duration of this study, therefore an important area of research would be to extend the duration of the online collaboration to allow for a longer period of observation and data collection. In view of this, the results showed that satisfaction does not seem to impact task performance, future studies with application of longitudinal design might give new insights into the benefits of having students who are more satisfied with their collaboration.

We hope that the results of this study will lead to a better understanding of collaborative learning in culturally heterogeneous groups in higher education. This line of research will further help educators, researchers and instructional designers to effectively integrate this new approach of instructional technology, which is responsive to culturally diverse learner groups in higher education.

References

- Chen, S., Hsu, C., & Caropreso, E. J. (2006). Cross-cultural collaborative online learning: When the West meets the East. *International Journal of Technology in Teaching and Learning*, 2(1), 17-35.
- Cox, T.H., Lobel, S.A. & McLeod, P.L. (1991). Effects of Ethnic Group Cultural Differences on Cooperative and Competitive Behavior on a Group Task. *Academy of Management Journal*, 34(4), 827-847.
- Curtis, D.D., & Lawson, M.J. (2001). Exploring collaborative online learning. *Journal of Asynchronous Learning Networks*, 5(1), 21-34.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Dillenbourg, P. & Jermann, P. (2007). Designing integrative scripts. Cognitive, computational and educational perspectives. In: F. Fischer, I. Kollar, H. Mandl, & J. Haake (Ed.), *Scripting computer-supported collaborative learning: Cognitive, computational and educational perspectives* (pp. 275-301). New York: Springer.
- Giesbers, B., Rienties, B., Gijsselaers, W. H., Segers, M., & Tempelaar, D. T. (2009). Social presence, Web-videoconferencing and learning in virtual teams. *Industry and Higher Education*, 23(4), 301-310.
- Hinds, P., Zhao, C., Wulf, V., Thomas, J., Fussell, S. and Zhang, Q. (2010). What's up with culture? CSCW 2010 , pp. 607-610, ACM.
- Hofstede, G. (1991). *Cultures and Organizations: Software of the Mind*. London: McGraw-Hill.
- House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (2004). *Culture, Leadership and Organizations: The GLOBE study of 62 societies*. Newbury Park, CA: Sage Publications.
- Koschmann, T. (1999). Computer support for collaboration and learning. *Journal of the Learning Sciences*, 8, 495-498.
- LaPointe, D. K., & Gunawardena, C. N. (2004). Developing, testing, and refining a model to understand the relationship between peer interaction and learning outcomes in computer-mediated conferencing. *Distance Education*, 25(1), 83-106
- Lim, J., & Liu, Y. (2006). The role of cultural diversity and leadership in computer-supported collaborative learning: a content analysis. *Information & Software Technology*, 48(3), 142-153.
- Nisbett, R. E. (2003). *The Geography of Thought: How Asians and Westerners think differently . . .and Why*. New York: Free Press.
- Rummel, N., & Spada, H. (2005): Learning to Collaborate: An Instructional Approach to Promoting Collaborative Problem Solving in Computer- Mediated Settings. *Journal of the Learning Sciences*, 14(2), 201-241
- Thompson, L., & Ku, H. Y. (2006). A case study of online collaborative learning. *Quarterly Review of Distance Education*, 7(4), 361-375.
- Wang, H-C., & Fussell, S. R. (2010). Groups in groups: Conversational similarity in online multicultural multiparty brainstorming. *Proceedings CSCW*, 2010 (pp. 351-360).
- Weinberger, A., Clark, D.B., Hakkinen, P., Tamura, Y., & Fischer, F. (2007). Argumentative Knowledge Construction in Online Learning Environments in and across Different Cultures: a collaboration script perspective. *Research in Comparative and International Education*, 2(1).
- William F. Laurance, D. Carolina Useche, Julio Rendeiro et al. (2012). Averting biodiversity collapse in tropical forest protected areas, *Nature*, 290-294.
- Zhao, N., & McDougall, D. (2008). Cultural influences on Chinese students' asynchronous online learning in a Canadian university. *Journal of Distance Education*, 22(2), 59-80.
- Zhu, C. (2009). *E-learning in higher education: student and teacher variables in the Chinese and Flemish cultural context*. PhD dissertation, University of Ghent, Belgium.

Fostering CSCL Adoption: An Approach to Professional Development Focused on Orchestration

Luis P. Prieto, Yannis Dimitriadis, Juan I. Asensio-Pérez, School of Telecommunications Engineering,
University of Valladolid, Paseo de Belén, 15, 47011 Valladolid (Spain)

Email: lprisan@gsic.uva.es, yannis@tel.uva.es, juaase@tel.uva.es

Sara Villagrà-Sobrino, Iván M. Jorrín-Abellán, Faculty of Education and Social Work, University of Valladolid,
Paseo de Belén, 1, 47011 Valladolid (Spain)

Email: sarena@pdg.uva.es, ivanjo@pdg.uva.es

Abstract: Conceptual barriers are often cited as an important obstacle for the integration of innovative, ICT-enabled practice (such as CSCL) in authentic educational settings. Although professional development (PD) can be used to overcome these barriers, there is a dearth of descriptions of PD programs that foster change towards CSCL practice. This paper presents an approach to PD centered on the notion of orchestrating CSCL activities, and exemplifies it through teacher workshops. The evidence from a mixed methods evaluation of two such workshops highlights the potential of this approach in driving conceptual change, and increasing the perceived feasibility of using CSCL, even in limited-time PD interventions. The results also show the limitations of the approach when used in isolation, hinting at multi-level coordinated actions (e.g. institutional, policy-level ones) to further foster CSCL practice.

Introduction

There exists an increasing concern within the CSCL community about the lack of perceived impact of CSCL research advancements in the everyday practice of our educational institutions (Looi, So, Toh & Chen 2011; Chan, 2011). Within the possible reasons for this lack of adoption, the CSCL community has acknowledged the complexity that coordinating CSCL activities entails, especially for teachers, in authentic educational settings. This coordination has been commonly referred to by researchers as *orchestration* (Prieto, Holenko-Dlab, Abdulwahed, Gutiérrez, & Balid, 2011b; Dillenbourg et al., accepted).

Existing literature on the endeavor of changing the classroom towards ICT-enabled practice has exposed different kinds of barriers, some of them obvious like the lack of ICT infrastructures, but others less apparent, such as teacher beliefs, attitudes and other conceptual obstacles (Ertmer, 1999). In order to overcome such barriers, very often teacher professional development (PD) programs have been proposed (Kagan, 1992; Zhao, Pugh, Sheldon & Byers, 2002; Chai, Hong & Teo 2009; Roschelle et al., 2011).

However, in the particular case of fostering CSCL adoption among teachers, there is a dearth of studies depicting concrete PD approaches and programs. Existing studies generally lack concrete descriptions of the PD actions involved (Chan, 2011; Looi et al., 2011), or do not examine the impact of the PD actions in the teachers' conceptions and actual practice (Zhao & Rop, 2001; Lin, Lin, & Huang, 2008).

In this paper, we propose a PD approach centered on enabling practical application of CSCL in a certain technological and pedagogical context, throughout the different facets and moments of CSCL activity orchestration, combining both conceptual and technological tools. In order to illustrate the approach, we describe in detail one short PD action formed by two workshops. These workshops were aimed concretely at fostering the orchestration of blended CSCL activities in higher education, using Virtual Learning Environments (VLEs, e.g. Moodle (1)) and other web-based tools. We also present a mixed methods evaluation study of the workshops' enactment with 36 teachers from multiple disciplines, focusing on the kind of conceptual change that the PD action brought about (e.g. in beliefs, perception of feasibility and self-efficacy), and assessing the impact in the teachers' actual everyday practice once the PD intervention finished.

Conceptual Change and Fostering CSCL Practice Through PD

Despite huge investments by governments in applying Information and Communication Technologies (ICT) to Education, technology is mostly used to support established practices rather than transform them (Cuban, 2001; Conlon, 2004). Research has shown that providing effective training opportunities for teachers to learn how to effectively redesign education by incorporating ICT is not simple (Lawless & Pellegrino, 2007), and that such training should be embedded into their daily practice (Löfström & Nevgi, 2008; Lawless, et. al., 2007).

Teacher beliefs (e.g. about learning, about technology) have been extensively linked to the integration of ICT in classroom practice (Ertmer, 1999; Orlando, 2009; Prestridge, 2012), including studies in the context of integrating CSCL environments in the classroom (Song & Looi, 2012). Thus, a challenge in fostering CSCL adoption is how to enable teachers not only to overcome technology barriers, but also conceptual ones (Ertmer, 1999; Ertmer & Ottenbreit-Leftwich, in press), empowering them to integrate appropriate technology into the learning process (Mishra & Koehler, 2006).

Kagan (1992) noted the potential of teacher education and PD programs to promote belief change among teachers. Teachers should experiment with the innovation that could involve technology, by means of critical issues discussion with peers and researchers, observing exemplary models, as well as providing them with opportunities for reflecting on the pedagogical beliefs and teaching practices (Chai, et al., 2009). Thus, there is a need to empower teachers to understand and deal with the complexities involved in implementing innovative practices in Technology Enhanced Learning (TEL) settings. PD strategies could provide the catalyst of change in engaging the teachers in collaborative knowledge building, leading to a deep understanding among them about teaching in a CSCL environment (Chai & Tan, 2009).

There is a certain shortage of research examples which describe PD programs in detail, and examine the impact that such actions have in helping teachers to conduct innovations in technology-enhanced learning scenarios. In the CSCL field, there have been efforts in PD that encourage teachers to work collectively through a design-based approach (Putnam & Borko, 2000). Others support teachers' inquiry in TEL settings, from a knowledge building community perspective (Chan & Song, 2010). There are studies of practicing teachers' collaborative online interaction for PD that report quantitative data about teachers' rate of participation (Zhao & Rop, 2001), while others portray the process of knowledge sharing and creation for teachers participating in virtual teams of a teacher professional community (Lin, Lin, & Huang, 2008). Taking an ecological perspective, Chan (2011) proposes actions at different levels to bridge the gap between CSCL research and practice in the context of scaling up and sustaining a knowledge building model. In similar multi-level proposals, from a design research approach, Looi et al. (2011) indicate the empowerment of teachers to orchestrate the classroom as an essential feature of meso-level actions towards CSCL practice.

Roschelle et al. (2011) note that there is a "need to have a new kind of professional development for teachers which [...] aims to develop teachers' potential as innovators". However, none of the aforementioned works depict how such concrete PD actions should be implemented, or study the impact of these meso-level PD actions on teachers after the intervention finishes.

An Orchestration-Focused Approach to PD Towards CSCL Adoption

The professional development approach proposed here draws on several principles and studies outlined above. However, the central concept behind the proposed approach is that of *orchestration*, which has been defined as "the process of productively coordinating supportive interventions across multiple learning activities occurring at multiple social levels" (Fischer & Dillenbourg, 2006). This notion captures the increased complexity of applying CSCL into authentic educational settings, even if there is a lack of general consensus about its main components (see Dillenbourg et al., accepted). Orchestration typically covers the whole lifecycle of the CSCL activity implementation, from its design and preparation to the actual enactment in the classroom (Prieto et al., 2011b). Despite the lack of consensus, researchers are reaching a common understanding that it provides a more holistic view of the problems and constraints involved when applying CSCL principles and technologies to authentic (formal) educational settings in everyday practice (as opposed to dealing with those aspects separately, e.g. in lab settings).

Thus, we propose to take the kind of holistic approach to the application of CSCL that orchestration represents (already hinted by Looi et al., 2011), combining it with PD frameworks such as TPACK (Mishra & Koheler, 2006) and the advice from previous PD research efforts regarding the need for integration into teachers' everyday practice (Löfström et al., 2008; Lawless et al., 2007). The main characteristics of our proposed approach are:

1. **Multi-aspect:** The PD action should address all (or, at least, as many as possible) of the different aspects and dimensions that conform the orchestration of CSCL in the concerned educational setting (e.g. Prieto et al., 2011b mention up to eight orchestration aspects). For example, it should not focus only on assessment techniques, disregarding other aspects like class management or how to adapt the activities in the face of such assessments.
2. **Whole lifecycle:** The PD action should address the whole lifecycle of CSCL activities, from their preparation and design, through their implementation with ICT and their enactment in the classroom, as well as its eventual evaluation and redesign afterwards (as opposed to e.g. centering the PD action only in design, disregarding how such design is afterwards implemented).
3. **Pedagogical and atomic patterns:** One of the main distinctive features of this approach is the way in which orchestration knowledge is made available to teachers. In order to provide starting points in the complex interplay between technology, pedagogy and content that innovative teacher practice requires, we propose to provide teachers with pedagogical patterns (Fincher & Utting, 2002) and atomic patterns (Prieto, Villagrà-Sobrino, Jorrín-Abellán, Martínez-Monés, & Dimitriadis, 2011) elicited from successful (authentic) CSCL practice. These patterns cover and combine the different aspects and phases of orchestration, at multiple levels of granularity.

4. **Technological tools for orchestration:** The PD action should not remain at the level of conceptual change towards CSCL (however important it is), but rather it should also provide hands-on experience with technological tools that are compatible with the principles being taught (and with the educational setting in which the teachers develop their practice).
5. **Modeling:** In order for teachers to get a more accurate idea of how the CSCL activities are enacted, and how the experience is like for the students, the PD action itself should be modeled using the same patterns, techniques and technologies that are being taught in the PD action.
6. **Authentic problems:** During the PD action, teachers should be able to work on problems that are authentic and relevant for themselves, e.g., defining how the orchestration strategies and technologies could be applied to a concrete course they teach, within their current contextual constraints.

In order to illustrate how this approach may be applied in practice to develop a PD action, Table 1 depicts the activities of two PD workshops developed following this approach, which aim at promoting blended CSCL practice at the university level, using interactive digital whiteboards, Virtual Learning Environments (VLEs, e.g. Moodle) and other Web 2.0 tools as the main technological support. These workshops combined Collaborative Learning Flow Patterns (Hernández-Leo, Villasclaras-Fernández, Dimitriadis & Asensio-Pérez, 2010) and atomic patterns elicited from blended CSCL practice with VLEs, as well as specific technological tools for the orchestration of such activities: the WebCollage (2) authoring tool and the GLUE!-PS system (3) to deploy and manage learning designs in VLEs. The concrete form of this design was also influenced by local contextual settings such as the emphasis on Moodle (the official institutional VLE), or the time-frame of the workshops (12 hours each, agreed with the university PD agency). Indeed, these workshops have been actually enacted in 2012, and the following section describes a mixed methods study performed during such enactment, to explore the potentialities and limitations of the approach.

Table 1: Design of two workshops as example of the orchestration-centered approach to foster CSCL practice.

Workshop 1 – conceptual emphasis		Workshop 2 – technological emphasis	
Task/Phase	Notes	Task/Phase	Notes
1. Read a sample CSCL scenario, answer questionnaire on initial thoughts	Online. Scenario is hypothetical, but plausible for the audience's teacher practice	1. Read a sample CSCL scenario and design, answer questionnaire on initial thoughts	Online. Scenario is hypothetical, but plausible for the audience's teacher practice. Learning design is derived from previous workshop (Pyramid pattern)
2. Brief explanations, design using the Pyramid pattern and design- and deploy-time atomic patterns	Face-to-face. Teachers activities in the session follow several phases that conform a Pyramid pattern similar to the one they are designing. Use of pen, paper and atomic pattern cards	2. Brief explanations, design with WebCollage and deploy with GLUE!-PS, on Moodle	Face-to-face. Teachers use the technologies involved in dyads. Teachers are scaffolded with a worksheet detailing tool usage
3. Propose similar design for each teacher's own courses	Online. Teachers work individually. Design is expressed freely, but should involve Moodle+Web 2.0 tools. Facilitators provide formative feedback on individual designs	3. Propose similar design for each teacher's own courses, and implement it in WebCollage	Online. Teachers act individually, using the provided WebCollage tool and worksheet. Facilitators provide formative feedback on the designs' feasibility and alternatives
4a. Reflect/debate on main features of submitted designs	Face-to-face. Multiple parallel debates following Think-Pair-Share pattern. Debates are traced through ICT tools mentioned in the workshop (e.g. Moodle, Google Docs)	4a. Finish deployment of individual design using GLUE!-PS and favourite VLE	Face-to-face. Teachers act individually, supported by the worksheets and facilitators
4b. Role-play scenario enactment, supported by enactment-time atomic patterns	Face-to-face. Represent a set of problematic situations (e.g. latecomers, ICT failure). Teachers in 6-people teams with different roles (teachers, students, observers/critics)	4b. Face problematic situations, and adapt designs supported by GLUE!-PS tool	Face-to-face. Teachers work in dyads. Worksheets of the changes needed are not provided until <i>after</i> the session (for future reference)
5. Second iteration of individual design, and questionnaires for reflection and evaluation	Online. Again, facilitators provide formative feedback on the submissions	4c. Reflect/debate on applicability to everyday practice	Face-to-face. Teachers form 6-person focus groups which discuss in parallel, and report their conclusions to the whole group (orally and through a shared document)
		5. Second iteration of individual design, and questionnaires for reflection and evaluation	Online. Again, facilitators provide formative feedback on the submissions

Studying CSCL Adoption in Teacher Workshops

Context and Methodology of the Study

Multiple CSCL researchers have highlighted the adequacy of using mixed method approaches (Creswell, 2012) in order to explore the different perspectives and multiple factors that affect CSCL situations (e.g. Strijbos & Fischer, 2007). This advice is also provided by the Evaluand-oriented Responsive Evaluation Model (CSCL-EREM, see Jorrín-Abellán and Stake, 2009), which we have used to design and structure our evaluation. Regarding the data gathering and analysis, we have adapted the mixed method evaluation approach described by Martínez-Monés, Dimitriadis, Gómez-Sánchez, Jorrín-Abellán, Rubia-Avi & Marcos-García (2006), in which quantitative and qualitative data gathering techniques are combined, triangulating the available evidence to enhance its credibility. The concrete flow of data gathering and analysis techniques used is depicted in Figure 1, including multiple techniques and data sources used before, during and immediately after the two workshops (e.g. eight interviews with participants, parallel focus groups during the workshop, observations, questionnaires, teacher-generated artifacts). Additionally questionnaire was provided to the participants six months after the second workshop, in order to assess medium-term effects of the PD action.

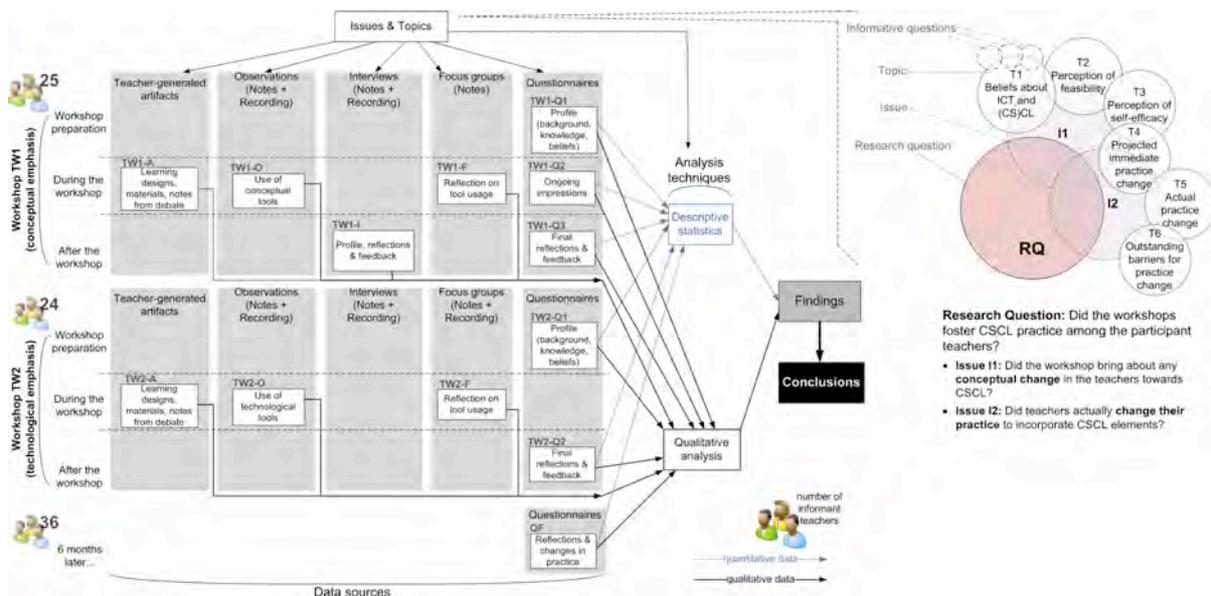


Figure 1. Data gathering and analysis flow of the study (left), adapted from Martínez-Monés et al. (2006); also includes an “anticipated data reduction” diagram (right), inspired from Miles and Huberman (1994).

Also, following the recommendations by Dillenbourg (2009) regarding research on orchestration, we have tried to evaluate the effects of our approach in authentic PD actions, situated within the usual activities of teachers. In this case, the two workshops, designed as described in the previous section, were enacted by a team of 5 researchers, between February and April 2012, as two separate 12-hour teacher workshops aimed at in-service university teachers at University of (*anonymized for review*): the first one aimed at conceptual training on CSCL and its orchestration; the second one centered on technological tools to make such orchestration possible, in blended learning scenarios using the official university Virtual Learning Environment (VLE). The workshops followed a blended learning format and were open for teachers from any discipline (although it was recommended that, in order to take the second workshop, teachers should have completed the first one, or an equivalent training in basic CSCL concepts). As indicated in Figure 1, 25 teachers from different disciplines (from Engineering to Medicine, Education or Law) attended the first workshop, and 24 attended the second one (with partial participant overlap between them – in total, 36 different teachers attended at least one of the workshops). This lack of consistency between the workshop’s participants was due to the situatedness of the PD action, which prevented the researchers from restricting access or enforcing attendance to both workshops.

Within this context, our main research question in this study (“did the workshops foster CSCL practice among the participant teachers?”, see Figure 1) is explored through two main evaluative tensions, or issues, regarding the conceptual change towards CSCL (I1) and the actual changes in teacher practice (I2) that the workshops brought about. These issues are in turn explored through the evidence gathered, which is grouped around six topics: teachers’ beliefs about ICT and CSCL (T1), teachers’ perceived feasibility of CSCL practice (T2), teachers’ perceived self-efficacy to orchestrate CSCL (T3), the immediate practice changes that teachers expected (T4), the actual teacher practice changes (T5) and the outstanding barriers for change towards CSCL practice (T6). This kind of “anticipated data reduction” schema (Miles et al., 1994) has been used to analyze the quantitative and qualitative evidence.

Findings and Supporting Evidence

In general, both workshops were considered a success by both the facilitators and the participant teachers (e.g. the two PD actions were valued at 8.42 and 9.23 respectively, in a 10-point scale [TW1-Q3, TW2-Q2], see Figure 1). Let us examine the evidences of conceptual and practice change brought about by the workshops:

Conceptual change towards CSCL adoption (I1)

Concerning the topic of the conceptual change achieved by the workshops in the area of the teachers' beliefs and attitudes towards ICT (topic T1), a majority of the participants in the first workshop reported some kind of conceptual change regarding (computer-supported) collaborative learning (17 out of the 22 participants that answered the post-workshop questionnaire - 77%) [TW1-Q3]. Among these, a majority expressed changes in their concept of collaborative learning in general (65%), like for example: "*[to the question: Did the workshop change your view of collaborative learning?] Totally. I had always reduced collaborative work to working in dyads, and essentially, I thought they were worthless. The success of this course is that I have seen that they can be really useful – if well designed – for the learning processes.*" [TW1-Q3], a fact that was also confirmed in the interviews: "*[when asked about usage and knowledge about CL prior to the workshop] I have seen that, if you do it well, you can learn a lot with it [...] designing it with time, thoughtfully, and applying these patterns, you can get more outcomes from a subject matter.*" [TW1-I]. Others highlighted the down-to-earth view given by the workshop, which made inherent difficulties of collaborative work surface (24%), or the importance of making a careful pedagogical design (17%): "*[to the question: Did the workshop change your view of collaborative learning?] [...] Maybe the most interesting was [learning about] the preparation of collaborative activities and seeing what difficulties they entail.*" [TW1-Q3]. Regarding those participants that reported that their views had not changed, in some cases it was because their views were already favorable to CSCL, while others voiced unchanged concerns on the efficiency and uncertainty of such methods, e.g.: "*[when asked about the feasibility of using the provided pedagogical patterns] It demands a time that I am not sure I have, and a big effort in collaborating, all to reach uncertain results.*" [TW1-Q2].

There was also evidence of teachers changing their perception of the feasibility of using CSCL in their everyday practice (topic T2), from the qualitative responses to the questionnaires, as well as from the in-depth interviews "*[to the question: Did the workshop change your view of collaborative learning?] The most important aspect is that real cases are presented and, from them, it is easier to gather ideas that are easily transferable to my courses. I came out with the impression that I can put in practice a collaborative activity, really; up to now it did not occur to me anything other than ordering a report and making students present it in class.*" [TW1-Q3]. Many participants emphasized the role of atomic patterns in this perception: participants in the first workshop valued the different kinds of atomic patterns as useful (averages of 4.30, 4.13 and 5.31 in a 1-7 Likert scale), and this was also brought to attention in some of the interviews "*[when asked about the main value of the workshop] the catalogue of patterns and the catalogue of routines [...] as a reference guide or a skeleton to structure activities, to begin making things*" [TW1-I]; "*[when asked about the added value of atomic patterns] It is the founding, the structure, the skeleton [...] you may have some ideas, but if you don't know anything about what to do [...] it is like LEGO blocks*" [TW1-I]. Others highlighted the fact that the workshop itself was modeled as the same kind of blended CSCL activity that was being promoted: "*[when asked about the workshop's usefulness] [...] Very useful, playing the "guinea pigs" like students in class, enables you to see points of view that you had forgotten.*" [TW1-Q3].

Regarding the teachers' perception of self-efficacy (topic T3), we observed how all the participants were able to orchestrate the example scenario given (in the first workshop), and implement such orchestration through the provided technological tools (in the second workshop). This ability was then transferred to the individual exercises of application to each teacher's courses in a large majority of the cases (100% and 75%, respectively) [TW1-A, TW2-A]. Even if we look solely at the evidences of the first workshop, we can see how the self-perceived ability to orchestrate the design they were doing throughout the workshop raised from an average of 4.83 (in a 1-7 Likert scale) to 5.23 after the first workshop [TW1-Q2, TW1-Q3]. Teachers mentioned the advantages of the workshop's practical orientation: "*[to the question: Did the workshop change your view of collaborative learning?] Now I have a much clearer idea of what is a collaborative activity and, especially, which patterns to use to encourage that those activities are really collaborative. [...] how to complement these learned [strategies] with the use of ICT, the ICTs that can be used and some of their affordances.*" [TW4-Q4]; "*[when asked about the changes brought about by the workshop] now I'm more convinced [...] because now I know how to put it in practice [...] up to know you never saw concrete examples, tasks.*" [TW4-I]. It is also worth noting that many teachers viewed the presented atomic patterns as not very novel, and reported using common sense rather than the provided strategies (especially in the role-played enactment, which required a timely response to events), a view that seems to be somehow tied to prior teacher experience: "*[when talking about the usage of enactment-time atomic patterns during roleplaying] I think we went directly to common sense [...] I also did the same when doing the individual [design], I used common sense and then asked myself*

'how is this called in the routines?' [...] and, as you look through them, you also see other [routines] and you open your mind a bit [...] expand common sense.' [TW1-I] [18 years of teaching experience].

Actual Changes in Teacher Practice (I2)

Regarding the actual changes in teacher practice towards CSCL since the second workshop ended (six months, as of this writing), the evidence is less conclusive, and can be analyzed at different levels. If we look at teachers' perception and intentions of changing their practice immediately after the workshops (topic T4), we find that teachers scored atomic patterns as close to their everyday practice (average of 5.22 in a 1-7 Likert scale) [TW1-Q2], and that teachers asserted they would likely use the proposed tools in their everyday practice in the immediate future (in a 1-8 Likert scale, avg=6.19, std=1.63) [TW2-Q2]. Moreover, teachers asserted they would be moderately likely to use the designs they deployed in Moodle in the immediate future (in a 1-8 Likert scale, avg=6, std=1.64) [TW2-Q2]. Qualitative responses included favorable assertions such as: *"[when asked about using the patterns/routines when doing a CL design in future teaching practice] Sure. Now that I have them, I would take a look at them [...] I would use this [method] I know"* [TW4-I], but also critical voices of the most skeptical about collaborative work: *"[when talking about the usage of routines in the individual design] The truth is I didn't [use them] [...] people want to appease the teacher who ordered the exercise [...] if you taught 15 things, they think they should introduce at least 10 of them"* [TW4-I].

The evidence of actual changes in teachers' everyday practice (T5) are much less prominent. Only one teacher reported putting the strategies into practice immediately after the first workshop: *"[when asked about the eventual usage of atomic patterns in real practice] Yes, yes... in fact, the next day I tried, not this [the individual design] but I did a pyramid in class [...] With pen and paper"* [TW4-I]. The questionnaire taken 6 months after the workshop reveals that not many teachers (5 out of 14 respondents – 35.7%) had incorporated elements of the workshop into their everyday practice [QF]. This apparent lack of significant impact can be attributed to recent massive lay-offs and subsequent re-organization at the university (*"[when asked whether the workshop's strategies had been put into practice] No. Without knowing the courses I was going to teach [...] The changes in the courses and their quantity [...] prevented me from trying new techniques."* [QF]), or to the lack of institutional support for some of the technological tools provided in the workshop.

However, leaving aside those circumstantial events, we can also look at which barriers teachers see for the adoption of CSCL strategies and technologies such as the ones presented in the workshops (topic T6). Some of the barriers mentioned are well known, such as the lack of adequate ICT infrastructures, training or technical support. Others relate more to teacher beliefs about ICT technologies in general (*"[when outlining one focus group's conclusions about whether they would use WebCollage/GLUE!-PS in real practice] some of us would not use it because we think [using ICTs] is complex and not necessary, versus other simpler ways of doing collaborative work"* [TW2-FG]), or towards collaborative learning (*"[when outlining one focus group's conclusions about run-time problems in blended CSCL] we have also considered whether these collaborative activities may subtract from the [content] learning [...] we should not mistake the means with the goals"* [TW2-FG]). Teachers were especially concerned with regard to the increased time and effort needed to apply (CS)CL in practice: *"[when talking about doing CSCL in a VLE as in the example courses shown] When I see all those links, folders, etc [...] I think it is too much work, non-pedagogical work involved [...] does it all compensate in order to reach what? To make a summary? [...] I'm a skeptic."* [TW1-I]. Many others expressed beliefs more related to students, especially their attitudes towards collaboration, or the sheer number of participants in some of the classes (*"[In the debate, when asked about collaboration in large groups] Too much effort for the teacher"* [TW1-A]).

Discussion

The quantitative and qualitative evidence provided above shows how a PD action based on the proposed approach can bring about conceptual change in teachers at different levels (in their beliefs about CSCL, as well as their perception of feasibility and self-efficacy). This is even more remarkable given the limited time-frame of the PD action (24 hours of blended work). There are, however, comparatively few instances of actual practice change towards CSCL practice among teachers (six months after the workshops).

Our evidence also found hints of the influence of multiple factors in the amount and nature of such conceptual change (as it is often the case in the "messy" environment of an authentic situation): the amount of teaching experience, prior beliefs about ICT, collaborative learning and CSCL, or how the workshop itself was orchestrated by the facilitators. Teachers also voiced several barriers for the adoption of the concerning CSCL practices, some of which are well known (teacher beliefs, ICT infrastructure, lack of training or support). Other barriers, however, are more striking, such as teachers' perception of students attitudes towards collaborative work, or the general consideration that collaborative work requires lots of effort and might not be feasible in larger student groups.

Our findings seem to support Looi et al. (2011)'s consideration that multiple factors limit teacher innovation in the classroom (which they adequately illustrate with the metaphor of a broken barrel which only

holds as much water as its shorter wood board). In the concrete case of our workshops, we could speculate that the potential for conceptual and practice change of the PD action is trampled over by other factors that such a short intervention cannot overcome (e.g. current instability of teaching loads at the university, tendency towards larger classrooms). This points towards the convenience of actions coordinated at different levels (policy, institutional, classroom) (Chan, 2011). Interestingly, the persistence of perceived barriers regarding the effort needed to setup and manage collaborative activities seems to indicate that further efforts on technologies and strategies for easing such orchestration are needed, in order to make CSCL practice less cumbersome, especially in large cohorts.

Naturally, due to the fact that it was gathered by observational methods in a situated PD action, these findings are hardly generalizable (nor was it our aim to make them so) to other contexts. The lack of consistency in the attendance to the two workshops is another weakness which limits our ability to trace clear evolutions of individual teacher conceptions and practices. Our naturalistic evaluation, however, provided hints of several factors influencing the impact of this kind of PD actions, which should be examined more thoroughly in later studies.

Also, the evidence from our study, along with the current evolution in many university settings towards larger cohorts, prompt us to iterate over the format and materials of this workshop, modifying the strategies and the technologies we are developing, e.g., in order to make more efficient the orchestration of large student groups. This is one of our main directions for future work in the near future, in accordance with a recently approved international R&D project.

Overall, we have presented a novel approach to teacher professional development that intends to foster CSCL practice within authentic formal educational settings, with an emphasis on the holistic approach known in CSCL as orchestrating learning. This emphasis is reified in the usage of conceptual tools that highlight possible synergies applicable to teachers' contexts (e.g. atomic patterns extracted from successful CSCL practice in authentic settings), and by the problem-orientedness and authenticity of the PD activities. By presenting an example of such PD action and initial evidences of its usefulness for conceptual and practice change, we highlight the potential usefulness of this kind of actions to foster CSCL adoption, even if coordinated actions at other levels are needed as well in order to provide long-standing impact in our educational systems.

Endnotes

- (1) <http://www.moodle.org> (Last visit: 25 Oct 2012)
- (2) <http://pandora.tel.uva.es/wic2/> (Last visit: 25 Oct 2012)
- (3) <http://gsic.uva.es/glueps/> (Last visit: 25 Oct 2012)

References

- Chai, C. S., Hong, H. Y., & Teo, T. (2009). Singaporean and Taiwanese pre-service teachers' beliefs and their attitude towards ICT: A comparative study. *The Asia-Pacific Education Researcher*, 18(1), 117–128.
- Chai, C. S., & Tan S. C. (2009). Professional Development of Teachers for Computer Supported Collaborative Learning (CSCL) through Knowledge Building. *Teacher College Records*, 111(5), 1296-1327.
- Chan, C. K. K. (2011). Bridging research and practice: Implementing and sustaining knowledge building in Hong Kong classrooms. *International Journal of Computer-Supported Collaborative Learning*, 6(2), 147–186.
- Chan, C. K. K., & Song, Y. (August 3–6, 2010). Towards a principle-based approach for knowledge creation in teacher professional development. Paper presented at the Summer Institute, University of Toronto, Canada.
- Creswell, J.W., & Plano Clark, V.L. (2012). *Designing and conducting mixed methods research*. Thousand Oaks (CA): Sage Publications.
- Cuban, L. (2001). *Oversold and underused*. Cambridge, MA: Harvard University Press.
- Conlon, T. (2004). A Failure of Delivery: the United Kingdom's New Opportunities Fund programme of teacher training in information and communications technology. *Journal of In-Service Education*, 30(1), 115-140.
- Dillenbourg, P. (2009). Exploring neglected planes: social signs and class orchestration. *Proceedings of the International Conference of Computer-Supported Collaborative Learning (CSCL 2009)*, 6-7.
- Dillenbourg, P., Dimitriadis, Y., Nussbaum, M., Roschelle, M., Looi, C-K., Asensio, J.I., et al. (accepted). Design for classroom orchestration. *Collective paper at Computers & Education*.
- Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: Strategies for technology integration. *Educational Technology Research and Development*, 47(4), 47-61.
- Ertmer, P. A., & Ottenbreit-Leftwich, A. (in press). Removing obstacles to the pedagogical changes required by Jonassen's vision of authentic technology-enabled learning. *Computers & Education*.
- Fincher, S., & Utting, I. (2002). Pedagogical Patterns: their Place in the Genre. Paper presented at the ITiCSE 2002, Aarhus, Denmark.

- Fischer, F., & Dillenbourg, P. (2006). Challenges of orchestrating computer-supported collaborative learning. Paper Presented at the 8Th Annual Meeting of the American Educational Research Association (AERA).
- Hernández-Leo, D., Villasclaras-Fernández, E. D., Dimitriadis, Y., & Asensio-Pérez, J. I. (2010). Pattern languages for generating CSCL scripts: From a conceptual model to the design of a real situation. In P. Goodyear, & S. Retalis (Eds.), *E-learning, design patterns and pattern languages* (pp. 49-64) Sense Publishers.
- Jorrín-Abellán, I. M., & Stake, R. (2009). Does ubiquitous learning call for ubiquitous forms of formal evaluation?: An evaluand oriented responsive evaluation model. *Ubiquitous Learning: An International Journal*, 1(3), 71-82.
- Kagan, D. M. (1992). Implications of research on teacher belief. *Educational Psychologist*, 27(1), 65-90.
- Lawless, K. A., & Pellegrino, J. W. (2007). Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. *Review of Educational Research*, 77(4), 575-614.
- Lin, F., Lin, S., & Huang, T. (2008). Knowledge sharing and creation in a teachers' professional virtual community. *Computers & Education*, 50(3), 742-756.
- Looi, C.K., So, H.J., Toh, Y., Chen, W. (2011). The Singapore experience: Synergy of national policy, classroom practice and design research. *International Journal of Computer Supported Collaborative Learning (CSCL)*, 6(1), 9-37.
- Löfström, E., & Nevgi, A. (2008). University teaching staffs' pedagogical awareness displayed through ICT-facilitated teaching. *Interactive Learning Environments*, 16(2), 101-116.
- Martínez-Monés, A., Dimitriadis, Y., Gómez-Sánchez, E., Jorrín-Abellán, I.M., Rubia-Avi, B., & Marcos-García, J.A. (2006). Studying participation networks in collaboration using mixed methods. *International Journal on Computer Supported Collaborative Learning*, 1(3), 383-408.
- Miles, M. B., Huberman, A.M. (1994). *Qualitative data analysis*. Newbury Park (CA): Sage Publications.
- Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Records*, 108(1), 1017-1054.
- Orlando, J. (2009). Understanding changes in teachers' ICT practices: A longitudinal perspective. *Technology, Pedagogy and Education*, 18(1), 33-44.
- Prestridge, S. (2012). The beliefs behind the teacher that influence their ICT practices. *Computers & Education*, 58(1), 449-458.
- Prieto, L. P., Villagrà-Sobrino, S., Jorrín-Abellán, I. M., Martínez-Monés, A., & Dimitriadis, Y. (2011). Recurrent routines: Analyzing and supporting orchestration in technology-enhanced primary classrooms. *Computers & Education*, 57(1), 1214-1227.
- Prieto, L. P., Holenko-Dlab, M., Abdulwahed, M., Gutiérrez, I., & Balid, W. (2011b). Orchestrating Technology Enhanced Learning: a literature review and a conceptual framework. *International Journal of Technology-Enhanced Learning (IJTEL)*, 3(6), 583-598.
- Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, 29(1), 4-15.
- Roschelle, J., Patton, C., Schank, P., Penuel, W., Looi, C.-K., Chen, W., Prieto, L.P., Villagrà-Sobrino, S., Dimitriadis, Y. (2011). CSCL and innovation: In classrooms, with teachers, among school leaders, in schools of Education. In *Proceedings of the 9th International Computer-Supported Collaborative Learning Conference (CSCL 2011)*, Vol. 3, 1073-1080.
- Song, Y., Looi, C.K. (2012). Linking teacher beliefs, practices and student inquiry-based learning in a CSCL environment: a tale of two teachers. *International Journal of Computer Supported Collaborative Learning*, 7(1), 129-159.
- Strijbos, J.W. & Fischer, F. (2007). Methodological challenges for collaborative learning research. *Learning and Instruction*, 17(4), 389-393.
- Zhao, Y., Pugh, K., Sheldon, S. & Byers, J. L. (2002). Conditions for classroom technology innovations. *Teachers College Record*, 104(3), 482-515.
- Zhao, Y., Rop, S. (2001). A critical review of the literature on electronic networks as reflective discourse communities for in-service teachers. *Education and Information Technologies*, 6(2), 81-94.

Acknowledgements

This research has been partially funded by the Spanish Ministry of Economy and Competitiveness Projects TIN2008-03-23, TIN2011-28308-C03-02 and IPT-430000-2010-054, and the Autonomous Government of Castilla and León Project VA293A11-2, as well as by the METIS European Project (531262-LLP-2012-ES-KA3-KA3MP). The authors thank the participants in the teacher workshops and the rest of the GSIC/EMIC research team, as well as the Learning Design Theme Team funded by the European Union through the Stellar Network of Excellence for Technology Enhanced Learning (FP7-IST-231913) for their contributions.

The Effect of Formative Feedback on Vocabulary Use and Distribution of Vocabulary Knowledge in a Grade Two Knowledge Building Class

Monica Resendes, Bodong Chen, Alisa Acosta, Marlene Scardamalia,
OISE/University of Toronto, 252 Bloor Street West, Toronto, ON, Canada
Email: monica.resendes@utoronto.ca, bodong.chen@utoronto.ca, alisa_acosta@yahoo.ca,
marlene.scardamalia@utoronto.ca

Abstract: This study examines the impact of formative feedback to enhance students' productive written vocabulary. Behavioral, lexical, and network structure analyses were applied to the work of two Grade 2 classes engaged in knowledge building in science. Two variations of feedback including vocabulary and contribution-based visualizations were integrated into the knowledge building practice of the experimental class. Behavioral and lexical measures were calculated with automated tools, and content analysis was used to evaluate depth of understanding. Moreover, the degree of vocabulary distribution throughout the communities was explored. Findings show that formative feedback embedded in knowledge building practices can help students grow their vocabulary, apply new words in productive ways in their writing, and advance community knowledge. Results also show that as students learn and use a more diverse range of words in the context of knowledge building, the more discursively connected they become, and the greater the knowledge distribution across the community.

Introduction

Literacy and the ability to work creatively with ideas are essential competencies for students to develop to become productive citizens in a knowledge society. Literacy as a foundational component of a 21st century education is emphasized in contemporary educational initiatives such as the "Partnership for 21st Century Skills" (see <http://www.p21.org>). As evident in this framework, literacy crosses all domains and underlies not only core content learning, but also the ability to innovate and collaborate, and to engage effectively with media and technology. In a broad sense, literacy entails an ability to read and write with understanding, use information productively from a range of sources, as well as use language effectively to build and communicate ideas. From a socio-cognitive perspective, developing literacy requires integrating language learning within authentic pedagogical practices that embed language use within inquiry and problem solving processes (Applebee, 1981; Bereiter and Scardamalia, 1987). This study explores an integrated approach to language learning that engages young students in creative knowledge work together with literacy practices. It examines how formative feedback supports designed to boost knowledge building discourse impact literacy skills, particularly growth in students' productive written vocabulary. The study also examines the extent to which new and important terms are distributed throughout shared discourse as students worked to collaboratively build knowledge in science.

Knowledge Building for Vocabulary Learning

Knowledge building pedagogy (Scardamalia & Beretier, 2003) is a socio-cognitive approach that can be described as "the production and continual improvement of ideas of value to a community" (p. 1370). This approach places advancement of community knowledge as the explicit and shared goal (Scardamalia, 2002). In knowledge building, students work together to participate in creative work with ideas in the effort to produce increasingly coherent explanations to shared problems of understanding. Knowledge building practices are enhanced by Knowledge Forum, an online environment specifically designed to support high-level knowledge work (Scardamalia, 2004). Knowledge building affordances embedded within the environment include the ability to co-author, reference, or build-on notes; scaffolds support high-level discourse moves such as "My theory," "I need to understand," or "This doesn't explain" to help frame thinking and writing, automated assessment tools support evaluation and exploration of discourse. Both online and offline, a knowledge building approach fosters collaboration and creative knowledge work, with shared discourse as its driving force. As such, it provides a rich context to engage students' in authentic literacy practices that involve individual and co-operative reading, writing, idea development, active research, and sustained collaborative dialogue (see Sun, Zhang, & Scardamalia, 2010).

As elaborated above, literacy is bound up with processes of productive knowledge work. A critical aspect of literacy involves use and growth of vocabulary. Research has shown that greater knowledge and use of vocabulary is a reliable predictor of reading and writing comprehension (Stahl, 1991) as well as verbal and

listening skills (Steahr, 2009). Studies also show that learning a new word is not a singular event, but happens over time, with increased and varied usage indicating deeper understanding (Nation, 2001). Integrated contexts of literacy that promote productive vocabulary use and growth thus engage students in meaningful activities related to new or difficult words, expose them to multiple and varied encounters with these words, and give them opportunities to utilize such words in speaking, reading, writing and listening (Stahl, 1991). Authentic literacy practices engage students in these activities not only in the interest of language acquisition, but in the service of authentic inquiry and problem solving; such instructional environments have been shown to be more effective for language learning than direct instruction with respect to depth of word knowledge, writing quality and expansion of vocabulary (Yonek, 2008; Stahl, 1991). With its focus on immersing students in shared discourse for solving problems of understanding, knowledge building practices present conditions highly conducive for effective vocabulary learning. Students are offered rich opportunities to introduce new vocabulary within inquiry-based work, negotiate and infer word meanings, and use available sources to help them deepen their knowledge of new words.

In this study, we explore the discourse of two grade 2 classes as they each engaged in two knowledge building units in science, with a focus on the life cycles of birds and salmon. We focus on development of productive written vocabulary as evidenced in students' writing on Knowledge Forum. Productive use of vocabulary entails that students display a diverse range of words in their writing in a way that conveys understanding. Richness in student vocabulary includes use of both domain-specific and epistemological terms or "academic words" (Coxhead, 2000). Productive use of domain-specific vocabulary is indicative of grasping core content and language, with frequent use of domain specific words indicative of integration into a discursive community (Chernobilsky et al., 2004). Similarly, "academic words" (eg. source, theory, hypothesis) refers to terms that occur at a reasonably high frequency rate in academic discourse; these words cross domains and generally correspond with higher level knowledge work. Academic words typically appear in students' discourse at a relatively late age, beginning in adolescence and increasing with post-secondary education (Laufer, 1994).

So, is it plausible to expect children of primary school age to use sophisticated vocabulary in their written work? According to research on reading progression (Chall, 1996), the spectrum of learning across which both reading comprehension and vocabulary usage take place is characterized by important developmental changes. According to this framework, in primary level grades students are still "learning to read"—gaining foundational phonetic knowledge—rather than "reading to learn", which involves higher level cognitive processes and does not begin to take place until approximately grades 4-6 (Chall, 1996). However, this progression is not a rigid series of sequential stages, but an overlapping continuum that is based on approximate grade and age levels; furthermore, the developmental steps are dependent to a considerable extent upon the learning environment itself (Chall, 1996).

Research shows that exposing students to specialized fields of discourse on a repeated basis in authentic language-learning settings can help foster the productive use of sophisticated words (Corson, 1997). Immersing students in settings that include speaking and listening along with reading and writing is particularly beneficial for lower-level readers (Beimiller, 1999). Similarly, research shows that even with a single exposure, a word encountered in a richer context is more likely to be learned than is one in a less rich context (Herman, Anderson, Pearson, & Nagy, 1987). Combining reading and writing activities with explicit vocabulary learning has been shown to be a highly effective strategy for language learning (Stahl & Fairbanks, 1986). In addition, the use of formative assessments to enhance learning is widely recognized (Black & William, 1998; Stiggins, 2004; Marzano, 2006). Formative assessments integrated within computer-supported learning environments have also been shown to be beneficial for learning (Tseng & Tsai 2007). Moreover, studies show that vocabulary-based feedback such as word or tag clouds provide useful overviews of knowledge that highlight key concepts (Hearst and Rosner, 2008) and aid in semantic exploration and comprehension of data by users (Bateman, Gutwin, and Nacenta (2008).

These findings support the notion that even students as young as the second grade can learn and use complex vocabulary productively if conditions and resources are conducive to such learning. A knowledge building approach has been shown to provide such conditions. For example, research has shown gains in vocabulary and comprehension as by-products of collaborative and creative work with ideas—knowledge building contexts with no direct focus on vocabulary learning and text comprehension (Scardamalia et al., 1992). Furthermore, children as young as junior kindergarten have shown gains in literacy using this approach (Pelletier, Reeve, & Halewood 2006). Looking at vocabulary growth in knowledge building students across grades 3 and 4, Sun et al. (2010) traced an increase of use of academic words of almost four percent on average, and found positive correlations between use of sophisticated vocabulary with depth of understanding. Where benefits in knowledge building work for literacy are reported, this study will be the first to focus on the role of formative feedback targeted to enhance students' vocabulary knowledge. Moreover, examining students' knowledge building calls for collaborative, emergent knowledge advancement in addition to individual assessments. This study will also

address learning gains at the group level by looking at the underlying network structure of rich vocabulary in the collective discourse to determine the extent to which vocabulary use is distributed throughout the community.

Method

Participants and Classroom Context

Participants for this study include 44 Grade 2 students attending the Dr. Eric Jackman Institute for Child Study in downtown Toronto. Two consecutive Grade 2 classes were studied—22 students (11 boys, 11 girls) from the 2010-11 grade 2 class, and 22 students (11 boys, 11 girls) from the 2011-12 school year. Both classes were taught by the same teacher, and engaged in the same activities for each knowledge building unit. For their knowledge building sessions, both Grade 2 classes split up in a rotation in which half the students went to the library and the other half engaged in inquiry. The 2010-2011 class was not subject to any treatments and provides what we will call the “benchmark” class. The 2011-2012 class functions as the experimental class. Within this class, two student groups (Group A and Group B) each received a different variation of the treatments, which are elaborated below.

Both Grade 2 classes in this study participated in a four month “Bird Study” knowledge building unit followed by a 4 month unit on “Salmon”. For both units, the Grade 2 students typically had one 45-minute session a week dedicated to knowledge building, referred to as “KB” time. During this period, students engaged in active research or whole group “KB talks” in which they discussed questions, ideas, and so on, related to their area of study. Typically, students were given 20 minutes to enter their ideas, questions, theories, etc., into the Knowledge Forum community space. For both units of study, students engaged in active research and used a variety of sources, including books, websites, and videos, to increase their knowledge on birds and salmon. Students in both classes also examined objects such as owl pellets, feathers, and nests, as well as raised salmon in a classroom tank as part of the “Lake Ontario Salmon Restoration Program”. Thus, students in both years had rich environments to support their knowledge building work. Although students were split into groups during knowledge building time in both classes, all students in a single class worked in the same knowledge space.

Design

Four knowledge building principles served as important design elements for this research, informing the two different treatments embedded in the knowledge building practices of the experimental class:

(i) Knowledge Building Discourse: This type of discourse constitutes collaborative dialogue that focuses on continual refinement and improvement of ideas and advances through a community’s continued efforts to deal with puzzling facts. An element of knowledge building discourse includes occasional periods of reflection on the state and direction of the community’s discourse itself. “Meta-discourse” can be described as discussion about discussion, and calls for community members to take a “meta-perspective” on their own dialogue. Meta-discourse serves as a type of formative evaluation that can help a knowledge creating community both assess their achievement up to the current point and decide on a future plan of action. Van Aalst (2009) identifies meta-discourse as a key condition of an innovation ecology that can enable knowledge creation. Studies also show that meta-discourse can help students in a range of important ways, such as recognizing shared knowledge advances, identifying setbacks, plotting out next steps, setting goals and drawing links between them, connecting ideas, articulating new and promising questions, and establishing deeper ties between authoritative knowledge and newly identified problems (Zhang et al., 2009; Zhang & Messina 2010; Zhang et al., 2011). In this study, special “KB Talks” devoted to meta-discourse were integrated within the students’ inquiry time as a pedagogical treatment geared towards enhancing students’ knowledge building dialogue. Questions addressed in these sessions included: Are we answering our questions; are we going deeper with our theories; are we bringing in useful information that is helping us develop our ideas; are we stuck on a problem; what can we do to get “unstuck”? While both 2011 and 2012 classes engaged in collaborative discourse both on and offline, the experimental 2012 class was subject to a series of special “KB talks” that focused on engaging students in meta-discourse. Both Group A and Group B participated in a total of eight meta-discourse sessions over the course of eight months.

(ii) Concurrent, transformative and embedded assessment: This principle speaks to the effort, on behalf of the community itself, to identify advancements or setbacks in its knowledge building endeavors on a continual basis. To help facilitate meta-discourse sessions, students in the 2012 class were given formative feedback in the form of simple visualizations to help them take a “bird’s eye view” of their own discourse. Two forms of feedback were tested: a.) Word Clouds—Students in both Group A and Group B were shown a series of different word clouds that visualized key concepts and vocabulary relevant to streams of inquiry that emerged in their own discourse. The aim was to introduce and further acquaint students to new or challenging words in a context deeply integrated with their knowledge building work. In this study, three different types of word clouds were used (see Figure 1): those that depicted the most frequent terms the students were using in their naturally-occurring dialogue over time (“Our Words”); those that depicted key words that experts frequently used when talking about those same phenomena (“Expert Words”); and a third which allowed students to see the extent to

which the words characterizing their discourse mapped onto the “expert” dialogue, by means of colour-coding (“Our Shared Words”). For instance, the expert terms featured on the “Expert Words” cloud that students were engaging in their own online discourse were coloured red on this visualization, while terms that students had not yet used remained black. While visualized to the experimental class by means of the word clouds, the “expert” vocabulary was available to both classes via research materials in the classroom, including books that were read or objects that were discussed during KB talks.

All word clouds were refined throughout the inquiry, with changes based directly on terms emerging from students’ writing on Knowledge Forum. These visualizations would help students gain a sense of the semantic field of their discourse, and would enable the community to trace the use and longevity of new terms in a discourse over time. For instance, growth and change of the “Our Words” cloud helped to make explicit the attention that different terms were receiving at different points in time, displaying to the community which terms were dominating the discourse, which potentially significant terms remained underused or unrecognized, whether terms stayed relevant or useful to the problem at hand, and so on. Lack of common vocabulary between students and authoritative sources, as evidenced in the “Our Shared Words” cloud, could show limits of student understanding while also depicting terms that could help to fruitfully expand the dialogue. Embedding discussion of these visuals into meta-discourse sessions was designed to help position them as objects of public discourse that helped to make explicit important elements of the online dialogue as it emerged and to serve as artifacts the community could rally around during group reflection.



Figure 1. Visualizations supported meta-discourse, including Word Clouds (left) and Meta-Discourse tool (right)

b.) Meta-Discourse Tool—In addition to word clouds, students in Group B were exposed to the Meta-Discourse tool (see Figure 1). This is a new tool embedded within Knowledge Forum that is specifically designed to help students take a meta-level perspective on their own discourse, and to support meta-discourse by giving students explicit feedback about the contribution makeup of their group dialogue at any given time. This tool allows students to monitor the types of discursive moves—corresponding to the scaffolds in Knowledge Forum—used by their community at any given time. While both groups in the experimental class participated in meta-discourse sessions, Group B was introduced to the Meta-Discourse tool from the first treatment in order to chart the contributions on their own working view, and used the graphs produced by the tool to mediate their reflective discussions in each of the eight sessions. Both the word cloud and the meta-discourse visualizations were shown to the whole class.

(iii) *Constructive use of authoritative sources*: This principle requires that students engage with “expert” texts and information in a way that is both critical and conducive to improving their own ideas. This practice involves encountering unknown terms and concepts, and applying them to students’ own ideas. In the experimental class, students were encouraged to explore unknown words and find relevant sources to help them understand new or challenging vocabulary. After meta-discourse discussions students moved onto writing in Knowledge Forum, often forming small groups or working in pairs to find resources to help them learn more about the important terms just discussed. Students engaged in co-operative reading, writing and discussion about these words, and worked to acquire definitions of new words as well as integrate them into group discourse.

(iv) *Symmetrical knowledge advancement*: This principle implies that knowledge and expertise flows within and between community members working on shared problems in the interest of improvement of ideas. The distribution of knowledge across a community is important in the context of vocabulary learning, especially in the early years. Research shows that children who acquire literacy skills in the early years of schooling are more likely to experience success at higher levels of education, with the reverse also holding true (Stanovich, 2000). Simply put, children who know more words can learn more words (Stahl, 1991). The collaborative meta-discourse discussions, coupled with visualizations designed to give students a meta-level perspective on critical aspects of their own discourse, were aimed at engaging all students in various literacy practices including reading, speaking, listening as well as writing, so that productive vocabulary use was distributed throughout the group discourse.

It is our hypothesis that students in the experimental class will demonstrate a greater degree of productive written vocabulary than the benchmark class from the previous year. We also predict that the more expansive the vocabulary, the greater the knowledge advancement of the community. Moreover, we predict that vocabulary use in the experimental class would be used and distributed across time and groups to a greater

extent than in the benchmark class. We also hypothesize that Group B from the experimental class would contribute more diversely than Group A or the benchmark class, and correspondingly exhibit greater knowledge advancement.

Data Analysis

The data source for this study was student discourse as archived on two Knowledge Forum databases generated over two consecutive years. These include: i.) Grade 2, 2011—248 notes across four views, from both the “Bird Study” (114 notes, 3 views) and “Salmon Study” units (134 notes, 1 view); ii.) Grade 2, 2012—203 notes across eight views from their “Bird Study” (175 notes, 7 views) and “Salmon Study” (90 notes, 1 view) units; and iii.) video of student “KB talks” and meta-discourse sessions supplement notes and provide qualitative information about students’ ideas.

The application of behavioural, lexical, and group-level dynamics, are summarized as follows:

(a) Behavioural Measures: The Knowledge Forum Analytic Toolkit (Burtis, 1998) was used to calculate the number of notes authored per student and the percentage of notes read per student.

(b) Lexical Measures: Lexical profiles were calculated for each student using the Knowledge Forum Analytic Toolkit. Researchers manually corrected spelling errors so that all words could be picked up by the automated tools. Three attributes were used to create students’ lexical profiles, and include the following: i.) academic words; ii.) 1st, 1000 words; iii.) domain-specific words. The Academic Word List (AWL) is composed of 570 written families external to the 2000 most frequently used English words but common in academic discourse. The 1st, 1000 words refers to a lexicon consisting of the most frequently used words in English, plus their grammatical variations. Greater use of high frequency words is indicative of a more limited vocabulary (Nation, 2001). With respect to domain-specific words, two inquiries were conducted to generate a single word list. Firstly, researchers consulted the Ontario Curriculum Standards document for Science and Technology and identified key words corresponding to the “Understanding Life Systems” stream. The words selected totaled 342 individual terms that ranged across Grades 1-10. Words selected from the curriculum document were divided into two levels according to the grade in which they appeared in the curriculum document. 84 words were identified at or below the Grade 2 level, and 258 words above the Grade 2 level. In addition to this, the author and classroom teacher consulted the external sources available in the classroom and identified terms critical to particular streams of inquiry as they emerged during the course of knowledge building work. These words appeared on the word cloud visualizations to help students expand their vocabulary repertoire. For analysis, a total of 64 “expert” words were combined with the 342 curriculum words to create a single comprehensive list. This cumulative list, which totaled 406 words, plus their grammatical variations, was used to measure domain-specific vocabulary.

(c) Depth of Understanding: To examine community knowledge advancement, two researchers used content-based analysis to select notes from the online discourse that represented “theorizing” work (see Chuy, Resendes, Tarchi, Chen, & Scaradmalia, 2011). Such notes exhibit students’ explicit attempt to produce explanations and express original ideas, and as such comprise useful examples of students’ productive writing and their ability to convey conceptual understanding. To evaluate depth of understanding, “theorizing” notes were then subject to further analysis according to two coding schemas developed by Zhang and colleagues (2007) to measure “scientificness” and “epistemic complexity” of ideas, each possessing four levels. Scientificness implies the degree to which an idea is scientifically accurate, while epistemic complexity represents the level of cognitive effort and written sophistication evident in an explanation. The level of idea complexity informs the meaning for scientificness, so scores for each note were multiplied for a single composite value (Zhang & Sun, 2008).

(d) Group Discourse Network Structure: On a group level, notes were analyzed using KBDex (Matsuzawa, Oshima, Oshima, Niihara, & Sakai, 2011), a tool developed for Knowledge Forum that is designed specifically to analyze the network structure of collective discourse based on co-occurrence of words. KBDex can reveal the network structure of a community’s discourse according to three levels. Firstly, it analyzes interconnectivity between students via shared vocabulary on a social level; secondly, it maps discursive connections on the level of individual notes, which shows the use and distribution of vocabulary in students’ writing; lastly, it traces connections at the level of individual words, which reveals semantic relationships between words and the conceptual content of the discourse. For this study, we analyzed each class’s discursive network on the social and individual word levels according to Degree Centrality (DC), Betweenness Centrality (BC) and Closeness Centrality (CC), which represent standard points of analysis in complex network science (e.g., Newman, 2010). Degree centrality measures the “popularity” or number of connections one node has with other nodes in the network. In this case, each network node represents a student or a word, with connections between students created through the use of the same word, and connections between words created when one word appears in the same written note as another word. So, the more discursive connections a student has with other students, or a word with other words, the more “popular” or centralized that student or word is in the network. Betweenness centrality provides a valuable measure at both a local and global level, and indicates the degree of connectivity of a node, as well as the “load” placed on the node by all other nodes. For this research, this measure reveals the

extent to which students or words are connected within a community and the degree to which they bridge various social clusters or discursive cliques, respectively. Closeness centrality measures the proximity of one node to all other nodes, and is indicative of how quickly information can flow through a network. Applied to this case, this measure reveals how closely connected students are to each other via the discourse they are engaging in, or, in the case of words, the semantic context in which they are being used. The particular domain-specific and academic words used by the students in each class, generated from their lexical profiles, were used to comprise two separate word lists for group analysis in KBDeX. In this way, the discursive relationships between students and words characterizing the collective discourse could be mapped.

Results

Did the experimental class show more productive written vocabulary than the “benchmark” class?

To explore significant differences across groups in student performance on behavioral and lexical measures, as well as on their demonstrated depth of understanding, a one-way ANOVA was conducted for each measure. Results show significant differences for the following measures: total domain words, $F(2, 43) = 7.77, p < .01$; unique domain words $F(2, 43) = 5.62, p < .001$; total words $F(2, 43) = 3.44, p < .05$; use of words above Grade 2, $F(2, 43) = 7.24, p < .001$; and depth of understanding $F(2, 43) = 11.5, p < .001$. Post-hoc tests (HSD) revealed that Group B used significantly more domain words in total ($p < .001$, Cohen's $d = 9.72$), as well as more unique domain words ($p < .01$, Cohen's $d = 5.91$) than the other two groups. Furthermore, Group B wrote significantly more words than Group A ($p < .05$, Cohen's $d = -72.64$), and both Group A ($p < .01$, Cohen's $d = 2.96$) and Group B ($p < .05$, Cohen's $d = 2.35$) outperformed the 2011 class with respect to use of words above Grade 2. These results suggest that formative feedback that is embedded in knowledge building practice helps young students to use increasingly rich and diverse vocabulary. These findings also suggest that visualizations reflecting student contribution patterns to group discourse prompt students to write more in total.

With respect to depth of understanding, both Group A ($p < .001$, Cohen's $d = 2.45$) and Group B ($p < .01$, Cohen's $d = 1.79$) performed better on depth of understanding than the 2011 class. This suggests that formative feedback coupled with collaborative reflective discussion can help students construct and communicate ideas in writing that reflect greater scientific accuracy and more elaborate explanations. Within the 2012 class, Group B wrote significantly more words than Group A, as noted, yet there was no significant difference in depth of understanding between groups ($M = 5.13, SD = 2.15$ vs. $M = 4.48, SD = 1.12$, respectively). A closer look at student work reveals that within Group A, a few students stood out as having especially high composite scores, which could help to explain group performance on this measure.

To what extent was key vocabulary distributed in the shared discourse?

The continual give and take of ideas to advance community knowledge is a foundational principle upon which knowledge building communities operate. In order to explore group-level dynamics in the community and the shared discourse, network structure analysis was conducted using KBDeX. As elaborated in a previous section, typical knowledge building sessions in both Grade 2 classes involved students splitting up into rotating groups. However, all students in both classes worked in the same knowledge space on the database and contributed their ideas to a shared online discourse. For this reason, group-level analysis was conducted across the 2011 and the 2012 class as a whole with this tool.

To explore any significant differences across classes with respect to the degree, betweenness and closeness centrality of each student community, a one-way ANOVA was conducted for each measure. A significant difference was found for both degree centrality $F(2, 41) = 11.17, p < .0001$ and betweenness centrality $F(2, 41) = 13.46, p < .001$. Post-hoc tests showed that both Groups A ($p < .01$, Cohen's $d = .17$) and B ($p < .001$, Cohen's $d = .19$) displayed greater degree centrality than the 2011 class. Furthermore, the benchmark class showed greater betweenness centrality than both Group A ($p < .001$, Cohen's $d = .0091$) and Group B ($p < .001$, Cohen's $d = .0090$). No difference was found with respect to closeness centrality. These findings indicate that the 2012 class community had a higher number of students who had a higher number of connections with other students, suggesting that a greater number of students were using more shared words more often. Moreover, a higher betweenness centrality indicates that there were more social clusters in the 2011 class, as opposed to the 2012 class in which each student was more highly connected to every other student. However, that the 2011 class also exhibits a highly connected community, albeit to a lesser degree than the 2012 class, indicates that knowledge building practices are themselves conducive to knowledge distribution across the community. In general, these results suggest that as students learn and use a more diverse range of words in the context of knowledge building, the more discursively connected they become, and the greater the vocabulary knowledge distribution across the community.

The same series of tests were conducted to trace differences in the network structure of individual words. A significant difference across groups was found for closeness centrality, $F(2, 43) = 3.34, p < .05$, with

the students in Group B using terms that were more semantically bound together in their discourse than in either of the other groups. This group of students exhibited more diverse vocabulary, as their discourse contained more terms semantically-related to those of other students. This condition is conducive to vocabulary learning as students have access to a wider range of vocabulary in a collective knowledge pool, and this appears to foster higher levels of word-associations leading to use of these terms in different contexts. In terms of degree centrality and betweenness centrality, no significant differences were found. This suggests that engagement in knowledge building practices encourages active use of important vocabulary in writing and making connections across various discursive streams in collective dialogue.

Discussion and Conclusions

This study explored the impact of formative feedback visualizations embedded within knowledge building practices to students' productive written vocabulary. Student work was analyzed on two levels. First, vocabulary use was assessed by calculating behavioral and lexical attributes. Content analysis was used to evaluate depth of understanding, as demonstrated in student online discourse. Second, the distribution of vocabulary use throughout the community was explored through network structure analysis of discourse on the level of students and words.

Results show that formative feedback that is productively integrated into authentic inquiry practices can facilitate vocabulary growth, use of new words in students' writing, and advances in community knowledge. On the whole, students in the experimental class used more domain-specific vocabulary more often and exhibited greater scientificness and complexity of ideas than students in the benchmark class. Within the experimental class, students who received formative feedback related to both vocabulary use as well as feedback regarding the various ways they were contributing to group dialogue, used more sophisticated words than students who only received feedback regarding vocabulary, but did not show greater knowledge advancement. This suggests that the vocabulary use for students receiving both kinds of feedback extended more widely beyond their theorizing work and into different contribution types, such as asking questions or reporting facts. It also suggests that engaging students in rich reflective discussion around formative feedback has a positive effect on students' knowledge advancement. Based on these findings, one possible recommendation for primary grade teachers is to encourage group reflection consistently throughout a knowledge building study, since meta-discourse sessions proved fruitful even for students as young as Grade 2. Another recommendation is that teachers take advantage of group discussion periods to integrate feedback visuals for students to collaboratively explore.

Finally, network structure analysis of students' collective discourse showed that all students in the experimental class were more discursively connected to one another and made more connections with other students via their shared discourse than students in the benchmark class. That this distribution of vocabulary knowledge is evident in Grade 2 is promising given that the disparity between students who demonstrate high literacy skills and those who show lower level skills accelerates notably after the primary level and into the junior grades. Results also suggest the importance of supporting meta-discourse, enhanced by formative feedback, as a routine component of knowledge building practice with young students.

Further research that explores students' verbal dialogue in addition to the content of their online contributions is needed to more fully explore and assess primary aged students' literacy levels and capacities for expanding vocabulary knowledge. Also, to better understand the impact of formative feedback on developing students' capacities in literacy concurrent with knowledge building, future research will focus on refining feedback designs and examining a wider range of literacy and knowledge distribution indicators.

References

- Applebee, A. N. (1981). Writing in the secondary school. Urbana, IL: *National Council of Teachers of English*.
- Bateman, S., Gutwin, C., & Nacenta, M. (2008). *Seeing things in the clouds: The effect of visual features on tag cloud selections*. Proceedings of HT '08, June 19-21, 2008. Pittsburgh, PA.
- Bereiter, C., & Scardamalia, M. (1987). An attainable vision of high literacy: Approaches to teaching higher-order skills in reading and writing. *Curriculum Inquiry*, 17(1), 9-30.
- Biemiller, A. "Teaching Vocabulary: Early, Direct, and Sequential." *American Educator*, Spring (2001).
- Black P. and D. Wiliam (1998), "Assessment and Classroom Learning", *Assessment in Education: Principles, Policy and Practice*, CARFAX, Oxfordshire, 5(1), 7-74.
- Burtis, J. (1998). *Analytic Toolkit report*. Toronto, ON: Ontario Institute for Studies in Education of the University of Toronto.
- Chall, J. S. (1996). *Stages of reading development* (2nd ed.). Fort Worth, TX: Harcourt Brace.
- Chernobilsky, E., DaCosta, M. C., & Hmelo-Silver, C. E. (2004). Learning to talk the educational psychology talk through a problem-based course. *Instructional Science*, 32(4), 319-356.
- Chuy, M., Resendes, M., Tarchi, C., Chen, B., & Scardamalia, M., (2011). Ways of contributing to an

- explanation-seeking dialogue in science and history. *QWERTY - Interdisciplinary Journal of Technology, Culture, and Education*.
- Corson, D. J. (1997). The learning and use of academic English words. *Language Learning*, 47, 671–718.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213-238.
- Hearst, M. & Rosner, D. (2008). *Tag clouds: Data analysis tool or social signaller?* Proceedings of the 41st Hawaii International Conference on System Sciences. Kauai, Hawaii.
- Herman, P.A., Anderson, R.C., Pearson, P.D., & Nagy, W.E. (1987). Incidental acquisition of word meaning from expositions with varied text features. *Reading Research Quarterly*, 22 (3), 263-284.
- Laufer, B. (1994). Lexical profile of second language writing: Does it change over time? *RELC Journal*, 25, 21-33.
- Marzano, R. J. (2006). *Classroom assessments and grading that work*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Matsuzawa, Y., Oshima, J., Oshima, R., Niihara, Y., & Sakai, S. (2011). KBDeX: A platform for exploring discourse in collaborative learning. *Social and Behavioral Sciences*, 26, 198-207.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Newman, M. (2010). *Networks: An introduction*. Oxford University Press, USA.
- Pelletier, J., Reeve, R. & Halewood, C. Young children's knowledge building and literacy development through Knowledge Forum®, *Early Education & Development*, 17(3), 323-346.
- Scardamalia, M. (2004). CSILE/Knowledge Forum. In *Education and technology: An Encyclopedia* (pp.183-192). Santa Barbara: ABC-CLIO.
- Scardamalia, M. (2003). Crossing the digital divide: Literacy as by-product of knowledge building. *Journal of Distance Education*, 17(3), 78–81. Learning Technology Innovation in Canada.
- Scardamalia, M., & Bereiter, C. (2003). *Knowledge building*. *Encyclopedia of Education*, (2), 1370-1373.
- Scardamalia, M., & Bereiter, C., Brett, C., Burtis, P.J., Calhoun, C., & Smith Lea, N. (1992). Educational applications of a networked communal database. *Interactive Learning Environments*, 2(1), 45-71.
- Stahl, S.A. (1991). Beyond the instrumentalist hypothesis: Some relationships between word meanings and comprehension. In P. Schwanenflugel (ed.), *The psychology of word meanings*, pp. 157-178. Hillsdale, N.H.: Lawrence Erlbaum Associates.
- Stahl, S. A. & Fairbanks, M.M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, 56, 72-111.
- Stanovich, K.E. (2000). *Progress in Understanding Reading: Scientific Foundations and New Frontiers*. New York: Guilford Press.
- Steahr, Lars S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. December. *Studies in Second Language Acquisition*, 31(4), 577-607.
- Stiggins, R. (2004). *Student-involved assessment for learning* (4th ed). Upper Saddle River, NJ: Prentice-Hall Inc.
- Strogatz, S. (2001). *Exploring complex network*. *Nature*, 410, 445-489.
- Sun, Y., Zhang, J., & Scardamalia, M. (2010). Knowledge building and vocabulary growth over two years, Grades 3 and 4. *Instructional Science*, 38(2), 147-171.
- Van Aalst, J. (2009). Distinguishing between knowledge sharing, knowledge construction, and knowledge creation discourses. 4(3) *International Journal of Computer-Supported Collaborative Learning*, 259-287.
- Yonek, L. (2008). The effects of rich vocabulary instruction on students' expository writing. (Doctoral dissertation). University of Pittsburgh, Pittsburgh, PA.
- Zhang, J., & Sun, Y. (2008) *Knowledge building measures that matter*. Paper presented at the Knowledge Building Summer Institute, University of Toronto, Canada.
- 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.

Acknowledgments

This research was funded by a grant from the Social Sciences and Humanities Research Council of Canada titled “Ways of contributing to dialogue in elementary school science and history.” We would also like to thank the students, teachers and principals of the Dr. Eric Jackman Institute of Child Study, University of Toronto.

Youth Roles and Leadership in an Online Creative Community

Ricarose Roque, Natalie Rusk, Amos Blanton

MIT Media Lab

75 Amherst St

Cambridge, MA

ria@media.mit.edu, nrusk@media.mit.edu, amos@scratch.mit.edu

Abstract: Teens within local community organizations often serve in leadership roles, such as camp counselors or program assistants. As they carry out their responsibilities, they gain work skills and the community benefits from their contributions. With young people spending more time online, how might they build similar skills while contributing to the online communities they engage in? In this paper we examine the experience of youth who have taken on leadership roles within Scratch, a creative online community. We identify the main challenges these youth encountered, the strategies they used to manage these challenges, and what they learned in the process. Their descriptions suggest a progression from learning to carry out their responsibilities in collaboration with other team members to eventually developing their own visions for improving the community. We have found that these roles provide pathways of participation and deeper engagement for youth interested in contributing to online communities.

Introduction

Scratch is a programming language that enables children to create interactive media, such as animations, games, and stories. Scratch is also an online community, where young people share their creations and connect with other members online (Resnick et al., 2009). Inspired by Seymour Papert's idea of a "computational samba school" (Papert, 1980; Zagal and Bruckman, 2005), the Scratch online community is designed to engage members of all levels of expertise, ages, and backgrounds in learning from one another as they create and play together. In Scratch, members explore others' projects, write comments, and download and remix each other's projects. The Scratch website also includes a discussion forum, where members can ask questions, converse about Scratch-related topics, and find collaborators. Since it launched in 2007, Scratch (<http://scratch.mit.edu>) has grown into a dynamic community with more than a million registered members, primarily between the ages of 8 and 16, and over 2.8 million shared projects.

As participation in the Scratch website grew, young people contributed in ways beyond what we had originally anticipated. More and more young people took the initiative to help others—from answering questions in the website discussion forums to creating interactive tutorials to share their skills. Youth also began to look for ways to help the Scratch Team, which we are members of, including reporting issues and suggesting ideas for improving the programming language and the website. The Scratch Team saw these emergent activities as opportunities to engage members in helping create a supportive environment for the entire community. We and other team members have incorporated youth's ideas in a number of ways. In this paper, we focus on one of our initiatives: creating explicit roles for youth to volunteer their time to help out in the Scratch community.

Youth within local community-based organizations often take on leadership roles, for example, serving as camp counselors, program assistants, and peer tutors (Roth & Brooks-Gunn, 2003). As they carry out their roles and responsibilities with support from adult staff, they learn to handle challenges and develop planning, problem solving, communication, teamwork, and other work and life skills (Salusky et al., 2012). In addition to youth gaining valuable experience, the community organizations also benefit from the new ideas, perspective, and energy of youth leaders (Pittman, Irby, Tolman, Yohalem, & Ferber, 2003).

As young people are increasingly spending time online (Lenhart, Purcell, Smith, & Zickuhr, 2011), how might youth—and the online communities they participate in—gain similar benefits in roles online? In this paper, we explore this question by looking at the experience of youth who have taken on leadership roles within the Scratch online community. We first describe the roles of *community moderators* and *collab counselors* and introduce the youth that we interviewed about their experience. We then examine the challenges these youth leaders encountered, the strategies they used to overcome them, and what they learned in the process. Finally, we reflect on the opportunities for learners and designers in providing leadership roles for youth in online communities.

Community Roles for Youth

In this paper, we focus on two key roles that youth have served within the Scratch online community: community moderators and collab counselors. We also refer to youth in these roles as "youth leaders," as they

are developing leadership skills and fulfilling authentic responsibilities in collaboration with the Scratch Team adult facilitators.

Community moderators assist in the management of the Scratch community's activities, particularly in the website's active discussion forums, which receive about 1,100 posts a day. They help answer questions, provide constructive feedback, keep discussions friendly and on-topic, and model respectful interaction in the community. They have access to a moderators' forum, where they can discuss community-related issues with the Scratch Team. The first group of community moderators were hand-picked by the Scratch Team in 2008. The selection process was shifted into an election model in 2010. Scratch members who are interested in becoming community moderators can nominate themselves and describe their interest in the role. The Scratch Team then selects a subset of nominees for community members to vote on. Since the community moderator role was created, there have been 14 moderators from 6 countries.

Collab counselors support the community in a series of online collaboration events called Collab Camps on the Scratch website. Collab Camp is a community-wide event in which participants form collaborative groups, called "collabs," to create a Scratch project. Prior to the start of each Collab Camp, we invite community members who have demonstrated the ability to give constructive and helpful feedback to become collab counselors. A collab counselor's primary responsibility is to give constructive feedback on projects created by participants of the Collab Camp. Like community moderators, collab counselors have a private discussion forum to ask questions and discuss strategies for giving constructive feedback. Since we created this role in 2011, there have been 9 counselors from 5 countries.

The Scratch Team, based in the MIT Media Lab, develops and manages the Scratch programming environment and online community. We, the authors of this paper, are part of the Scratch Team. As adult facilitators, we regularly interact with moderators and counselors, answer questions, select or encourage members to take on these roles, and discuss the latest trends and issues in the community.

We discuss youth participation in Scratch from the perspective of youth development literature, which focuses on children and adolescents developing a broad range of skills through active participation in programs, typically in out of school time (Eccles & Gootman, 2002). Some youth development studies focus specifically on adolescents' development of leadership skills as they carry out responsibilities, such as leading activities for younger children or organizing community action projects (e.g., Conner & Strobel, 2007). To date, only a few studies have examined youth leadership in online environments, and those have focused on youth working with each other as part of a group or team (e.g., Cassell, Huffaker, Tversky, & Ferriman, 2006; Turkaya & Tirthalia, 2010). The current paper focuses on youth carrying out roles with responsibilities within a broader online community, in which they help to manage online activities for people of diverse ages and backgrounds.

Studying the Experiences of Youth Leaders

To understand the experiences of the youth in these roles, we collected observations and online activity data from the Scratch website, such as their comments, forum posts, projects, and self-reported age and gender. These observations and online activity enable us to see how they participated in the community. To understand more deeply how they saw their experiences, we also conducted semi-structured interviews. We posted a message in the moderator and counselor forums inviting them to be interviewed about their experiences in their roles. Three moderators and one counselor responded to our message. Table 1 provides brief descriptions of each youth leader. In these interviews, we asked them questions (e.g., "What has been challenging?", "Why do you continue to take on this responsibility?") to surface their challenges and the lessons they learned to overcome them.

Table 1: Brief portraits of the youth leaders

<i>Jacob</i> , a 15-year-old from Belgium, was one of the first community moderators to be invited by the Scratch Team. He started using Scratch when it first came out in 2007 and began making creative and sophisticated games that became well-known in the community.
<i>Fayth</i> , a 22-year old from the United States, was invited to be a moderator by the Scratch Team. Fayth is the oldest of the moderators and counselors. She discovered Scratch in a college course and became an active member of the community, creating many art and animation projects.
<i>Sam</i> , a 14-year-old from Canada, was chosen as a community moderator in the first community-wide moderator election. Sam began using Scratch in 2008, and has shared many game projects on the website.
<i>Jessica</i> , a 17-year-old from the United States, was one of the first collab counselors to be invited by the Scratch Team. She started using Scratch in 2009 and has created many simulations and interactive projects that express her love of math.

We used grounded theory strategies (Charmaz, 2006) to analyze all our data from interviews, observations, and documentation. Our coding and analysis led to the identification of the key challenges and strategies that the youth used to fulfill their responsibilities.

Learning from Challenges

The youth interviews suggested four types of challenges that youth leaders encountered: (a) learning to carry out their role as part of a team; (b) managing their new identity within the role; (c) interacting constructively in an online medium with limited context; (d) and facing broader challenges within the community. In this section, we examine each of these challenges, the strategies youth used to manage the challenges, the skills they developed in the process, and how these experiences connect with related studies in the youth development literature.

Fulfilling Responsibilities Through Collaboration

One of the first steps in becoming a moderator or counselor is learning what one's responsibilities are and how to fulfill them. Sam described how he needed to read several pages of guidelines on how to be a community moderator before getting started—and then additional instructions on the steps for carrying out each of the responsibilities.

I found that I had several pages of moderator guidelines before me before I was able to do anything. After that I found that I had even more to read on how to actually do what I needed to do, such as close topics or respond to reports.

All four youth we interviewed remarked at how valuable their fellow moderators and counselors as well as the Scratch Team were in learning how to fulfill their responsibilities. Moderators and counselors worked together closely, using their respective discussion forums to communicate and coordinate around their shared responsibilities. Jacob summarized the overall workflow of moderators as a collaborative process:

We have the moderators forum where you post a topic about this conversation. We use it a lot—whenever there is something going on and we see what people have to think about it. It's an important source of information. Because when you see how others do it, you can change yourself to match it. . . . We basically do everything together as a group. And when someone does something wrong, it's corrected by others. It's a continuous collaboration, moderation.

Learning from each other was especially emphasized by moderators, who described not only learning new forum moderation tools, but also learning how to answer questions, respond to reported issues, and to maintain a friendly and respectful tone in the discussion forums. When new moderators come into the role, they are encouraged to ask many questions and observe other moderators. For example, Sam explained that he learned “right away to ask for help rather than guess what to do” whenever he was unsure how to respond to an issue or question that arose.

Learning from each other went in both directions, as experienced moderators and the Scratch Team learned from new moderators. Jacob described how new moderators come in with “fresh perspectives” that they can share with the rest of the team. “They have been on that side and now they are moderators and they share their opinions and it makes us think more like the community thinks.” In moderator forum discussions, the Scratch Team and older moderators explicitly asked new moderators to participate and share their opinions.

Based on the descriptions of their experiences, the nature of cooperation in the moderator group resembles a community of practice. Members learn through peripheral as well as productive actions (Lave & Wenger, 1991). For example, through participating in discussions in the moderation team, Sam learned how the team thinks through decisions very carefully, which he says he had not realized before as a regular member of the community. The experience of working together on a team not only help youth learn to carry out their responsibilities, but also how to collaborate with other team members to accomplish shared goals. This ability to collaborate effectively on a team is recognized as a fundamental skill for the development of youth leadership (Conner & Strobel, 2007). For Scratch moderators and counselors, this experience extended to learning to work together on a team not only with other youth but also with adults.

Managing New Identities

When youth took on a role within the Scratch community, the title (e.g., “Collab Counselor”) appeared under their name in the website discussion forums. As they began to interact with their peers in the rest of the community, moderators and counselors noticed how their peers perceived and interacted with them differently.

Jessica described how Scratch members, especially the newer and younger ones, saw her as an authority figure. While Jessica understands that she had a special role in the community, she wanted members to see her as a “big kid,” someone who is helpful and experienced but not an authority figure:

I don't really want them to think I have more authority, necessarily. I like [to convey] the feeling, especially [to] younger Scratchers . . . of sorta being a “big kid” and helping them along but not necessarily being in charge. . . . I try to construct my comments as suggestions and not “now go do this I'm in charge” because of I'm not. [Laughs.] I don't know, I guess that's the main thing, especially the younger Scratchers thought I had more authority than I did when in fact I was just trying to help them.

She carefully wrote her comments as suggestions, making sure her comments were perceived as suggestions from a peer, rather than an authority figure. Similarly, Jacob tried to act in the community as he would normally act, regardless of being a moderator. He did not want to seem “distant and unreachable and important.” He saw his role as nothing more than extra tools to do what he would normally have done in the community.

Despite the unwanted perceptions, they also recognized the value in having a formal title. Jessica found it useful to introduce herself as a “collab counselor” to members, especially to members she had never interacted with before, so they could understand why she was giving them detailed and constructive feedback. When intervening in a conflict between two Scratch members, Jacob found his title useful to explain why he became involved.

While youth leaders in face-to-face contexts also deal with perceptions of authority (e.g., Dean, 2010), moderators and counselors must negotiate this perception in an online medium. For example, signals taken for granted in person, such as facial expressions and physical appearances, are not apparent, especially in a website where members are anonymous. To protect children's privacy on the Scratch website, we do not display their age or gender, and discourage members from sharing personal information. Youth leaders must also learn to manage their identity in a community where membership is ephemeral: new members are constantly joining, and old members are leaving. To overcome these challenges, moderators and counselors actively worked to make sure that they were seen as someone who is accessible and helpful, rather than distant and important. In addition, they learned to leverage these perceptions to help them fulfill their responsibilities effectively. Learning to negotiate one's identity when taking on new roles is seen as a positive developmental outcome of community-based youth programs (Polman & Miller, 2009), and the descriptions of the youth leaders in the Scratch community suggest that they learned to manage their new identity as they interacted with their peers.

Using Empathy and Interpreting Signals

Communicating within the text-based medium of Scratch comments and discussion forums can be challenging, as members usually have never met in person and what they say can, at times, be ambiguous. Jacob talked about sarcasm online and the difficulty of detecting it. For example, when a member writes, “That's the best project ever,” they may mean one of two things: either they are genuinely expressing praise or are sarcastically implying that it is not a good project.

When responding to posts like these, moderators often tried to take the perspectives of the individuals involved. For example, Fayth encountered a situation where one Scratch member posted an Internet joke that was funny to some, but was offensive to others. When handling a tricky situation like this, Fayth would ask herself to think about the people she is interacting with:

With the Scratch community, we have to deal with people of almost every age. From what I heard, the youngest users are around 5 years old, and some of the oldest, they're like grandparents already, and then there's everyone else in between. And so, I guess it's kind of like, when moderating the Scratch community, you have to think about, “Well, who am I talking to? Am I talking to them in the right way so that they can understand me?”

Fayth tried to empathize with the people involved, to understand why they might say or do the things they've done. She then took a step back and tried to “clarify everything to come to some sort of solution that is beneficial to everyone.”

For counselors who are responsible for giving feedback on projects, sometimes the lack of context made it difficult to understand who the creators are and what they needed most from feedback. Through her work as a counselor, Jessica interacted with diverse types of projects and creators. She said her biggest challenge was to respond to genres of projects she was less familiar with, especially game-related projects. Because most of her projects have been simulations and math-related projects, at first she felt that she may not be able to contribute any valuable feedback, especially if it was an advanced game project.

To better understand what project creators might need, Jessica looked to other cues such as what the creators write in their projects notes and other projects they've created. As she described:

Simply reading the project notes gives it away. Often Scratchers will include hints about what they're particularly proud of: "We spent hours perfecting the timing!" . . . Of course, glancing at some of their other projects might provide a hint too.

Jessica used these new cues to understand what project creators cared most about in their project. Rather than writing a comment about what she would do and what she cares about, Jessica crafts her comments so that they are most relevant to the creator and their project vision.

The ability to give others' constructive feedback is a key skill that youth in local community-based programs say they learn through their participation (Dworkin, Larson, & Hansen, 2003). The youth took on the challenge of communicating effectively and constructively in this medium by striving to consider as much as possible about the individual members in their responses, including reading their project notes to understand their intentions and viewing other projects they had made. The ability to consider others' perspectives and communicate with consideration for differences are considered core skills for adolescents to develop (Lerner, 2009), and their roles provided opportunities for the youth moderators and counselors to practice and refine these skills.

Developing a Vision of Community

When young Scratch members took on these roles, they transitioned from being a creator—designing and sharing projects on the Scratch website—to a role where they learned new responsibilities and experienced new ways to interact with the entire community. Through these experiences, they became aware of community-wide dynamics and challenges they had not known of before. For example, Jessica saw more clearly the challenges newcomers encounter on Scratch and became more sympathetic to them over time.

People want to work with [expert Scratch members] . . . so they may not give new people as much of a chance in their groups. I noticed a lot of [new Scratch members] never managed to get a group together. Probably [new Scratch members] didn't really know many people from the Scratch website. So that made it hard for them.

When discussing the community, all four youth leaders described what they wished for the community and what they wanted for their peers to experience. Fayth described how she wanted to help others in the community have the kinds of positive experiences she had. Similarly, Sam believed that the community should be "welcoming to everyone and really nice" and "appropriate for all ages."

These visions for community translated to how they wanted to improve the community. As a moderator, Sam found that he spent a significant amount of his time responding to reported issues and other forum maintenance tasks, such as moving posts to relevant topics and closing threads. However, rather than only reacting to and cleaning up content, Sam wanted the moderators to also focus on initiating positive interactions:

One of the things that we don't do enough is put positive things in the community. We take out the negative things, but we don't put positive things as much as we should. And I think that's something that should be changed.

As Sam became more experienced as moderator, he also began working with the Scratch Team to conceptualize a "welcoming committee" in Scratch, a group of Scratch members who are interested in helping newcomers get started. As of this writing, this committee has been implemented on the website, with newcomers being greeted by youth volunteers. Jessica is now helping the Scratch Team to manage the growth of the committee.

When our youth leaders first began, they operated under the expectations of their roles and worked hard to fulfill them. However, as they began to interact with their peers and actually act on their responsibilities, they started to develop a vision for the community that, at times, extended beyond the expectations and boundaries of their role. They became motivated to do more to support their peers in having positive and constructive experiences in the community. This process of youth leaders envisioning and contributing to improvements fits with the ideal of youth development programs in which youth not only develop leadership skills, but also contribute new ideas, perspectives, and energy to address problems within the community (MacNeil, 2006; Camino & Zeldin, 2002).

Discussion

In this paper, we examined how young people developed leadership skills within these roles online. In their experiences, we see a progression in their development. When they first came into their roles, they needed to learn their various responsibilities, looking to each other to learn how to fulfill them. Even after they learned the ropes, they continued to learn by collaboratively working with other moderators and the Scratch Team. By taking on a role and title, they were faced with the task of learning to negotiate their new identity within the Scratch community, managing perceptions of being an authority figure, but using the recognition of their role and title to carry out their responsibilities. They were also learning to interact with diverse members of the community, which developed into new connections, but also new challenges to understand people who had different interests and backgrounds. And, as they became more aware of the dynamics and challenges across the community, their participation transformed. From fulfilling responsibilities outlined by the Scratch Team to developing their own visions for themselves and for the community, these youth leaders saw new ways to create and sustain a supportive community.

Embedded within the youth's experiences are structures we designed to support them—structures that we continually redesigned with input from our youth leaders. Moderators and counselors discussed and coordinated within separate website discussion forums, which became valuable spaces for them to learn from one another. We developed “guidelines” which describe their responsibilities and suggest ways for them to interact with other community members. We iterated on these guidelines together with the youth leaders, especially as they developed new strategies to carry out their responsibilities of moderation or giving constructive feedback. And as they developed new ideas to support the community, we worked together with youth leaders to design new structures, such as the “welcoming committee” to greet newcomers.

The youth's contributions through these roles have benefited the entire community. While the Scratch Team accepts overall responsibility for the website, we depend on discussions with youth moderators and counselors as a way to collectively think through choices related to the community and the design of the website. These interactions between the Scratch Team adult facilitators and the youth have developed into partnerships, where we work together to achieve shared goals. The youth leaders have provided valuable insights into the community—insights they have gained from their authentic participation. These insights have influenced our views of the community and have helped us better maintain the community. Finally, the community not only benefits from the contributions of their role (e.g. keeping forum discussions friendly and giving constructive feedback), but these youth illustrated to other members how these roles can be pathways of participation.

The experiences of these youth highlight the opportunities that youth leadership roles in online settings can provide for the youth, their adult partners, and the communities they participate in. While many youth in the community already help out independent of these roles, these roles created explicit and visible pathways of participation. Such visibility can be valuable for youth who may be interested in contributing, but may not stumble into these kinds of activities on their own. This visibility is especially important in open and large online communities, where most actions—while public and persistent online—may be buried in the rapidly changing and increasing activity of the community. From these roles, we saw how youth expanded their vision and crafted new ideas for what was possible in the community. These youth envision a community where they and other young people can create and share their projects in a supportive and safe environment. Designers of online communities can create these opportunities for their members and for their community through these youth leadership roles. Such participation can foster an environment where members are actively taking ownership of the community and giving back in multiple ways. As Jessica, the collab counselor, so aptly stated, “When everyone helps a little bit, we all benefit.”

References

- Camino, L. & Zeldin, S. (2002). From periphery to center: Pathways for youth civic engagement in the day-to-day life of communities. *Applied Developmental Science, 6*(4), 213–220.
- Cassell, J., Huffaker, D., Tversky, D., & Ferriman, K. (2006). The language of online leadership: Gender and youth engagement on the Internet. *Developmental Psychology, 42*, 436-449.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London, UK: Sage Publications.
- Conner, J.O. & Strobel, K. (2007). Leadership development: An examination of individual and programmatic growth. *Journal of Adolescent Research, 22*, 275-297.
- Dean, C. (2010). *The ecology of peer tutoring: Perspectives of student staff in one high school writing center* (Doctoral dissertation). Retrieved from <http://digitalcommons.library.umaine.edu/etd/488/>
- Dworkin, J.B., Larson, R., & Hansen, D. (2003). Adolescents' accounts of growth experiences in youth activities. *Journal of Youth and Adolescence, 32*, 17–26.
- Eccles, J. S., & Gootman, J. A. (Eds.). (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.

- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation (1st ed.)*. Cambridge University Press.
- Lenhart, A., Purcell, K., Smith, A., & Zickuhr, K. (2010). *Social media & mobile Internet use among teens and young adults*. Pew Internet & American Life Project. Retrieved from <http://pewinternet.org/Reports/2010/Social-Media-and-Young-Adults.aspx>.
- Lerner, R. M. (2009). The positive youth development perspective: Theoretical and empirical bases of a strengths-based approach to adolescent development. In S. J. Lopez & C. R. Snyder (Eds.), *Oxford handbook of positive psychology*. (2nd ed., pp. 149–163). New York: Oxford University Press.
- MacNeil, C. A., (2006). Bridging generations: Applying “adult” leadership theories to youth leadership development. *New Directions for Youth Development*, 109, 27-43.
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books, Inc.
- Pittman, K., Irby, M., Tolman, J., Yohalem, N., & Ferber, T. (2003). *Preventing problems, promoting development, encouraging engagement: Competing priorities or inseparable goals?* Washington, DC: The Forum for Youth Investment.
- 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, 879–918.
- Resnick, M., Maloney, J., Monroy-Hernandez, A., Rusk, N., Eastmond, E., Brennan, K., et al. (2009). Scratch: Programming for all. *Communications of the ACM*, 52, 60-67.
- Roth, J. L., & Brooks-Gunn, J. (2003). What is a youth development program? Identification and defining principles. In F. Jacobs, D. Wertlieb, & R. M. Lerner (Vol. Eds.), *Enhancing the life chances of youth and families: Public service systems and public policy perspectives: Vol. 2. Handbook of applied developmental science: Promoting positive child, adolescent, and family development through research, policies, and programs* (pp. 197-223). Thousand Oaks, CA: Sage.
- Salusky, I., Larson, R. W., Wu, J., Griffith, A., Raffaelli, M., Sugimara, N. & Guzman, M. (2012). *How youth develop responsibility: What can be learned from youth programs*. Manuscript submitted for publication.
- Tversky, D., & Tirthali, D. (2010). Youth leadership development in virtual worlds: A case study. *Procedia Social and Behavioral Sciences*, 2, 3175–3179.
- Zagal, J. P., & Bruckman, A. S. (2005). From samba schools to computer clubhouses: Cultural institutions as learning environments. *Convergence: The International Journal of Research into New Media Technologies*, 11(1), 88–105.

Acknowledgments

We would like to thank the Scratch community members who participated in this research and the Scratch Team for their helpful feedback. The writing of this paper was supported by a grant from the National Science Foundation (NSF-CDI-1027736). The views expressed are those of the authors and do not necessarily represent the views of the Foundation or MIT.

Using Eye-Tracking Technology to Support Visual Coordination in Collaborative Problem-Solving Groups

Bertrand Schneider, Stanford University, schneibe@stanford.edu
Roy Pea, Stanford University, roypea@stanford.edu

Abstract: In this paper we present the results of an eye-tracking study on collaborative problem-solving dyads. Dyads remotely worked on contrasting cases to study how the human brain processes visual information. In one condition, dyads saw the gaze of their partner on the screen; in a control group, they did not have access to this information. Results indicate that this real-time mutual gaze perception intervention helped students achieve a higher quality of collaboration and a higher learning gain. Implications for supporting group collaboration are discussed.

Introduction

Joint attention is defined as "the tendency for social partners to focus on a common reference and to monitor one another's attention to an outside entity, such as an object, person, or event. [...] The fact that two individuals are simultaneously focused on the same aspect of the environment at the same time does not constitute joint attention. To qualify as joint attention, the social partners need to demonstrate awareness that they are attending to something in common" (Tomasello, 1995, pp. 86-87). Joint attention is fundamental to social coordination: young infants communicate emotions in a state of synchrony with their caregivers, in turn helping them achieve visual coordination when learning language (Stern, 1977). Parents use *deictic gestures* (i.e., pointing at a focus of interest to establish joint visual attention) to signal important features of the environment to their children (Bates et al., 1989). Professors and mentors teach by highlighting subtle nuances between students' and experts' conceptual understanding (Roth, 2001). Groups of students rely on coordination between their members to reach the problem solution (Barron, 2003), in turn influencing their level of abstract thinking (Schwartz, 1995).

We argue that the construction of perceptual joint attention rests significantly though not entirely (1) on two primary channels of communication: people can either point at things *physically* (i.e., using deictic gestures) or *verbally* (i.e., by describing the object of interest). These two mechanisms are not terribly efficient because misunderstanding can happen on several levels: verbally, communication is prone to misinterpretation from the receiver. This is likely to happen when experts are teaching novices, because novices are still learning the perceptual skills to isolate subtle features or patterns that separate them from experts. Physically, there is an extra step of taking the point of view of another person. From a spatial and social point of view, this is not a trivial mental operation (especially for children as demonstrated by Piaget in his studies of egocentrism and in more recent studies on the role of 'theory of mind' in human development; Leudar, Costall & Francis, 2004).

Previous work in CSCL used eye-trackers to study joint attention in collaborative learning situations. Richardson, Dale and Kirkham (2007) showed that common knowledge grounding positively influences the coordination of visual attention. Sangin (2009) studied pairs of students remotely working on a concept map and found evidence that knowledge awareness tools (i.e., displaying the level of expertise of each member of the dyad) was associated with a higher density of gaze-coupling and a higher quality of collaboration. Jermann, Nuessli, Mullins and Dillenbourg (2011) used synchronized eye-trackers to assess how programmers collaboratively work on a segment of code; they contrasted a good and bad dyad, and their results suggest that a productive collaboration is associated with high joint visual recurrence. Finally, Cherubini, Nuessli and Dillenbourg (2008) designed an algorithm that detects misunderstanding in a remote collaboration by using the distance between the gaze of the emitter and the receiver. Taken together, those evidences suggest that eye-trackers are a promising way to understand and predict the factors responsible for a high-quality collaboration.

Based on those studies, our goal is to develop new ways of supporting the establishment of *perceptual joint attention* (as opposed to cognitive, or social joint attention). We use eye-tracking technologies to share users' gaze during collaborative learning. More specifically, our first attempt involves dyads studying contrasting cases (Schwartz & Bransford, 1998). Our hypotheses are as follows: first, we expect dyads with access to their partner's gaze to have a higher quality of collaboration because such information will disambiguate their focus of attention and better enable "common ground" for learning conversations (Clark & Brennan, 1991). Secondly, we assume that a better collaboration will positively impact participants' learning gain (Barron, 2003).

General Description of the Experiment

The experiment had three distinct steps: during the first 12 minutes, dyads worked on 5 contrasting cases in neuroscience. They had to collaboratively explain how visual information is processed by the human brain based on what they have learned from the models described in Figure 1. They then read a text on the same topic for 12 minutes. Finally, they answered a learning test with questions on the terminology used, concepts taught and questions in which they needed to transfer their knowledge to a new situation.

Methods

Participants

Participants were 42 college-level students from a community college (average age 23.0, SD = 8.3; 28 females, 14 males). Dyads were randomly assigned to the two experimental conditions: the treatment group was in the “visible-gaze” condition (N = 24) with 16 females and 8 males; the control group was in the “no-gaze” condition, with 16 females and 8 males (N = 20). There was no significant difference in terms of GPA between the two conditions: $F(1,36) = 0.29$, $p = 0.59$ (for “visible-gaze”: mean = 3.09, SD = 0.87; for “no-gaze”: mean = 3.22, SD = 0.59). All participants were taking an introductory class in psychology and were required to participate in an experiment as part of their course.

Material

During the first step of the experiment, dyads worked on the contrasting cases shown in Fig. 1. To force collaboration, the answer of lesion 1 (top left) was visible only to the first member of the group while the answer of lesion 6 (top right) was shown only to the second member of the dyad. This kind of “jigsaw” method is commonly used to assure that one member of the dyad does not solve the problem alone (Aronson et al., 1978). The text used in the next step is available online(2). Finally the learning test contained 15 questions: 5 terminology questions (participants were asked to provide the name of a specific brain region or pathway), 5 conceptual questions (participants had to predict the effect of a specific lesion), and 5 transfer questions (subjects had to use their new knowledge to solve a vignette; e.g. “patient X is likely to have a lesion in region Y of the brain; should he be allowed to drive?”). All the material was exactly identical in the two conditions.

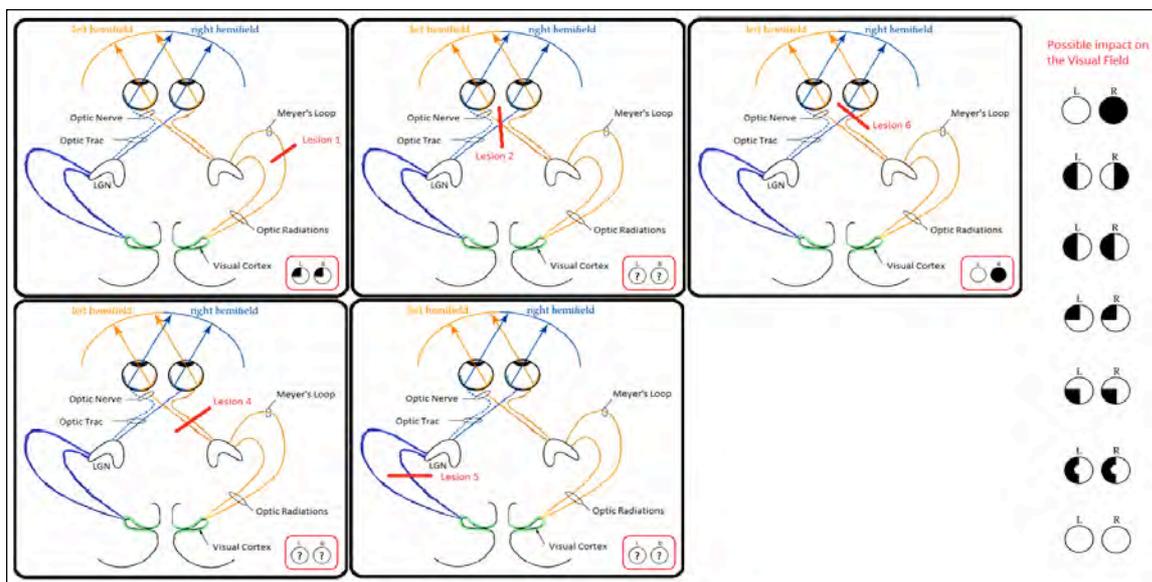


Figure 1. The dyads worked on the five contrasting cases above. Possible answers are shown on the right side. Answers of two cases (#1, top left and #6, top right) were given to subjects. Participants had to solve the three remaining cases (#2, top middle and #4, #5 bottom left and right of the screen).

Design

We used a between-subjects design with two conditions. In the “visible-gaze” condition, dyads were able to see the gaze of their partner on the screen. In the “no-gaze” condition, they could not. In the former condition, the gaze was only visible during the first step of the experiment (i.e., when dyads had to solve contrasting cases).

Procedure

Upon their arrival, participants were welcomed and thanked for their participation. The experimenter then explained that they would need to collaborate and suggested that they introduce themselves to their partner.

They were also told that each member of the dyad would be in a different room but would be able to communicate via a microphone. The experimenter explained that they would learn basic concepts in neuroscience, and he described the structure of the experiment (12 minutes of contrasting cases, 12 minutes of reading a text and as much time as needed for the learning test). Each participant then followed the experimenter to different rooms, where he calibrated their personal eye-tracker. The contrasting cases were then briefly presented to each participant and the experimenter ensured they understood the goal of the task. Subjects then worked on the contrasting cases and tried to determine how different lesions affected the brain's visual field. After 12 minutes, the screen automatically switched to a text explaining how the human brain processes visual information. The experimenter told participants they should read the text individually and then discuss it with their partner. After 12 minutes, the screen automatically switched to the learning test. The experimenter then told the subjects to individually complete the test and stopped the audio link. Participants took as much time as they needed for completion. They were then debriefed as the experimenter explained the goal of the study.

Measures

Because no participant had previous knowledge in neuroscience, learning gains were computed from the final learning test, which had three sub-dimensions: *conceptual questions* (predicting the effect of a particular lesion), *terminology* (naming brain regions or neural pathways) and *transfer questions* (solving a word problem using the concepts learned). The quality of collaboration was rated using dimensions developed in Meier, Spada and Rummel (2007), who assessed collaboration on a 5-point scale across 9 dimensions (sustaining mutual understanding, dialogue management, information pooling, reaching consensus, task division, task management, technical coordination, reciprocal interaction and individual task orientation). The evaluation of this rating scheme demonstrated a high inter-rater reliability, consistency and validity, which renders it an appropriate tool for assessing collaboration. Finally, we categorized each participant as being a "follower" or a "leader" in the activity. We acknowledge subjects are likely to shift roles while solving contrasting cases. This measure can be considered as an aggregate estimation over the whole activity of the dyad's dynamic profile. We used indicators to categorize the dyad's members: 1) who starts the discussion when the experimenter leaves, 2) who speaks most, 3) who manages turn-taking (e.g., by asking "what do think?", "how do you understand this part of the diagram?"), and 4) who decides the next focus of attention (e.g., "so to summarize, our answers are [...]. I think we need to spend more time on diagram x"). We also collected eye-tracking data during the experiment: approximately 30 data-points per second were captured for each participant. This gave us ~1'000'000 gaze points in total. Within those measurements, we also collected participants' pupil size.

Results

In this section, we compare main effects for learning gains and collaboration scores across our two experimental groups. We then characterize the dyads of our experiments in terms of their gaze patterns by analyzing our eye-tracking data. We also compare process variables in terms of their predictive effect as mediators. Finally, we conclude by conducting a small qualitative analysis of two dyads (one from each experimental group) to suggest mechanisms explaining the main effects found.

Learning and Collaboration

As predicted, we found that participants in the "visible-gaze" group outperformed the dyads in the "no-gaze" condition for the total learning gain: $F(1,40) = 7.81, p < 0.01$. For the sub-dimensions, they also scored higher on the transfer questions $F(1,40) = 4.47, p < 0.05$. The difference would likely to be significant with a larger sample for the terminology questions $F(1,40) = 3.59, p = 0.065$ and for the conceptual questions $F(1,40) = 2.11, p = 0.154$, since the effect sizes are between medium and large (Cohen's d is 0.62 and 0.5, respectively).

The treatment group ("visible gaze") also had a higher quality of collaboration as measured by Meier, Spada and Rummel's (2007) rating scheme (the total score is an average across the 9 sub-dimensions described in the "measure" section): $F(1,19) = 11.73, p < 0.01$ (mean for the treatment group = 0.89, $SD = 0.48$; mean for the control group = -0.08, $SD = 0.79$). More specifically, those dyads were better at *sustaining mutual understanding* ($F(1,19) = 5.15, p < 0.05$), *pooling information* ($F(1,19) = 7.53, p < 0.05$), *reaching consensus* ($F(1,19) = 22.57, p < 0.001$) and *managing time* ($F(1,19) = 4.98, p < 0.05$). A second judge double-coded 20% of the video data; inter-reliability index using Krippendorff's alpha was 81.63%. An alpha higher than 80% is considered as a reliable agreement between judges (Hayes & Krippendorff, 2007).

Additionally we categorized each member of the dyad as "leader" and "follower" (Fig. 1). Interestingly we found an interaction effect between those two factors (experimental conditions and individuals' status) on the total learning score: $F(1,38) = 5.29, p < 0.05$. Followers learnt significantly more when they could see the gaze of the leader on the screen.

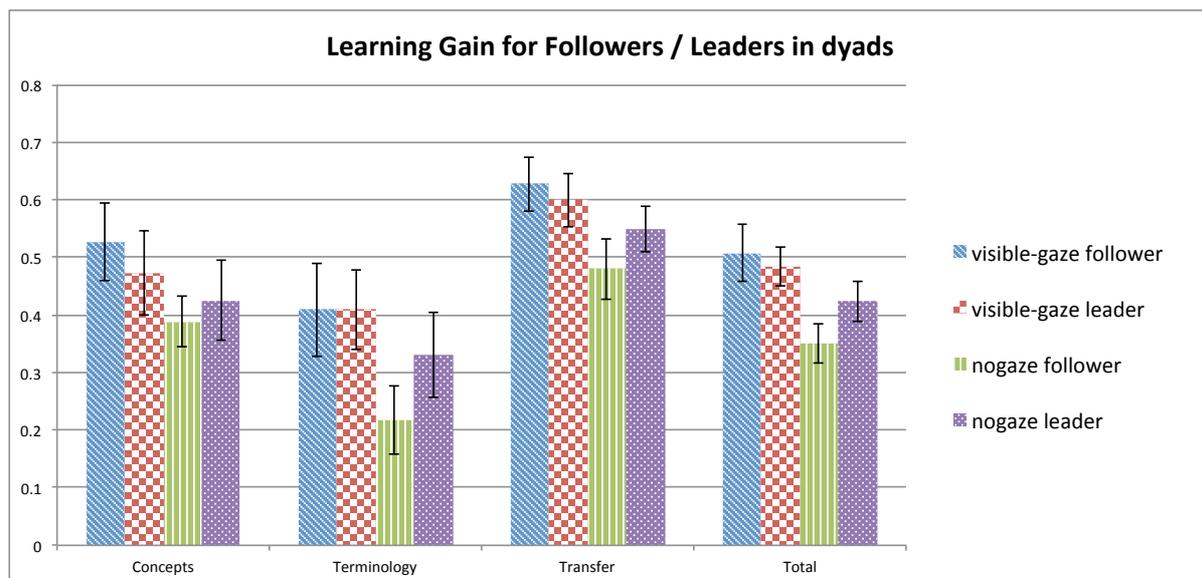


Figure 1. The total scores of the learning gain and the three sub-dimensions measured: conceptual understanding, participants' recall of the terminology, and transfer questions (crossed with two factors: experimental conditions and individuals' status in the dyad).

Eye-tracking Data

We isolated four kinds of measures from the eye-tracking data: first, we counted the number of fixations on the five contrasting cases and on the region showing the potential answers. Secondly, we aggregated the number of saccades between two regions from the six previously mentioned (i.e., 5 cases and 1 area for the answers). Thirdly, we defined a "joint attention" measure, where we counted how many times both participants looked at the same case on the screen. Previous research has shown that subjects need ~2 seconds to focus their attention on an object after a peer mentioned it (Richardson & Dale, 2005). We followed those guidelines to create our measure: for each data point, we checked whether the other member of the dyad was looking at the same area of the screen during the following two seconds. Fourthly, we used the size of the subjects' pupil as an indication of their cognitive load. Since eye-trackers react differently to different eyes' physiology, we divided each measure by the total number of data points for each subject. This yielded the percentage of fixations, percentage of saccades and percentage of joint attention. For cognitive load, we also subtracted the smallest value from each measure of a particular participant to take into account differences in eyes' morphologies. Participants' pupil size is not always a reliable measure, especially when the lighting conditions vary; however, since the room we used for the experiment did not have a window and thus had a constant lightning, we included those results for our analysis.

We excluded 5 subjects from those analyses because of missing data (i.e., the eye-tracker crashed during the activity). Three participants were in the "no-gaze" condition and two participants in the "visible-gaze" condition. We thus have 37 subjects when measuring the number of fixations and saccades and 16 dyads (32 subjects) when measuring joint attention. Due to space constraints, we will describe only a subset of our results.

We found that participants in the "no-gaze" condition had significantly more fixations on case 1 ($F(1,35) = 9.69, p < 0.01$), and case 3 ($F(1,35) = 4.92, p < 0.05$). Participants in the "visible-gaze" condition spent more time looking at the answers ($F(1,35) = 10.41, p < 0.01$). In terms of cognitive load, we did not find a significant difference between our two conditions: $F(1,35) = 1.09, p = 0.3$ (mean = 1.44, SD = 0.34 for "visible-gaze"; mean = 1.31 SD = 0.41 for "no-gaze"). The interaction between experimental condition and status in the dyad (i.e., leader or follower) is not significant: $F(1,29) = 2.51, p = 0.12$, but the effect size is between medium and large (partial eta squared = 0.08). It would be interesting to have more subjects to see if this result becomes significant. The pattern is similar to the one described for the learning test (i.e., followers have a higher cognitive load than leaders in the "no-gaze" condition, and a lower load than leaders in the "visible-gaze" condition).

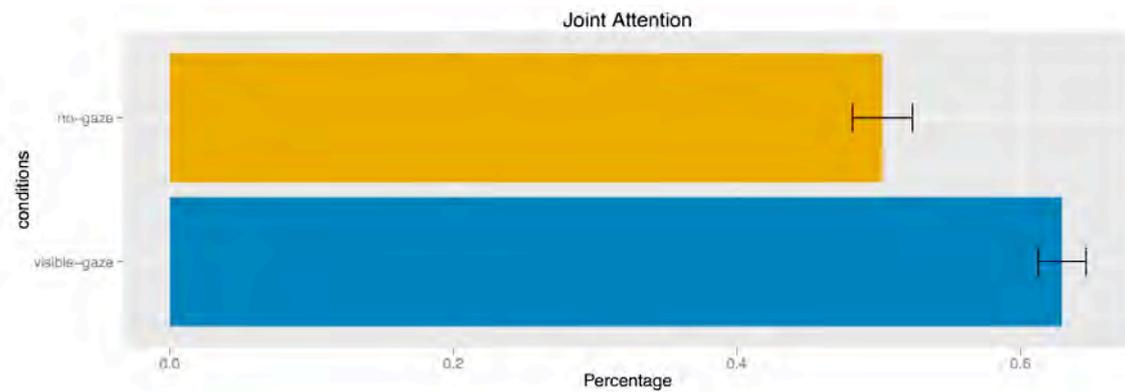


Figure 3. Percentage of Joint Attention in each experimental condition.

Participants in the “visible-gaze” condition achieved joint attention more often than the participants in the “no-gaze” condition (see Fig. 3): $F(1,30) = 22.45, p < 0.001$. This result holds when taking dyads (and not individuals) as the unit of analysis: $F(1, 14) = 16.36, p < 0.001$. The percentage of joint attention is one of the only measures correlated with a positive learning gain: $r = 0.39, p < 0.05$.

Model for Potential Mediators

In this section, we tested which process variables were most strongly associated with a positive learning gain. One may hypothesize that the quality of collaboration, the amount of effort produced by the participants, or the number of moments of joint attention may predict students’ learning. We tested for multiple mediation using Preacher and Hayes’ bootstrapping methodology for indirect effects (Preacher and Hayes, 2008). We used 5,000 bootstrap resamples to describe the confidence intervals of indirect effects in a manner that makes no assumptions about the distribution of the indirect effects. Significance is determined by checking if a confidence interval does not contain zero. We tested our model with the following candidates for being a mediator: collaboration, percentage of joint attention, cognitive load. GPA was used as covariate, since our goal is to find mediators regardless of participants’ grades. Results for multiple mediation indicated that only joint attention (CI: [0.03; 0.19]) was a mediator for learning (see Fig. 4).

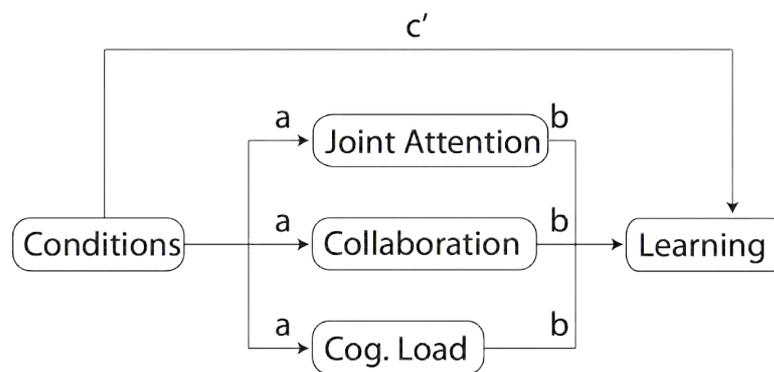


Figure 4. Mediation model for our data: we tested whether cognitive load (measured by the pupils’ sizes), quality of collaboration (measured by Meier, Spada and Rummel’s rating scheme) and percentage of joint attention (estimated with the eye-tracking data). Only joint attention was found to be a significant mediator.

Vignette

The previous section provides quantitative data on the effect of a gaze-awareness tool on students’ remote collaboration. However it does not provide us with any explanation for causal mechanisms. Table 1 tries to suggest answers to this question by comparing two dyads in terms of their gaze patterns. We compared two groups: one in the “visible-gaze” condition (left side) and one in the “no-gaze” condition (right side). The main goal of this comparison is to illustrate how our intervention changed the behavior of our participants. More specifically, we focused on four dimensions: students’ ability to coordinate themselves, create convention, build hypotheses and share theories.

Visible-gaze (P54-P55)	No-gaze (P07-P08)
<p><i>Lea (L) is the leader, and Flo the follower (F).</i></p> <p>(v1) establishing common ground (0:51) F: "I see this light blue dot" L: "I think that's my gaze" F: "Oh cool... that's so bizarre!"</p> <p>(v2) building anticipation (1:22) L: I have an answer for the... <i>[gaze moving to case 1]</i> F: <i>[gaze moving to case 1]</i> L: ... further most left one. F: Okay. Where the lesion is the orange colored thing.</p> <p>(v3) sharing hypotheses (3:45) L: maybe lesion two is... <i>[gaze moving back and forth between the two hemifields, eyes and optic nerves]</i> F: <i>[her gaze is moving from Lea's gaze to the other lesions]</i> L: those are both... they would be disrupted... I think that lesion two would be... <i>[gaze moving to the second answer]</i> F: <i>[gaze moving to the second answer]</i> L: the second one. F: why do you say that? L: because they are both going to be equal <i>[gaze moving to the lesion where two optic nerves (one from the right hemifield, one from the left hemifield) are severed]</i> F: <i>[gaze drifting to the same point]</i> oh right.</p> <p>(v4) creating implicit conventions (6:20) F: let's look at two again <i>[both gazes move to lesion 2]</i> L: everything is sort of cut off... F: well it's just the two in the middle L: yeah so that would be... left-left <i>[gaze moving to the left hemifield of the left eye, followed by Flo's gaze]</i> and right-right <i>[gaze moving to the right hemifield of the right eye followed by Flo's gaze]</i> F: <i>[gaze moving to the second answer, followed by Lea's gaze]</i> ... which would be the second one. L: yeah, which is the second one.</p> <p>(v5) sharing theories (7:34) L: so for the fifth we are not sure [...]. <i>[both gazes are exploring different cases on their own]</i> L: so maybe the further away from the eye it is, the less severe <i>[gaze moving from the eyes to the LGN on lesion 1]</i> F: <i>[gaze moving to lesion 1]</i> Maybe... what was lesion one again? L: that was the top left and top right <i>[gaze moving to the 4th answer, followed by Flo's gaze]</i>, the fourth one down. F: Oh... Ooooh... Hum <i>[gaze comparing cases 1 and 5]</i> L: so the one you had was right by the eye, and it was completely crossed out <i>[gaze on lesion 6]</i> F: <i>[gaze moving to lesion 6, then 5]</i> so maybe this would be similarly only a quarter of the eye L: <i>[gaze on answer 5]</i> yeah, maybe it would be the third one from the bottom <i>[followed by Flo's gaze on answer 5]</i> F: maybe... hum... <i>[gaze jumping from lesion 5 to answer 5]</i></p>	<p><i>Laurie (L) is the leader, and Fiona is the follower (F).</i></p> <p>(n1) establishing common grounds (0 - 0:30) L: Hi! F: Hi! <i>[laughing]</i> I don't get this stuff. L: I don't either! F: okay, so I have one with the answer <i>[looking at her case]</i> L: yeah I have an example too. <i>[looking at her case]</i></p> <p>(n2) sharing answers (2:37 - 3:45) F: so... <i>[gaze moving to lesion 1]</i> do you see lesion 1? L: yes <i>[gaze moving to lesion 1]</i>. F: I think it blocks Meyer's loop somehow. L: yes F: so the answer would be the left and the right... L: <i>[gaze moving moving between answer 4 and 5]</i> F: both the visions, they're blocked by one fourth. So it's not like completely blocked. So the answer would be that one. L: but how is it... hum... <i>[gaze still moving between answer 4 and 5]</i>. So I think it's the fourth answer down? Where the quarter is blacked out on the top? On the left? F: yes both right and left vision.</p> <p>(n3) sharing theories (8:19) F: <i>[gaze on lesion 5]</i> you said that lesion 5 would be the third from the bottom, right? <i>[gaze on answer 5]</i> L: <i>[gaze moving from lesion 5 to answer 5]</i> yeah I think so because it's blocking the left lower part <i>[gaze moving back and forth between lesion 1 and 5]</i> F: hu <i>[gaze moving back and forth between lesion 4 and 6]</i> L: but then again it kinda doesn't make sense because if the answer for lesion 1 was the top left, F: <i>[gaze moving to lesion 1, but then going back to lesion 4]</i> hu hu L: then wouldn't it be blocked on the opposite side of where the lesion is? F: <i>[gaze moving from lesion 1 to 5]</i> that's what I thought...</p> <p>(n4) sharing hypotheses (5:20 - 6:10) L: okay lesion 4... F: lesion 4 would be L: <i>[gaze moving from the lesion to the third answer]</i> I think it is the one that's half and half, the third one from the top. Because it blocks... <i>[gaze moving to the eye]</i> F: <i>[gaze moving to the third answer]</i> the left part of the vision? L: yeah I don't know F: maybe L: <i>[laughing]</i> F: Hum... maybe. I don't know <i>[laughing]</i>, whatever you say <i>[both laugh]</i></p>

In terms of coordination, we found a strong difference between our two dyads. More specifically, the *sequence of actions* was reversed: in the “visible-gaze” dyad, the leader would start talking about a lesion, and the follower’s gaze would go to the same area on the screen *before* the leader mentioned the lesion’s number (v3). In the “no-gaze” dyad, the follower would have the double burden of finding the lesion of interest and following the leader’s explanation in parallel (n2). We argue that our intervention facilitated coordination, and helped the follower anticipate the leader’s explanations. Secondly, we found interesting conventions in the “visible-gaze” dyad (v4): when Lea says “that would be... left-left, right-right”, neither of them ever explicitly stated that she referred to the diagram’s eyes and hemifields. Rather, they implicitly build the convention of moving their gaze as a deictic gesture to complement their explanations. Thirdly, we hypothesize that our intervention helped students share their cognition, even though they did not master the expert terminology of the domain: sentences

as vague as “they are both going to be equal” (v3) suddenly made sense when Lea pointed her gaze at the optic nerves to show that half of the information from each hemifield would be disrupted. This is particularly interesting because novices often lack the vocabulary to effectively communicate their assumptions. In our case, it provided Flo with additional information about the symmetry of the brain and helped her build her own hypotheses. Finally, we observed a tighter coupling between subjects’ attention in the “visible-gaze” condition (v5): gazes would “dance” together during a longer period of time and focus on the same lesions even though they were not explicitly mentioned. In the control group, the follower would briefly attend to the same lesion as the leader and then continue to explore other lesions (n3). This suggests that the theories built during the activity were more the results of the dyad’s shared cognition (in the “visible-gaze” condition), and more the results of individuals’ contribution in the “no-gaze” condition.

Discussion

Our findings demonstrate the importance of supporting joint attention in collaborative learning activities. We conducted a study where students needed to learn from five contrasting cases in a remote collaboration. In one condition, subjects could see the gaze of their partner on the screen as it was being produced. In the other, they could not. Our results reveal that this simple intervention was associated with subjects in the first group producing a higher quality of collaboration and learning more from the contrasting cases. In particular, subjects characterized as *followers* saw their learning gain dramatically increase. This result was partially confirmed by a similar pattern found for students’ cognitive load: followers in the control group spent more effort than leaders while learning less; followers in the treatment group spent less effort than leaders but learned more. We also found that subjects in the “no-gaze” condition spent more time on cases 1 and 3; this suggests that they took more time (and probably had more difficulty) sharing their answers. Participants in the “visible-gaze” condition had a higher percentage of joint attention, which proved to be a significant mediator for learning.

These results provide strong evidence for the important contributions of **real-time mutual gaze perception**—a special form of technology-mediated shared attention—to the learning gains and collaboration quality of collaborative learning groups. Additional qualitative observations suggest that our intervention helped students on four dimensions: by supporting coordination, creating conventions, sharing cognition and by making knowledge-building a collective process rather than an individual one.

One might argue that a shared pointer could achieve a similar effect. We believe that real-time mutual gaze perception has several key advantages over a shared pointer: first, there is a cognitive overhead associated with consciously moving a cursor to a region of interest, which may interfere with the learning task. A gaze awareness tool does this work automatically, without requiring additional effort from the user. There is also a certain amount of uncertainty associated with a cursor that stopped moving; is your partner thinking, being distracted, or waiting for you? By looking at the videos of our experiment, we saw that members of a dyad would perform some sort of “micro-monitoring” of their partner’s behavior, where they would check on their partner’s gaze every few seconds. We believe that a continual flux of gaze information reduces uncertainty and helps students regulate the dynamics of their dyad. In summary, we hypothesize that our gaze awareness tool enabled some behaviors that would not be possible with a shared pointer. Future studies are needed to demonstrate the unique affordances of each of those interventions.

This study has limitations. First, we studied a very specific kind of collaboration: situations where members of a dyad were communicating via a microphone and sharing a computer screen. It is not clear whether this kind of awareness tool would have the same effect in a co-located situation; one may assume that joint attention is easily achieved in face-to-face or side-by-side collaboration, but key papers in the learning sciences suggest that it may not be the case (e.g. Barron, 2003). Future studies using eye-tracking goggles on interactive surfaces will answer this question. Secondly, students had a very limited amount of time to work on the contrasting cases. It is unclear how this limitation impacted students’ performance. Thirdly, we only cursorily evaluated the transcripts of the dyads. More fine-grained coding schemes would provide additional clues as to how joint attention facilitated collaborative learning; the interaction effect between followers and leaders is especially interesting and should be analyzed in greater depth. Lastly, one may argue about the sub-categories describing the learning gains (e.g. it is debatable whether the questions about predicting the effect of a lesion are effectively measuring conceptual understanding); however, because we are not making particular claims about those sub-categories, and since the same pattern is repeated across our three learning sub-dimensions (i.e., the interaction effect between followers and leaders), we do not consider this issue to be a serious limitation of our findings.

In future work, we plan on evaluating the result of our qualitative observations. More specifically, we want to quantitatively measure the four dimensions we uncovered and show that those processes are significantly different across conditions. Secondly, a next logical step is to investigate this phenomenon in a more natural

setting (e.g., in a co-located situation). Eye-tracking goggles could offer an interesting tool for this purpose. Thirdly, it would be interesting to see if those results generalize beyond contrasting cases; it may be that this intervention is only effective for perceptual tasks. Finally, our results suggest that supporting joint attention between novices and experts would bring interesting results, as real-time mutual gaze perception provides a form of “inter-identity technology” (Lindgren & Pea, 2012). As followers, novices could more easily share their understanding of concepts without having to know the expert terminology; additionally, it would disambiguate experts’ explanations by providing perceptual clues to novices (Hanna & Brennan, 2007).

Endnotes

- (1) Attentional alignment is also established partly by body position and orientation (Kendon, 1990).
- (2) The text used in the second part of the study is accessible here: <http://www.scribd.com/doc/98921800>. Originally retrieved from Washington University in St-Louis (<http://thalamus.wustl.edu/>)

References

- Aronson, E., Blaney, N., Sikes, J., Stephan, G., & Snapp, M. (1978). *The jigsaw classroom*. Beverly Hills, CA: Sage Publication.
- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12(3), 307-359.
- Bates, E., Thal, D., Whitesell, K., Fenson, L., & Oakes, L. (1989). Integrating language and gesture in infancy. *Developmental Psychology*, 25, 1004-1019.
- Cherubini, M., Nüssli, M., and Dillenbourg, P. 2008. Deixis and gaze in collaborative work at a distance (over a shared map): a computational model to detect misunderstandings. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications* (Savannah, Georgia, March 26 - 28, 2008). ETRA '08. ACM, New York, NY, 173-180.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. Levine, & S.D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.
- Hanna, J. E. & Brennan, S. E. (2007). Speakers’ eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, 57, 596-615.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures*, 1(1), 77–89.
- Jermann, P., Mullins, D., Nuessli, M.-A., Dillenbourg, P. (2001). Collaborative Gaze Footprints: Correlates of Interaction Quality. In *Proceedings of CSCL 2011*. Hong Kong.
- Kendon, A. (1990). *Conducting Interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.
- Leudar, I., Costall, A., & Francis, D. (2004). Theory of Mind A Critical Assessment. *Theory & Psychology*, 14(5), 571–578.
- Lindgren, R., & Pea, R. (2012, July). Inter-identity technologies for learning. *Proceedings of the International Conference of the Learning Sciences (ICLS2012: The Future of Learning)*. Sydney, Australia.
- Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning*, 2, 63–86.
- Preacher, K. J., and Hayes, A. F. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods* 40, 3 (2008), 879.
- 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, 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.
- Roth, W. M. (2001). Gestures: Their role in teaching and learning. *Review of Educational Research*, 71, 365–392.
- Sangin (2009). *Peer knowledge modeling in computer supported collaborative learning*. (Doctoral Dissertation). Retrieved from <http://infoscience.epfl.ch/record/133637>.
- Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences*, 4, 321-354.
- Schwartz, D. L. & Bransford, J. D. (1998). A time for telling. *Cognition & Instruction*, 16, 475-522.
- Stem, D. (1977). *The first relationship*. Cambridge, MA: Harvard University Press.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

‘CO-ALIENATION’ MEDIATED BY COMMON REPRESENTATIONS IN SYNCHRONOUS E-DISCUSSIONS

Baruch B. Schwarz, Yifat Ben-David Kolikant, Maria Mishenkina,

The Hebrew University of Jerusalem, Mt Scopus, Jerusalem 91905, Israel, baruch.schwarz@mail.huji.ac.il

Abstract. Synchronous e-discussions have become common social practices in and out of educational institutions. Socio-cultural psychologists have suggested that intersubjectivity is central for maintenance of face-to-face communication. We study here how communication is maintained in synchronous discussions and whether intersubjectivity is reached. Four university students used a CMC tool to discuss an educational issue. One week later, each student was interviewed on his/her views on the issue. Then, the technique of cued retrospective reporting was used to uncover how each student interpreted each move of the synchronous discussion. The cross analysis of the interviews and the reporting showed that actions were not co-ordinated. Agreements and disagreements were not shared, and order of actions was quite whimsical. We conclude that intersubjectivity was not established. However, communication was maintained through a process of *co-alienation* – the juxtaposition of incompatible alignments of representations through a common external representation. Although co-alienation is problematic, we show that discussants could learn from the e-discussion.

Four Senses of Intersubjectivity to Account for Maintenance of Communication

The notion of intersubjectivity was elaborated to account for maintenance of communication in practices that lead to development and learning. The social practices of synchronous (electronic) discussions are newcomers that seem strange to educators or psychologists: People seat alone by their computers and interact with others they don't see and often don't know. The interactions are often lopsided, interrupted and rudimentary. How can learning occur? Is it possible to discern any kind of intersubjectivity in synchronous discussions? We approach these questions in this paper. A necessary step in this enterprise is to come back to the different senses intersubjectivity has received to account for maintenance of communication.

The historical origins of *intersubjectivity* are philosophical: From Husserl, to Heidegger, Levinas or Gadamer, philosophers have asked the transcendental conditions of the possibility of dialogue, of the existence of the other. These philosophical steps led to consider intersubjectivity as an idea that explains how empathy develops between people or how people reach new meanings together. Socio-cultural psychologists adopted this togetherness view of intersubjectivity to become a central idea in their theory of human development. Intersubjectivity was primarily meant to express a general idea of overlapping of subjectivities or *prolepses* (Rommetveit, 1979; 1985) – communicative moves speakers take for granted things that have not been (yet) discussed. Even before language is mastered, a baby is born anticipating a sympathetic, interactive social environment. Another sense of intersubjectivity refers to the construction of a *shared meaning* by people to interpret social and cultural life, or a specific situation. This sense is very frequent in adults' guidance of children's explorations. It has also been adopted to describe some situations of collaboration among peers, for example, when jointly solving problems. A different sense given to intersubjectivity – *shared divergences of meanings*, is necessary to understand how practices such as lying, disputes and jokes “work”. Lying is genuinely inter-subjective as it operates between two subjective definitions of reality. Disputes, for example, are fueled by the common recognition of divergent views on a situation. So far, the senses of intersubjectivity we reviewed have in common the sharing of subjective states by two or more individuals, and stress that shared cognition and consensus is essential in the shaping of our ideas and relations.

But intersubjectivity may occur without any sharing. To make clear this new sense of intersubjectivity, Matusov (1996) observed a succession of discussions on play craft among elementary school children mediated by a teacher. Matusov showed that in the disagreements some students raised no idea or no opposition to any idea. Rather, although a fierce disagreement arose among children, the disagreement for some turned around ownership of the play crafted while the other students thought it was motivated by a better play craft. Matusov concluded that a new sense of intersubjectivity should be elaborated, a process of *coordination of participants' contributions in joint activity*. In such cases, an observer can extract from the social activity a thread of compatibility of actions, that is, an interpretation encompassing the goals of each participant, and which is coherent. Matusov considered other situations in which meaningful communication can be maintained – situations in which teachers or students care for others and help them understand, develop, or own what they already master. Communication in this case is maintained but calls for another sense of intersubjectivity, that of

agency to features concerns by teachers (or students) that are shared by the ones about which they care (Matusov, 2001).

The studies mentioned above on the different types of intersubjectivity involved face-to-face communication. The diversity of communication practices grows extremely rapidly, as various forms of e-discussions gain popularity. We show that the sense of intersubjectivity as sharing of subjective states among several individuals, cannot be retained for certain synchronous discussions. The question is then how communication is maintained? And is it possible for discussants to learn from such discussions? Those are harsh questions that we will approach in this paper in the case of synchronous e-discussions with a particular tool for facilitating *collective argumentation* in synchronous discussions (Schwarz & Glassner, 2007). This particular tool is representative of a quite large set of tools for facilitating e-discussions that displays a representation of the on-going discussion, including its whole history. We will claim that this gradually produced representation (called here an *argumentative map*) which is shared by the discussants while participating in their discussion, enables maintenance of meaningful communication among them, although possibly no idea is shared and no coordination of actions is attained during the discussion. In an example, we will show that the argumentative map produced by the discussants will serve as an artifact for projecting personal interpretations of the synchronous discussion so far, which would have probably led to ruptures if communication would have been face-to-face. We will point at the cohabitation of incompatible prolepses, or what we call *co-alienation* along with new kind of communication. We will show that the discussants can learn from such discussions. We will discuss whether in this case, a new sense should be conferred to intersubjectivity. But, before we delve into the analysis of the synchronous discussion, we first review what is known in communication studies about synchronous e-discussions. As we will see, this review appears to be highly relevant to the issue of maintenance of communication.

A priori, e-discussions seem detrimental to maintenance of communication, hence, by definition, to intersubjectivity: In spoken discussions, joint attention is very often maintained. Turn-taking is determined by facial expression, pause of the other, intonation, or simply interruption. Overlaps can occur but they generally occur through gestures, or facial expressions. In contrast, descriptive studies of computer-mediated communication (CMC) suggest that maintenance of communication may often turn to problematic: time lag between when a message is sent and when it is responded to, disrupted turn adjacency, overlaps with unrelated messages from other participants between an initiating message and its response (Murray, 1989) are characteristic problems in e-discussions. According to Herring (2001), these problems are responsible for incoherence, and for topic decay – the fact that discussants rapidly discuss less and less the topic at stake during e-discussions. On the other hand, e-discussions are attractive: CMC communication is based on the availability of a *persistent* textual record of the interaction. Persistent conversation appears to aid the user's cognitive processing. The predilection towards meta-humor and meta-play in CMC may be attributed in part to the fact that CMC persists as text on a screen and is subject to conscious reflection in ways that spoken language is not, thereby facilitating a heightened meta-linguistic awareness (Cazden, 1995). This short review suggests that studying maintenance of CMC communication is worthwhile.

Studying Maintenance of Communication in Synchronous Discussions

The context of the present study is educational. Discussions took place in a University course on *The Role of the Teacher in Classroom Learning*. One of the themes of the course was to focus on the role of the teacher in structuring interactions in group work. In addition to theoretical lectures, the teacher provided a CMC graphical tool for facilitating group e-argumentation, and for reflecting on ways teachers could intervene during discussions (Schwarz & de Groot, 2007). This tool named *Digalo* enables the management of discussions and the representation of their argumentative processes and components among participants. Using *Digalo* consists of co-creating maps built of written notes inside different shapes that represent the permissible argumentative moves for the discussion (or what is called the *ontology* chosen for the discussion), and different arrows representing different connections between the shapes. The ontology chosen in the course included Claim, Argument, Explanation, Comment, and Question. There are three kinds of arrows: support, opposition, and link. The support and opposition arrows help challenging, refuting, elaborating, or agreeing. In each of the contributions, participants add one shape or more and arrows/links to shapes built by others to articulate own claims, arguments, etc, and then write their contribution inside of the shape(s) they chose. Figure 1 shows the *Digalo* map on which we will focus in this paper.

11 graduate students participated in a university course. This article focuses on one of the discussions in the course. Judith is a 26 years old graduate student in Educational Psychology. Fatima and Rim are students in Education in their twenties. Ahmad is a 47 years old teacher. In a first session, students were introduced to *Digalo* and conducted two “warm-up” discussions during which they learned technicalities. They were then arranged in groups of 3-4 and asked to resolve a moral dilemma. The presentation of the ontology was accompanied by a suggestion to follow a series of ground rules of conversation developed to instill dialogic and

dialectical talk (e.g., “you should provide reasons to support your viewpoint”, “You should challenge any idea”, “you should answer to challenges or questions with respect”).

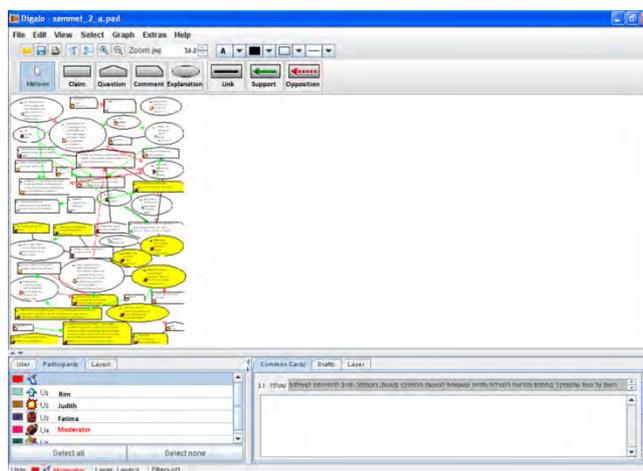


Figure 1. A Digalo Map

At the end of the first session, the students were given two articles to be read before the next lesson. Both articles related to the issue of guidance in discussions. Both articles suggest that the total absence of guidance during discussions is rarely productive but that moderation, a kind of guidance which is caring but not intrusive, may be one way to help students in their discussions. One moderator was designated for each group, and was asked to insure that discussants follow the ground rules we presented. One week after, all students sat by their own computer at distance from each other. The issue to be discussed was “Is it indispensable for research on classroom learning to account for the teacher's activity?” The discussions lasted around 30 minutes. Two weeks after the discussions, we meet with each of the discussants. The first part of the meeting was an interview in which the students were asked on the role of discussions in learning and on the role of the teacher in discussions. In the second part of the meeting, each student underwent a cued retrospective reporting (Van Gog, et al., 2009) in which he was presented the progressive reconstitution of his argumentative map by using the replay option of the Digalo software: whenever the interviewer clicked, the map grew by one argumentative move. Students were asked to explain their moves, thereby give their interpretation of each argumentative move, such as peers’ or moderator’s interventions, own reactions to these interventions, and to figure out the goals and expectations. Each of the meetings lasted around two hours.

Three Incompatible but Cohabiting Stories by Three Discussants

How was the discussion perceived by the participants in each of its moves? To answer this difficult question, we had at disposal the Digalo map they gradually produced, and the cued retrospective of Judith, Fatima, and Ahmad with the replay of the map (Rim, did not agree to participate in the retrospective). The preliminary interview of the discussants provided an indispensable background to interpret the e-discussion. We will insert some insights from the interviews to complete the discussants’ interpretations of the discussion in their cued retrospectives. As we will see, interpretations were quite divergent: they did not convey shared understandings, or even coordinated actions. Since the nature of communication was so precarious, the division of the discussion in episodes was a difficult enterprise: some moments could represent boundaries for some discussants, while they could be apprehended as in the middle of a move by others. Our approach was to propose a division into episodes according to clear changes in patterns of interaction. The discussion includes 49 contributions. The first episode – Turns 2-20 is characterized by the non-intervention of the moderator, and the quite equal participation of the three discussants. The second episode – Turns 21-37, is characterized by the participation of all discussants, including the moderator, who sends messages to all discussants. The pattern of interaction is quite centralized. In the third episode – Turns 38-49, the moderator continues to send messages to all discussants but only one discussant is responsive.

1st episode: Starting the discussion without moderator – accumulating similar contributions

The beginning of the discussion is characterized by short contributions in which no challenge is raised. Rather, the map seems to grow with contributions that express the same point: at Turn 8 and 13, Fatima writes *In my opinion research should refer to the position of the teacher, no one can overlook him!* and *I agree with you* (what is written in the e-discussion is in underlined italics to differentiate it from what is said in the interview);

at Turn 11, Rim writes *In my opinion, any research on learning and instruction should refer to all the directions, students, teachers, parents, school*. Judith writes at Turn 12 *We all agree. Is it possible?* and ostensibly links this recognition of agreement to all previous interventions with arrows of support. Ahmad did not make any contribution in the first episode because he faced technical problems. However, as he later reported, he watched the map, and thought how to act in this situation. We will see that all participants interpreted this situation differently and that their interpretations relied on their beliefs regarding the role of discussions in learning, and the role of the moderator and of the discussants in e-discussions.

Ahmad: Discussants do not really refer to each other; they agree instead of thinking in depth. When Ahmad observed this discussion in his cued retrospective, he reacted: *the participants began by throwing out ideas... and then everybody agrees, agrees, agrees, and in fact, I don't know about what*. In his view, what happens here is *disagreement which is called an 'agreement'*. When Ahmad read Judith's contribution at Turn 12 (*We all agree. Is it possible?*), he felt that *this means that we are afraid of ourselves. Is it possible that we agreed? On what did we agree? We agreed on one sentence [...], the fact that the teacher should be at the center of research and should look at the students. But we didn't formulate this properly*. He recalled the importance of collaborative learning to contrast it from the *'railroad parallel tracks'* (sic) that characterized the beginning of the talk. Ahmad thought that discussants should slow down their pace and begin to read others' contributions to *uncover their real meaning*. Ahmad raises an additional idea, the importance of *formulating own opinion* or of *creating own position*. He criticized the Arab society to which he belongs, that does not enable young people to think autonomously: *the Arab society is clearly patriarchal. The father decides on everything, and everybody has... sometimes, this is not the father, this is the grandfather [...] and everybody must tell him 'yes' and nobody argues with him, nobody argues with him about things that are important in their lives. The society was built in a way that only when you get old, you can think, you are allowed to think. Before you are old, you must listen, you must accumulate wisdom*. It seems that Ahmad sees the problems of his society through the lenses of Rim and Fatima's contributions.

Judith: It is impossible to learn from this discussion. As she read the contributions of the other discussants in her auto-confrontation, Judith reminded that she had no interlocutor for a real discussion. For example, her reaction to Rim's contribution at Turn 7 (*I agree with your opinion Ahmad that the teacher should be a partner*) is: *It seemed to me that it's not enough articulated [...] it means to say 'yes' to the main issue, but this is not enough [...]. I felt that I don't have any interlocutor with whom to argue*. She compares her pace in her contributions to theirs to say: *see how much I write [laughing] ...I felt that they don't move*. Judith was interested in changing this situation: *I tried again and again, as much as I could, to move things*. She commented on her contribution at Turn 12 (*We all agree. Is it possible?*), that she linked with arrows of support to all previous contributions) as *'a bit cynical'*. For her, the contributions of the others are like building bricks, and each discussant should bring personal contributions of high quality, original, and warranted: *When you express only your opinion, without going deep into it, without explanations, without bringing citations, warrants to what you say [...], it's superficial, it's to remain in 'I think that'*. It is then understandable that she is disappointed by the beginning of the discussion. In spite of her disappointment, Judith continued reacting to her interlocutors and expressed her opinion with the hope that it would develop as a more interesting discussion in which she will have opportunities to explain her position. Differently from Ahmad, she put the responsibility for the learning in the individual rather than in the group. However, she thought that *the discussion quickly became purposeless*.

Fatima: The discussion is successful because people express themselves and the moderator is not intrusive. As Fatima read Judith and Rim's contributions, she thought that the discussion was *excellent, not only good* at that stage: *each of the discussants expresses what she has*. For her, this situation is surprising and does not necessitate any moderator. Fatima asserts that the discussion is successful because it does not stop after the first contribution: *They [the discussants] forgot that there is no moderator and continue talking onward*. When reading Judith's Turn 12, she understands that Judith wants to serve as the moderator. She does not identify any cynicism in Judith's intervention (as meant by Judith). Rather, she interprets this turn as a sincere request to figure out the intentions of the discussants: *She [Judith] wanted to take the role of the moderator who tries to understand whether we all are on the same wavelength, whether we agree on the same things or not*. In her opinion, this role is not necessary: *the discussants are interested in this topic ... We want to go on although we have no moderator*. Fatima's satisfaction is consistent with the opinions she uttered in her interview: for her, discussions contribute to learning by the simple expression of diverse opinions. She expresses her surprise that the discussion goes on without any guidance, since, according to her experience, *students do not speak* and the role of the teacher is to cause them to speak. Understandably, since students speak, the moderator is unnecessary.

Second episode: Ahmad tries to change the flow and the quality of the discussion

The discussion which began with ungrounded agreements between the discussants, developed into a harsh dispute: Fatima and Rim argued that the student is at the center (Turns 14, 16, 19), against Judith who argued that opposed this view and added that it is impossible to take into consideration all the variables that are relevant to the educational field (Turn 18). Initial agreement quickly turned to a harsh polemic. At this point, Ahmad succeeded to enter the discussion. At Turn 21, he writes: *Let's make some order here, we should know on what there is an agreement.* At Turn 27, he intrudes into the gist of the discussion by asking the question: *What do you think about the saying according to which the student is at the center and the teacher only disturbs in his learning?* However, the discussants do not change the way they discuss the issue. The map fills up with short interventions popping out at a dizzying pace. Ahmad, who tries to stop this trend by writing (in Turn 30) *Please stop throwing out things without seriously referring to what is written* immediately receives Rim's reaction *We don't simply throw out* (Turn 31). In Turn 33, he links Turn 31 to her contribution in Turn 11 where she previously wrote *In my opinion, any research on learning and teaching should take into consideration all the directions – students, teachers, parents, school,* to ask (in Turn 33) *UI what about opening the idea in contribution 11.* This subtle move seems to point at Rim's incoherence, but she is not sensitive to it.

Ahmad: Efforts to organize the discussion, and to foster collaborative learning. In his auto-confrontation with this episode, Ahmad complains that the lack of reference to others' ideas originates from the fast pace of the discussion: *Too much, too fast, things they threw up. I had the feeling that it's raining.* Ahmad tried to lead discussants to scrutinize the ideas of each other and to refer to previous contributions expand them, or, in his own terms, *to open them.* In his opinion, the presence of different discussants facilitates the consideration of multiple perspectives and naturally leads to the necessity to explain. Consequently, Ahmad tries to slow down the pace of contributions. His comment of his contribution at Turn 30 (*Please stop throwing out things without seriously referring to what is written*) is: *So, I explained that you should be more focused, more ordered [...], that you should stop, that you should think. Because there was such a deluge of contributions.* This comment conveys a quite high emotional state against the behavior of the discussants. We already saw that Ahmad entered in the middle of the discussion (e.g., in Turn 27: *What do you think about the saying according to which the student is at the center and the teacher only disturbs in his learning?*). Viewing Turn 27 he explained: *I came to tell those who agreed: listen... there are other viewpoints...From the beginning, I didn't know on what they agree and on what they disagree. I try to position a conflicting viewpoint, an antithesis...to arouse the issue.*

Reaction to Ahmad in Ahmad's eyes: The discussants do not agree with his description of the situation. In his cue retrospective with the reaction to his injections, Ahmad assessed that the discussants adopted a negative position: When in Turn 31, Rim writes: *We don't simply throw out*, Ahmad comments: *This means that Rim thinks that what she says is important and is not superfluous. This means that she doesn't agree with what I said [at Turn 30].* Ahmad considers Rim's contribution as a disagreement on the definition of the situation in which the discussants are evolving. To handle this disagreement, he attempts to model his viewpoint in Turn 33 by writing *UI what about opening the idea in contribution 11* with a link to Rim's opposition at Turn 30. Ahmad comments: *I linked this [Turn 33] to 31 to say that if people don't refer to that, this means that they write for no reason [...] because she said that this is not without any purpose.* And Ahmad continues: *her contribution [Rim's contribution] at Turn 11 in which she says 'In my opinion, any research on learning and teaching should take into consideration all directions, students, teachers, parents, school' is good, there is a lot of content here, but one should develop it, I mean that we should elaborate upon it, explain and warrant it. It was written and no one referred to it. It's a pity that it passes by without mention [...]. This was my intention, that if one does not develop these ideas, he writes without any purpose.* (Rim did not react to this intervention). Ahmad's comments show how he used the Digalo map to convey suggestions to the discussants to improve the quality of the discussion by referring to previous contributions. When Ahmad affirms *I linked this to 31 to say that it was written and no one referred to it so it has no sense*, he makes clear that the growing map is a central artifact for constructing a shared meaning. The way Ahmad reacts to Rim's disagreement totally fits his positive approach to criticism and his demarcation from blind agreement. In the rest of the discussion, Ahmad acted to lead discussants to refer to what their peers write by requesting them to refer to specific contributions, by presenting them an idea that opposed what they wrote to one discussant.

Judith: The moderator was not noticeable. Judith attempts to justify the fact that she and her peers did not react to Ahmad's contributions by pointing at technical mistakes that Ahmad did, despite his good intentions. For example, in her comment on Turn 30 (*Please stop throwing out things without seriously referring to what is written*), she admits that this intervention can help the group in organizing and promoting the discussion: *The moderator belongs to the group. And he represents the goal which is common to all members of the group [...] to take a decision, or to reach a conclusion. And this is the role he has – to help in this goal.* Judith comments that in his interventions Ahmad complies with his roles of moderator. However, Judith estimates that he failed because he did not benefit from the visual aspect of Digalo: *He says important things, but does not locate*

himself properly ... he does not think about that. Do people listen to him, don't listen to him, see him or don't see him? And indeed, it seems as if Judith did not notice Ahmad contributions at all and she assumes that the same happened to the rest of the discussants. In her comment on his contribution in Turn 30, she explains: *I didn't see him at all; in my opinion, it's very important, and if nobody notices you, you can't organize a discussion.* She suggests that Ahmad could have created a personal reference at the right place *in the middle of the map and he could link it to everybody's last contributions with arrows; not to put himself like that, on the side, in the corner.* For Judith, the growing map is a central artifact for constructing shared meaning. Judith does not put all responsibility for this failure on Ahmad but mentions that the map is in disorder, and contributions hide one another, something which makes it difficult to organize a discussion.

Besides this critique, Judith praises Ahmad for his efforts to improve the quality of the discussion. While she notices that his contribution in Turn 30 (*Please don't throw out things without seriously referring to what is written*) is a bit harmful, she asserts that the idea is good: *the discussants really do not devote enough thought on their contributions.* Although she claimed that she didn't notice Ahmad, it seems that Judith reacted once to Ahmad's effort to promote the quality of the discussion, and when he wrote that one should take advantage of what the authors of the papers wrote to see what pros and the cons are to write in Turn 36: *We saw in Howe's research that interventions by the computer are equivalent to interventions by the teacher. But I claim that still, a teacher was necessary to organize this computer event. So it's impossible not to refer to the teacher.* We have here one of the rare moments in which a reaction is made on the basis of a shared understanding. Also, as Judith attends Rim's disagreement and Ahmad's reaction, her interpretation is: *The moderator tries a new way, as he didn't succeed in explaining himself [...] He shows her what he means, what is my problem with what you said.* This interpretation fits Ahmad's explanation of his act. But Judith was surprised to see Ahmad trying another way to lead discussants to act more effectively, rather than simply doing it instead of Rim. The difference between Ahmad's and Judith's reactions corresponds to their perception of the role of moderator. Judith lacks the vision of a moderator as a manager that Ahmad has. In her cued retrospective Judith appreciated how Ahmad tried to promote the quality of the discussion. However, she sees in his failure a lack of clarity.

Fatima: The intrusion of the moderator turns the discussion to uninteresting. In her cued retrospective, Fatima is quite resolute about her feelings towards what she sees as the intrusion of the moderator. She considers Turns 30 and 33 as examples of bad guidance. Turn 30 (*Please stop throwing out things without seriously referring to what is written*) is totally unacceptable: *This is not OK...he was not with us at the beginning, and he can't tell us such things.* She justifies Rim's reaction in Turn 31 (*We don't simply throw out*). Such an intervention from the part of the moderator *does not encourage the continuation of the discussion.* She interprets Ahmad's link to Rim's contribution in Turn 31 in his contribution (Turn 33: *UI what about opening the idea in contribution 11*) as a way to ask questions on claims that were far away and he wants us to return to Turn 11, what's the link? And Fatima goes on explaining: *He has to know that many things happened since then. So, his reaction here [in Turn 30] is not interesting at all and is connected to nothing [Fatima does a gesture of 'going away' with her hand].* Fatima feels that the rest of the activity of the moderator is even worse as he does not show any interest in her and Rim's reaction. Her feeling is based on the pace of his contributions in Turns 30, 32, and 33: *He didn't wait for our reactions. Rim reacted but it doesn't seem that he waits that somebody else reacts, or asks a question anymore.* This kind of behavior leads Fatima to stop contributing, yet she kept reading the discussion, which she found to be "not very interesting": *I only read. In the discussion, I read all the contributions and this was not very interesting.* Fatima's feeling is mediated by the fast development of the map (done by the other participants) which conveys for her a lack of consideration from the part of the moderator. Fatima's interpretation is consistent with her feeling of satisfaction in the first episode, when Ahmad did not intervene. It also fits what she said in her interview about the detrimental role of the teacher in discussions between students. In accordance with this view, Fatima opts not to react to Turn 30 that she considers as harmful. Rather, she reacts to Turn 27 (*What do you think about the saying according to which the student is at the center and the teacher only disturbs in his learning?*). She explains this choice by asserting: *In that way, he doesn't say that all what we said before doesn't matter or is nothing...in some way, he took what we already said and the main issue and assembled them together. He gave us a second question, a question more or less connected to the main question and to what we said before.*

The end of the discussion: a common summary or the moderator's aggressive coercion?

We can't enter into details in the end of the discussion. Its dynamics changed in comparison to what happened so far: The moderator referred to Judith's contributions, and she reacted to them. Ahmad gave to Judith ideas that complete or challenge the ideas she brought forward. Judith identified this new style as that of a *challenging moderator*. Judith responded by adding more ideas or more details to ideas already expressed. She felt that there is a learning layer in the discussion, and also learns from the inventive strategies he adopts to go deeper in

the discussion and to improve its quality. In contrast, the two other discussants vanished from the scene and did not contribute even when Ahmad turned to the whole group to refer to the discussion in a reflective way. Ahmad's attempt to instigate a common conclusion encountered Judith's strong opposition, who was happy that she could finally develop her own ideas and who thought that the opinions brought forward by Rim and Fatima are not academic enough. Judith saw in their desertion from the discussion an additional proof of their weakness and of their lack of motivation, whereas Fatima and Rim's e-silence is for them all but a lack of attendance; it is loaded, full of anger, inscribed there on a map that shows their presence.

Conclusion: Co-Alienation Mediated by a Common Representation as a New Manifestation of Maintenance of Communication

The idea of intersubjectivity has been developed to explain how people maintain communication in various social practices. Matusov has discerned several kinds of intersubjectivity that can be classified as (a) overlapping of subjectivities, (b) sharing of (divergences of) meaning, (c) coordination of participant's contribution in joint activity and (d) human agency. By definition, all kinds of intersubjectivity describe diverse forms of states of subjectivity **shared** among several individuals. These instances of intersubjectivity were recognized as being of the highest importance for learning and for development of productive guided participation (Rogoff, 1990). The particular setting of this research has uncovered what we consider as not being classifiable to any of these categories: the interview uncovered very different beliefs on learning, on the role of discussions for fostering learning, on the role of moderators, and on whether Digalo helps discussants and the moderator. We anticipated that these beliefs would come to the surface in the discussion. We expected that disagreement would arise and would be fueled by the common recognition of divergences. The confrontation of the discussants with their previous e-discussion in the cued retrospective showed a different picture. Ahmad, Judith, and Fatima had divergent interpretations of their synchronous discussions, but this divergence was not shared: what was meant to facilitate construction of knowledge from the part of Ahmad was interpreted as brutal interference by Fatima; What was meant to be an exchange of ideas (by Fatima and probably by Rim) was seen as a shallow discussion that does not lead to learning by Judith and Ahmad. There is even not any clear cut between the presence and the absence of discussants: Ahmad's absence in the first episode for technical reasons was interpreted as a welcome 'presence', a tacit agreement to give students the opportunity to talk about the issue at stake without interfering. Clearly, actions are not *co-ordinated*.

How, with all those unshared divergences and the absence of coordination of actions, communication could be maintained? One may argue that the question is not a real one, since during the first and the third episodes there is no real communication among all discussants. However, 'not-reacting-to-a-request' or 'not-participating-actively-for-a-while' are two behaviors that are inherent to synchronous discussions. Moreover, these behaviors are communicative in the sense that they are posted on a shared object, the argumentative map, and each of the participants intends to convey a message. Even in the third episode during which Fatima and Rim remained silent, they are not really out: in her auto-confrontation, Fatima makes clear that she wanted her silence to be posted on Ahmad's face! Fatima is 'in', attentive to Ahmad's moves and eager to convey her anger. So, what kept the group together in this weird communication?

We saw that the nature of tools for synchronous discussion enables to communicate differently from in f2f communication. The discussion map shows all previous moves, their authors, as well as the interlocutors to which they were directed. Discussants use the map to reason, as well as to communicate. For example, Ahmad selected two contributions by Rim, one recent and one remote, to point at some inconsistency. And this map is always present, even for participants who are silent. The growing map mediates the interpretations of the discussants in a way which is radically different from f2f settings. In f2f interaction, interpretations are updated at any moment and adjacent interventions influence more interlocutors than non-adjacent ones (Felton & Kuhn, 2001). In contrast, at the time an actor intervenes in synchronous discussions, he and his interlocutors can see the traces of a whole history. Their interpretations at any time of the discussion are sometimes mediated by previous misunderstandings inscribed in the map, and the discussants cannot instantaneously dissipate misunderstandings as easily as in face-to-face interactions through facial expressions or intonation. The big discrepancies in emotional states – satisfaction against anger, disappointment or boredom against interest, suggest that the map hosted unshared cohabiting misunderstandings. Each participant seemed isolated, the presence of the other being transient. This observation adds up to the analysis of the *cued retrospective reporting* to affirm that **the e-discussion developed without the establishment of any state of intersubjectivity** in a sense of coordination of contributions. In the absence of state of intersubjectivity, we considered different traditions that explain the propensity people have in participating in dialogues. According to a cognitive psychology perspective, dialogues demand a lot from discussants. However, people manage to participate in dialogues. Garrod and Pickering (2004) have proposed the idea of *interactive alignment* to explain this propensity: This is a process by which people align their representations at different linguistic levels at the same time. They do this by making use of each others' choices of words, sounds (in f2f communication),

grammatical forms, and meanings. Interactive alignment ensures that interlocutors operate on common representations. So in speaking, each partner generates his utterance on the basis of what he has just heard from the other and can leave out redundant information without the risk of misunderstanding. Similarly in listening, aligned representations at the levels of the situation model, semantic interpretation, and syntactic form enable the listener to fill in the gaps at these levels. Is the idea of interactive alignment applicable to synchronous e-discussions in general? Not exactly, but it is useful! But it provides some inspiration about for an alternative idea: Instead of aligning representations at different linguistic levels at the same time on the basis of what has just been heard, discussants interact with a growing map. This map has two contradictory characteristics. On the one hand, it changes instantly; messages arrive often at a hectic pace, sometimes simultaneously from different interlocutors. On the other hand, the map is stable; it mostly remains unchanged, with an accumulating history. The first characteristic seems to invite discussants to *align their representations*. However, discussants are not obliged to react to these messages. The second one leads discussants to rely on past persistent interpretations perpetuated by the map. Anyway, the high pace of communication gives the illusion to each discussant that he/she and his/her peers posted their beliefs about teaching, learning and moderating and interacted with them, and makes clear his/her (dis)agreements. However, what happened was a *co-alienation* – the juxtaposition of incompatible alignments of representations through a common external representation.

At that stage, it is premature to discuss the educational relevance of co-alienation and to decide whether its emergence is a priori welcome or should be avoided. But the question whether co-alienation is utterly bad or whether it can lead to learning, seems to us a wrong question. The right question is how people can learn from such communication. And indeed we can ask whether the discussion between Ahmad, Judith, Fatima and Rim lead to any kind of learning. On the one hand, the discussion in itself seemed quite shallow, scattered with persistent misunderstandings. However, Ahmad and Judith earned invaluable insights during this lopsided discussion. True, this is their cued retrospective that demonstrated clear gains but it is impossible to know whether this reflective activity revealed or promoted those gains. Anyway, synchronous discussions should be considered in their larger educational contexts. They rarely happen as isolated activities but rather belong to a series of activities. The cued retrospective was arranged in the present study for experimental purposes, but it resembles common educational settings in which synchronous discussions are reflected on. The precariousness of communication in synchronous discussions, the state of co-alienation we described, can be dangerous. However, it can serve, with appropriate activities, to improve interactions in discussions among learners.

References

- Cazden, C. B., (1995). Play with language and meta-linguistic awareness: one dimension of language experience. In J. S. Bruner (Ed.), *Play—Its Role in Development and Evolution* (pp. 603-608). Penguin, New York.
- Garrod, S., & Pickering, M. J. (2004). Why conversation is so easy. *Trends in Cognitive Science*, 8(1), 8–11.
- Herring, S. (2001). Computer-mediated discourse. In D. Schiffrin, D. Tannen, and H. Hamilton (Eds), *The Handbook of Discourse Analysis* (pp. 612-634). Oxford: Blackwell Publishers.
- Matusov, E. (1996). Intersubjectivity without agreement. *Mind, Culture, and Activity*, 3, 25-45.
- Matusov, E. (2001). Intersubjectivity as a way of informing teaching design for a community of learners classroom. *Teacher and Teaching Education*, 17, 383-402.
- Matusov, E. (2011). Irreconcilable differences in Vygotsky's and Bakhtin's approaches to the social and the individual: An educational perspective. *Culture & Psychology* 17(1) 99-119.
- Murray, D. E. (1989). When the Medium Determines Turns: Turn-taking in Computer Conversation. In H. Coleman (Ed.), *Working with Language* (pp. 251-266). Mouton de Gruyter, New York.
- Rommetveit, R. (1979). On the architecture of intersubjectivity. In R. Rommetveit & R. M. Blaker (Eds.), *Studies of language, thought, and verbal communication*. New York: Academic Press.
- Rommetveit, R. (1985). Language acquisition as increasing linguistic structuring of experience and symbolic behavior control. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives*. Cambridge: Cambridge University Press.
- Schwarz, B. B. & Asterhan, C. S. C. (2011). E-moderation of synchronous discussions in educational settings: A nascent practice. *Journal of the Learning Sciences*, 20(3), 395-442.
- Schwarz, B. B. & De Groot, R. (2007). Argumentation in a changing world. *The International Journal of Computer-Supported Collaborative Learning*, 2(2-3), 297-313.
- Van Gog, T., Kester, L., Nievelstein, F., Giesbers, B., & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior*, 25, 325-331.

Ethics for Design-Based Research on Online Social Networks

R. Benjamin Shapiro, Tufts University, Medford, MA, ben@cs.tufts.edu
Pilar N. Ossorio, University of Wisconsin-Madison, pnoosorio@wisc.edu

Abstract: Design-Based Research (DBR) allows learning scientists to investigate new processes, contexts, and technologies for learning. Social Networking Sites (SNSs) offer researchers rich new opportunities to create educational interventions that are deeply connected to learners' lives and relationships. We discuss legal and ethical challenges, and possible solutions to them, that face educational researchers as they begin to do DBR on SNSs. Addressing these issues will be crucial to design researchers wishing to use SNSs as sites for learning, and also offers an opportunity for the CSCL community to shape SNS research far beyond our field.

Introduction

Online social networking sites (SNSs) are powerful research tools because of their wide reach and deep connectivity to users' lives. In 2011, approximately 42% of U.S. adults belonged to a SNS (Hampton, Goulet, Ranie, & Purcell, 2011). In 2010, 73% of online U.S. teens used SNSs (Lenhart, Purcell, Smith, & Zickuhr, 2010). Facebook had over 1 billion users in October of 2012; one twelfth of the entire world's population now uses Facebook's mobile apps (Facebook, 2012), highlighting the enormous connectedness of social media to users' daily routines. SNS-based research offers learning scientists the opportunity both to understand human social activity, and to use SNSs for experimental interventions, such as increasing civic participation (Bond et al., 2012) or teaching about science (Shapiro, Squire, and ERIA, 2011). While observational studies offer researchers the ability to analyze activity – including learning – as it is already occurring, Design-Based Research (DBR) enables researchers to support and study new kinds of educational interactions (Brown, 1992; Collins, 1992). Though the number of studies conducted in SNS environments is growing, few DBR studies have been conducted to date. Instead the research has largely focused on publicly available data about extant activities.

Doing DBR on social networks offers researchers new opportunities to connect to learners' lives, and to understand how learning happens across levels of *space, time, and scale*. For example, a recent design experiment conducted by political scientists collaborating with Facebook staff reached over 60 million people and led to over 300,000 more of them voting in the 2010 United States elections (Bond et al., 2012). With SNSs we can build learning experiences that are deeply intertwined with learners' personal interests, and those of their friends and loved ones, such as by using information gleaned from users' posts on their Facebook feeds, or the content of web pages they've "liked." We can virally scale participation in learning environments by using the social sharing mechanisms that are central to SNSs, encouraging learners to invite their friends to learn along with them. We might use slow-moving interaction designs similar to social games like *FarmVille* to immerse learners in distributed embedded phenomena (Moher, 2006) that stretch over long periods of time and that differ depending upon users' physical locations. All of these possibilities highlight new opportunities for researchers to support and to understand learning at both individual and collective scales, over differing time scales, and to inform and study these experiences using new kinds of data.

Standard methodologies, tools, and norms for doing DBR on SNSs have not yet emerged, and we currently lack ethical frameworks for working with the unprecedentedly private data that SNS DBR makes available to researchers. Observational studies are already pushing boundaries; DBR will push further, such as by offering access to more information, exposure of information to, and about, peers. As learning scientists begin to develop DBR programs for SNS, we must also begin to develop a legal and ethical groundwork for doing our work in an appropriate manner. This groundwork can ultimately inform not just study design but the design of CSCL tools as well. Fields beyond the learning sciences are beginning to consider these issues as well (Introne, et al., 2013), and educational researchers have an opportunity to shape both the policy and the research practice landscape of SNS DBR.

The authors of this paper are, respectively, a learning scientist and a legal scholar specializing in bioethics who is a member of her university's IRB and the U.S. Secretary of Health and Human Services Advisory Committee on Human Research Protections (the analysis in this paper is our own and is not an opinion of any committee or governmental agency). We have begun working together to lay the necessary legal and ethical groundwork for DBR on SNSs. This paper illustrates the possibilities for DBR on SNSs using an example drawn from our own work, then discusses some of the legal and ethical considerations that educational researchers, as well as Federal regulators and university IRBs, must wrestle with for work like it to precede. Our *legal* analysis is grounded in study of US regulations including the Code of Federal Regulations, the Federal Policy for the Protection of Human Subjects (the "Common Rule"), as well as case law, though the *ethical* issues we raise are globally applicable.

An Illustrative Scenario

Consider the following vignette describing an online science education game:

Katrina started playing Anatomy Pro-Am (APA) after receiving an invitation from her Facebook Friend Riley. Now, they play APA together daily. This afternoon, they work on the case of Mr. Badger, whose primary care physician suspects he may have liver cancer. They look for abnormalities in his CT scans. At first, they work independently. Katrina examines the images and decides everything looks OK. She clicks Looks Healthy. When she does, Riley's work appears. Katrina is surprised to see that Riley labeled some spots as cancerous.

Their friend Marcus comes online and they invite him to give a third opinion. He agrees with Riley, but Katrina thinks those spots might be fat in the liver. To settle the argument, they search the APA Almanac and Google Images for reference images of fatty and cancerous livers. As it turns out Mr. Badger's liver looks more fatty than cancerous. Both Riley and Marcus change their diagnoses, and the team saves its work. The players then receive feedback comparing their judgments to those of experts.

Finally, the team's work is sent to McLuhan Hospital, where Mr. Badger is a patient. Their work, along with that of hundreds of others, will be used to generate a representation that will guide McLuhan's radiologist in fine-tuning her analysis of Mr. Badger's case.

APA is a real project (see Figure 1) in which we wrestle with ethically creating and studying an online learning environment that is deeply integrated with social networking. We hypothesize that game players will learn scientific content, increase their feelings of self-efficacy, and increase their interest in scientific careers. In a pilot study, we found that middle-school girls who played a precursor to APA became significantly more likely to believe they could become physicians, as well as better understand the goals, tools, and challenges of cancer medicine (Shapiro, 2011). Moreover, we hypothesize that APA, and projects like it, can demonstrate new ways for the public to participate in science and medicine. In the above vignette, it is through providing crowd-sourced support for medical diagnosis that players both learn science and participate in medicine.



Figure 1. Anatomy Pro-Am.

Third-party Facebook applications (e.g., games) utilize programmatic interfaces to Facebook data, known as APIs, to obtain information about SNS users and their “Friends,” and often to act on a user’s behalf (for instance, by posting information to her profile). Thus far, commercial software developers have been the primary users of these APIs, but they also enable researchers to create rich, interactive experiences for study. However, unlike commercial software developers, academic researchers are subject to regulations governing human subjects research, as well as ethical guidelines about how to conduct research. These legal and ethical frameworks shape the kinds of studies we can conduct, as well as the manner in which we conduct them. As yet, however, we lack standards specific to SNS SBR. For example, in the above vignette, is it acceptable for Marcus (a teenager) to participate in the educational intervention without his mother’s consent? The standard educational research approach is to seek parental permission for studies of minors’ learning. But as we describe below, this may significantly impair SNS DBR, offer little actual protection to minors, as well as impose a burden on researchers and participants that exceeds the legal requirements that govern SNS DBR. Or,

considering the issue of privacy, would it be appropriate for us to use Facebook APIs to crawl through and save a complete record of all of Katrina's interactions across Facebook so that we can build a stronger model of relationships between player profiles and their learning through game play? Both of these possibilities are ones in which commercial developers routinely engage, but both are outside the bounds of what most researchers would find appropriate or routine in traditional (non-SNS) DBR.

Here, we address three ethical, regulatory, and technical challenges raised by SNS-embedded DBR: (i) whether adolescents who participate in such research through commercial portals, such as Facebook, should be categorized as children for regulatory purposes; (ii) the extent to which researchers may collect data about SNS participants and their Facebook "friends"; (iii) how CSCL researchers might construct their technical systems so as to maximally protect participants while minimally hindering legitimate research.

Consent and the Adolescent Player

University researchers studying players of an SNS-based educational intervention, such as a game like APA, are engaged in human subjects research and require either Institutional Review Board (IRB) approval or exemption, and players' informed consent/assent (U.S. Department of Health and Human Services, 1991). A surprising question is whether parental permission is required for researchers to study adolescent players (adults give consent for their own research participation; parents give permission, rather than consent, for their children's research participation) (U.S. Department of Health and Human Services, 1983).

Design researchers have good reason to want their experimental SNS applications to be accessible to as many learners as possible. SNSs offer excellent viral recruitment potential, and most Facebook applications enable users to recruit their Friends to participate. These mechanisms are attractive to researchers because they both enable greater scale of participation and because they allow research on learning interventions to be ecologically valid. Research on how social sharing mechanisms can drive participation can increase ecological validity if those mechanisms mirror those used by commercial applications. From this perspective, it is desirable that adolescents participating in our DBR can invite their Friends to participate too, and that those Friends can accept those invitations and participate immediately. Other commentators on SNS research have considered adolescents to be children, and have stated that parental permission will usually be necessary for such research (Bull et al., 2011; Moreno, Fost, & Christakis, 2008). This would considerably damp the participation growth curves of our applications compared to commercial applications.

What The Federal Regulations Say

The Common Rule defines children as: "Persons who have not attained the legal age for consent to treatments or procedures involved in the research, under the applicable law of the jurisdiction in which the research will be conducted." This definition coordinates research regulation with other laws, such as laws permitting adolescents to consent on their own to treatments for sexually transmitted diseases. The regulatory definition of a child permits reasonable research on the effectiveness of those interventions, such as STD treatments, to which adolescents can consent. The regulations are clear that so long as an adolescent can consent to an intervention, she can participate in research about that intervention without parental permission.

The federal Children's Online Privacy Protection Act (COPPA) defines children as *persons under the age of thirteen* (15 U.S.C § 6501 *et seq.*, 1998). COPPA prohibits commercial operators of online services (such as Facebook) from engaging in unfair or deceptive practices with respect to the collection, use, or disclosure of personal information from and about children (i.e., those 12 and younger) on the Internet. Congress deliberated extensively before determining that 13 - 17 year olds ought to be treated the same as adults in the online context.

Facebook and other SNSs have responded to COPPA by prohibiting access to their services by people under thirteen. They are not legally required to ban users 12 and younger, but do so in order to avoid the costs of complying with COPPA's requirements that information collection and sharing practices be disclosed to, and parental permission obtained from, the parents of young users. Although some parents help their younger children to evade these age limits, SNSs take measures to restrict underage access, including deleting accounts identified as belonging to underage users (boyd & Hargittai, 2011).

In light of COPPA and the Common Rule's provisions, we believe IRBs should allow adolescents who enter an experimental game through a commercial website that limits access to people thirteen or older to consent on their own to research on participation and learning in that intervention. The existing commercial context of the research, in which 13 year olds have, according to COPPA, obtained legal age to consent to participation, including to data collection, triggers the common rule provisions for treating these adolescents as adults for the purposes of IRB review. In the non-research context, adolescents can legally use SNSs to play games and to provide identifiable, private information about themselves to commercial application developers, and so they should be able to consent in the research context. Parental permission should not be required. This approach is not a waiver of the requirement for parental permission. Rather, we argue that people thirteen and older should not be categorized as children for IRB review of SNS DBR. Researchers will nonetheless need to ensure that they take all possible steps to articulate to adolescent participants how their data will be used and the risks of participation.

Other education and social media researchers with whom we have spoken about doing SNS research report that their universities' IRBs have required them to obtain parental permission for all minors. This requirement contradicts U.S. Federal regulations governing research. While the regulations do permit local IRBs to impose additional protections for participants, we suggest that researchers might use the argument presented here to negotiate with their IRBs for greater latitude to conduct ecologically valid DBR.

What About Parents?

Of course, parents may control their adolescents' online activities. Parents who monitor their adolescents' Internet activities can observe the consent process and read the information provided, and parents are always free to prevent their adolescents' participation in IRB-approved research. IRBs may also impose additional protections for vulnerable populations (45 U.S.C § 46.111b, 2005).

Furthermore, we believe that researchers should make use of the online environment to deliver innovative, truly informative consent processes for *anybody* participating in online research. This is especially important given the general public's substantial ignorance about data collection over the Internet. For instance, while 44% of American parents "are extremely or very concerned that their children might have information about them used for targeted advertising," only 9% of parents whose children use SNSs believe their children's data have been used in this manner (boyd & Hargittai, 2011). In reality, 100% of SNS participants have their data mined for targeted advertising. In light of these data, the traditional approach of IRBs, to assume that parents will be able to weigh the risks of participation in research for their children, seems grossly insufficient to protect youth. Furthermore, many youths interact with social media from across a range of locations and devices, making parental mediation of activity extremely difficult (Yardi & Bruckman, 2011).

Given these and other findings, it is important that online researchers take advantage of their medium to create high quality consent processes that explain how data will be analyzed and clearly inform participants of possible risks. Because participants may only understand the full risks of participation after the fact, researchers should create easily accessible mechanisms for post-hoc withdrawal from participation in research. Allowing participants to delete data about themselves after they have provided it creates additional opportunities for participant education, such as through discussion between adolescents and teachers, parents, and peers, to shape informed consent. Many commercial application developers provide less than straightforward tools for accessing and controlling the information that they maintain; academic researchers could use the more stringent ethical criteria of our field to create and demonstrate higher standards for participant protection.

Scaling from Individuals to Networks

The APA vignette also raises regulatory and ethical questions about collecting information on members of participants' social networks (i.e., their Facebook "Friends"). An important research question for environments like APA is whether they can help overcome race and gender disparities in science. Understanding the demographics of player networks will help answer this question. For instance, to learn how biases shape APA participation, one might examine the demographic characteristics of Facebook Friends whom our main player Katrina invites to work on particularly hard problems vs. easy problems vs. the demographics of her network generally. The average Facebook user has 190 Friends (Backstrom, 2011); it is not plausible that all would consent to data collection. Requiring consent from non-players in order to characterize the overall characteristics of Katrina's network would introduce statistical bias and diminish such a study's rigor.

Facebook APIs offer researchers access to a great deal of identifiable information about game players and their Friends, much of which is necessary for providing game user interfaces. Should design researchers who use Facebook APIs be required to not store identifiable information on non-playing Friends? Should researchers be prohibited to access this identifiable Friend information at all, should they be permitted to access the information only to save it in non-identifiable form, or should they have unrestricted access to it under the guise of Facebook being a public space? In considering this question, we also must figure out how researchers should weigh the difference between Facebook's legal status as a public space (see below) and many users' expectation that it is semi-private.

Perhaps the closest already-understood analogy involves collection of family history information from research participants. Although researchers have engaged in this practice for decades, it became controversial in 2001: regulators temporarily halted all human subjects research at Virginia Commonwealth University after a research participant's father objected to the collection of sensitive family history information (Bolkin, 2001). Commentators and regulators emphasized that researchers who collect identifiable private information about relatives are engaged in human subjects research on the relatives and must obtain their consent, unless an IRB waives this requirement (U.S. Department of Health and Human Services, 1991). However, if the information is not identifiable, or is identifiable but not private, then the researcher is not engaged in regulated human subjects research on relatives, and their consent is not required. Such unregulated research has been called "human non-subjects research" (Brothers & Clayton, 2010).

Under the regulations, information is private if a person reasonably expects it will not be observed or recorded, or if it is provided for a specific purpose and the provider reasonably expects it will not be made

public (U.S. Department of Health and Human Services, 1991). Social conventions regarding information privacy on SNSs are still developing. Facebook warns users that some information is “always publicly available,” including one’s name, profile picture, network, username and Unique Identifier (UID), and any additional information one chooses to make public (Facebook, 2012). To date, state and federal courts that have addressed the question have held that individuals do not have a reasonable expectation of privacy in information posted to a SNS (Newell, 2011); however, the issue is new enough that courts have not yet considered many aspects of online privacy.

The emerging legal consensus – if not that of the social media research community – is that much information on SNSs is not private, so for now IRBs will often be justified in treating the collection of identifiable information about Facebook Friends as human non-subjects research. Even in the absence of clear regulation, however, investigators have ethical duties to minimize risks to people whose data they use. Even data that may be innocuous in its raw form may be embarrassing when crystalized into accountable facts (e.g., that someone exhibits racial bias in his/her daily choices; we describe a hypothetical SNS-based analysis of this below). A primary way to reduce risk to research participants is limiting their identifiability as research participants, while simultaneously mitigating the consequences to them of such identifiability. A standard practice for doing so is to anonymize collected data as early in the research process as possible.

Reidentification Risk

Though anonymization can be a way to protect the identities of participants and their associates, the challenge of successfully anonymizing data, while still permitting useful research, is surprisingly difficult (Ohm, 2010). Even seemingly sanitized datasets can be de-anonymized, such as by using network structure to re-identify individuals (Narayanan & Schmatikov, 2009), or statistical inference combined with non-random identifier generation algorithms to predict social security numbers (Acquisti & Gross, 2009).

Perhaps the most notable (or notorious) case of re-identifiability of an *academic* SNS dataset thus far is the Tastes, Ties, and Time (T3) project (Lewis et al., 2008). Researchers at Harvard and University of California, Los Angeles downloaded the Facebook profiles of freshmen at a “diverse private college in the Northeast U.S.” for four years, and combined those data with housing records obtained from that university in order to study relationships between online participation and physical space. This longitudinal dataset offered deep insights into the social behavior of four cohorts of undergraduates across different levels of *space, time, and scale*. The research was IRB-approved, and the data released with assurances that “all identifying information was deleted or encoded” (ibid.).

However, nearly immediately after the data were released, other researchers quickly identified the data as belonging to Harvard College students (Zimmer, 2010); some students could be easily identified based upon the sparseness of the space of individual characteristics present in the dataset (students from state X, majoring in Y, interested in Z). For example, only one student was present in the dataset for a given year with the home state of Mississippi. It would be easy to combine this data with publicly available information to identify many of the individuals present in the data set. Thus even though researchers made good faith efforts to protect participants (by removing their names from the dataset), the information that was released was enough to reveal participants’ identities and reunite it with the large amount of personal data that researchers harvested.

New Identifiability Challenges for Design Researchers

Design-based research raises new issues that projects like T3 have not yet contended with. T3, like almost all other SNS-based research to date, is exclusively observational in nature. Researchers collected data on already unfolding online activities. But design-based research on SNSs goes a step further, to providing the context for new activities that are situated in the existing context of sites like Facebook. This raises new challenges because it pits two important aims – both of urgent importance to design researchers – against each other: the needs of being an experience provider and facilitator, and the de-identification needs of ethical research. Most Facebook games (including our own) are long-running socially connected experiences. They are microworlds where users can log in repeatedly over time and have a stable profile, connected to their Facebook profiles and Friend networks, that follows them from session to session. Designers facilitate the social elements of these experiences for users by disclosing information to their peers (Friends) about their participation. For example, in the above vignette, Riley, Katrina, and Marcus are all informed of their friends’ participation in the APA game. In order to do any of this, application developers must maintain a considerable amount of identifying information about participants. If Katrina’s game play history were totally disassociated with her identity, then she would be required to start anew each time she played, which would significantly alter the nature of the educational project. Similarly, if users could not see which of their friends play the game, including perhaps their competencies at different skills the game demands, then opportunities for collaborative learning (such as through creating teams of complementarily skilled players) would be substantially diminished. Insofar as peer-driven discovery is the major means by which Facebook applications grow their user bases, this would limit the potential of an educational project to succeed, as well as the generalizability of the research. Ultimately the need for collecting and (sometimes) disclosing identifiable information is irreconcilable with the protective heuristic of anonymizing data as early in the research process as possible.

Technical Guidelines

The thicket of issues raised above is primarily ethical and legal in nature. We cannot find salvation from the described tensions among important values in technical solutions alone. However, it may be possible to architect systems for design-based research on SNSs in ways that reduce risk to participants and their SNS Friends.

In a typical Facebook game scenario, the user plays a game in his or her web browser (and increasingly on mobile devices). The game appears to be a part of Facebook, but resides on a different server (Figure 1 illustrates how a game hosted at a university appears seamlessly integrated with the rest of the Facebook interface). When the user first navigates to the application from a link on Facebook, Facebook's servers, the user's browser, and the application servers exchange information about the user. Applications can request, among other things, access to the user's profile, the ability to send messages on behalf of the user, to see Friends' information, and the ability to do all of this and more when the user is not even logged in. If the user has not previously agreed to the access that the application requests, then Facebook asks the user for permission. If the user agrees, access keys are sent from Facebook, via the user's browser, to the application developer. These keys are used in subsequent Facebook requests to authenticate access to users' profiles. The user may subsequently revoke these tokens using through Facebook's.

Once an application has these keys, it can use them to interact with the user's profile via the Facebook APIs in two distinct ways, each with strikingly different implications for users' privacy: The application's servers may use these APIs to directly obtain information from or post information to Facebook. Or, the application may execute JavaScript in the user's browser that interacts directly with Facebook via these APIs to obtain the information, and then parlay that information back to the university-based application servers. The former approach enables application developers to shift as much of the computational burden away from the user's browser as possible, as well as to maximize the amount of information that application developers have access to. The latter approach limits the demands on application developers' servers, may have worse performance characteristics from the user's perspective, but also has some interesting potential uses for increasing users' and users' Friends' privacy.

Consider the APA vignette above and suppose we wish to understand how racial bias creeps in to Katrina's decisions about whom to invite to collaborate with her. To study this, we would want information about the demographics of Katrina's network overall, such as the race of each of her Friends, as well as about the people Katrina specifically chooses to invite to work with her. We might be interested in comparing her choices in game play with her general choices about whom to interact with, and so also grab race data about the people Katrina chats with or is tagged in photos with. In the end, our analysis does not require these raw data, only aggregate statistics about Katrina's network and her collaborators (e.g., that 10% of her Friends are African American, that she is just as likely to chat with African Americans as other Friends, but far less likely to ask an African American to help her on a difficult game challenge).

We could obtain these aggregate data using either the server- or the browser-based approaches. In the server-based approach, our university-based servers would use Facebook APIs to crawl through Katrina's social network, retrieving and accumulating information about each Friend, her messaging history, etc., eventually distilling this information down to the necessary aggregates. In the browser-based approach, our code, running in Katrina's browser, would do much the same work. The difference is that in the latter scenario the raw data need never exist on our university servers. Only the aggregated information would be sent to us. This is strongly preferable from the standpoint of protecting SNS members' privacy, as it allows researchers to ask questions about participants who are situated within their networks without requiring researchers to ever see raw information about Friends that could be considered private.

This approach also has the benefit of tying data collection to explicit actions that users take to participate in research. A server-based approach allows researchers to harvest data about users and the Friends at will, with no active involvement by users. Researchers could periodically harvest information from all users' profiles and networks without users knowing about this ongoing activity. In contrast, if data collection code runs in the user's browser, it will only be active when the user has actively chosen to use our research-driven application. Unlike traditional research, when the event of participating in data collection is explicit and signaled to subjects by their unusual interaction with researchers, online research (such as by Facebook APIs) permits ongoing data collection about users once they have agreed to participate. A server-driven data-collection approach permits data collection months or even years after a user has consented, when he or she may not even remember doing so or even be aware that it is continuing. The browser-based approach requires explicit action by the user to re-enter the research space, and so permits the user to make an explicit, conscious choice about continuing to provide data to researchers. We believe that this is ethically preferable. Of course, this approach is only suitable to some research questions and methods, but exploring software architectures like this is a first step toward building systems that enable design-based research while maximizing participant protections. We hope other researchers will explore the space of possible technical systems designs that permit SNS-based design research while maximizing participant protections.

Conclusions

We have described several of the challenging questions facing design-based educational researchers hoping to utilize social networks like Facebook to create the learning environments of the future. A major promise of social media is peer-driven discovery, and we are excited about the prospects of creating learning experiences so compelling that learners voluntarily share them with their friends. Traditional IRB processes, particularly the requirement of parental permission that educational researchers usually obtain, would seem to preclude the possibility of viral growth in participation (and learning). But our analysis of the legal regulations governing commercial and academic work in this area supports the surprising conclusion that adolescents thirteen and older should be treated as adults for the purposes of research consent.

We examined the question of whether information about consenting users' peers (i.e. their Facebook Friends) should be accessible to researchers. Though the legal landscape in this area is changing quickly, there is currently a basis to permit researchers to use information about participants' broad networks without the explicit consent of others in the networks. Nonetheless, there is a need for researchers to develop ethical standards that are more stringent than the law alone requires, and to devise such standards in ways that both permit promising research and protect online users' privacy.

IRBs have been criticized for their inconsistent assessment of social, psychological, and economic harms and benefits in other research contexts (Department of Health and Human Services, 2011), and they will undoubtedly need guidance for assessing SNS research. Design-based research raises new ethical issues regarding risk minimization. IRBs must be educated regarding the array of technical options for addressing ethical issues. Further, funding agencies such as NSF and NIH might issue calls for research on building open source participant-protective cyberinfrastructure for SNS DBR, as well as work with HHS to encourage IRBs to consider the use of such infrastructure as a means of participant protection when evaluating research designs.

OHRP recently circulated a proposal to revise the Common Rule (Department of Health and Human Services, 2011). The proposal mentions the Internet as an emerging experimental sphere; however, it contains no discussion of how research in this sphere ought to influence regulatory revisions, or how such research would be reviewed under the new proposals. In the revisions, OHRP takes the position that all informational risk in research can be minimized by adequate data security precautions. While we believe data security is important, we suggest that appropriate research design also affects the risk calculus of SNS research. As the T3 example above highlights, even anonymized observational data bears identification risks for participants. We do not yet know how to balance the technical needs of creating persistent online experiences with the ethical need to minimize participant identifiability.

Lack of ethical guidance can stymie academic SNS research, potentially rendering academia irrelevant to an important and growing domain of online activity, and educational researchers unable to utilize SNSs to create and study better learning environments. The private sector is charging ahead, creating de facto standards for data use. Privacy invasion, or the imposition of risk, is not made acceptable because somebody else is already doing it. However, the goals of most academic research surely have as much social value as the goal of selling more products to SNS users. Permitting marketers greater access than academic researchers to peoples' online information is a dubious ethical outcome. CSSL researchers who figure out how to productively navigate the complex array of issues raised in this paper will not only impact the kinds of technologies and contexts that we can create to support collaborative learning, but have an impact far beyond educational research, impacting fields like public health and political science, as well as shaping research policy for the 21st century.

References

- Acquisti, A. & Gross, R. (2009). Predicting social security numbers from public data. *PNAS* 106(27). 10975–10980.
- Backstrom, L. (2011) Anatomy of Facebook. Webpage of the Facebook Data Science team. https://www.facebook.com/note.php?note_id=10150388519243859 Accessed October 28, 2012.
- Barab, S., et al. (2009) Transformational Play and Virtual Worlds: Worked Examples from The Quest Atlantis Project. *IJLM: International Journal of Learning and Media*, 1(2), published online.
- Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012) A 61-million-person experiment in social influence and political mobilization. *Nature*, 489, 295-298.
- Botkin, J. R. (2001) Protecting the privacy of family members in survey and pedigree research. *JAMA* 285, 207-211.
- Boyd, D. & Hargittai, E. Facebook privacy settings: Who cares? *First Monday*, 15 (8), published online.
- Brothers, K. B., & Clayton, E. W. (2010) Human Non-Subjects Research: Privacy and Compliance. *American Journal of Bioethics* 10(9), 15-17.
- 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.
- Bull, S. S., et al., (2011) The Cold Pressor Task: Is it an Ethically Acceptable Pain Research Method in Children? *Journal of Pediatric Psychology*, 36(10), 1071-1081.

- Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15–22). New York: Springer-Verlag.
- Department of Health and Human Services, Additional Protections for Children Involved as Subjects in Research. Federal Register 48, 9814 (March 8, 1983); codified at 45 Code of Federal Regulations 46, Subpart D.
- Department of Health and Human Services, Federal Policy for the Protection of Human Subjects. Federal Register 56, 28003 (June 18, 1991); codified at 45 Code of Federal Regulations 46.
- Department of Health and Human Services, Human Subjects Research Protections: Enhancing Protections for Research Subjects and Reducing Burden, Delay, and Ambiguity for Investigators. Federal Register 76, 44512 (July 26, 2011).
- Facebook, Information We Receive and How it is Used, <http://www.facebook.com/about/privacy/your-info-everyoneinfo>, visited January 5, 2012.
- Facebook (2012). One billion – key metrics. <http://newsroom.fb.com/imagelibrary/downloadmedia.ashx?MediaDetailsID=4227> Accessed online October 28, 2012.
- Hampton, K. N., Goulet, L. S., Rainie, L., & Purcell, K. Social networking sites and our lives. Pew Research Center's Internet and American Life Projects, Internet and American Life. Pew Research Center, Washington, DC, 2011.
- Introne, J., Levy, K., Munson, S., Goggins, S., Wash, R., Aragon, C. (2013) Design, influence, and social technologies: techniques, impacts, and ethics. Workshop at the 2013 ACM Conference on Computer Supported Cooperative Work. <http://distworkshop.wordpress.com/>
- Jianquiang, D. S., et al. (2011) Farmer's tale: a facebook game to promote volunteerism. CHI 2011, 581-584.
- Lenhart, A., Purcell, K., Smith, A., & Zickuhr, K. (2010) Social Media & Mobile Internet Use Among Teens and Young Adults. Millennials: A Portrait of Generation Next Pew Research Center's Internet and American Life Projects, Internet and American Life (Pew Research Center Washington, D.C., 2010).
- Lewis, K., et. Al. (2008) Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks*, 30, 330-342.
- Moher, T. (2006). Embedded phenomena: supporting science learning with classroom-size distributed simulations. CHI '06 Proceedings of the SIGCHI conference on human factors in computing systems.
- Moreno, M. A., Fost, N. C., & Christakis, D. A. (2008) Research ethics in the MySpace era. *Pediatrics*, 121(1), 157-161.
- Narayanan, A., & Shmatikov, V. (2009). De-anonymizing social networks. Paper presented at the 30th IEEE Symposium on Security and Privacy.
- National Research Council, Learning Science Through Computer Games and Simulations. Committee on Science Learning: Computer Games, Simulations, and Education. M. A. Honey, M. L. Hlton, Eds. (The National Academies Press, Washington, DC, 2011).
- Newell, B. C. (2011) Rethinking Reasonable Expectations of Privacy in Online Social Networks. *Richmond Journal of Law and Technology* 17, 12.
- Ohm, P. (2010) Broken promises of privacy: responding to the surprising failure of anonymization. *UCLA Law Review*, 57, 1701-1777.
- Shapiro, R. B. (2011). Anatomy Pro-Am. <http://benshapi.ro/projects/anatomy-pro-am/>
- Shapiro, R. B., Squire, K., and the Educational Research Integration Area (2011). Games for Participatory Science. *Educational Technology*.
- Steinkuehler, C., Duncan, S. (2008) Scientific Habits of Mind in Virtual Worlds. *Journal of Science Education and Technology*, 17(6), 530-543.
- Yardi, S. & Bruckman, A. (2011). Social and technical challenges in parenting teens' social media use. In Proceedings of ACM Conference on Human Factors in Computing Systems (CHI'11). Vancouver, Canada. May 7-12, 2011.
- Zimmer, M. (2010) But the data is already public: on the ethics of research in Facebook. *Ethics and Information Technology* 12(4), 313-325.

Acknowledgments

This work was funded by NSF DRL119383 and IIS1227530, support from the University of Wisconsin-Madison Graduate School, the Morgridge Institute for Research, and the Wisconsin Institute for Discovery.

Understanding Collaborative Program Comprehension: Interlacing Gaze and Dialogues

Kshitij Sharma, Patrick Jermann, Marc-Antoine Nüssli, Pierre Dillenbourg
CRAFT, École Polytechnique Fédérale de Lausanne, Switzerland
{kshitij.sharma, patrick.jermann, marc-antoine.nuessli, pierre.dillenbourg}@epfl.ch

Abstract: We study the interaction of participants in a pair program comprehension task across different time scales in a dual eye-tracking setup. We identify four layers of interaction episodes at different time scales. Each layer spans across the whole interaction. The present study concerns the relationship between different layers at different time scales. The first and third layers are based on the utterances of the participants while the second and fourth layers are based on participants' gaze.

Introduction

In Computer Supported Collaborative Learning (CSCL), one main open challenge is to use technology to measure the dynamics of interaction. We report recent developments in eye-tracking which show how gaze can be used to reflect cognitive and collaborative processes at various time scales. Thereby we scale up the social unit of analysis from individual to pair and scale down the temporal unit of analysis from the whole interaction to shorter interaction episodes.

With respect to the social unit of analysis, gaze has traditionally been used to assess individual cognition (e.g. eye-tracking studies about reading, program comprehension, etc.). However, in the context of CSCL, a methodology is needed to describe collaborative gaze. Various measures of "gaze togetherness" have been used to indicate the quality of collaboration in dyadic interaction. In general, good collaboration features convergent gaze. Gaze togetherness increases significantly especially during verbal and deictic references. These measures of togetherness are however related to a global time scale and don't consider the evolution of gaze focus during interaction.

With regards to the temporal granularity of analyses, studies have emphasized upon overall measures of individual attention. For example, studies have reported the proportion of time that subjects spent fixating different parts of the interface (Romero et al., 2002; Bednarik & Tukiainen, 2006; Sharif & Maletic, 2010; Hejmady & Narayanan, 2012; Pietinen et al., 2008; Pietinen et al., 2010;). These measures indicate overall gaze behavior (and may be correlated with expertise) but they cannot serve as real-time indicators of collaboration that could be used to provide immediate feedback. In the context of CSCL, the dynamics of interaction and dialogue are important indicators for collaborative knowledge building (e.g. Stahl, 2000). New gaze indicators are needed to reflect the knowledge building at the micro level.

Time scales have been used to describe behavior at various levels. Eye-trackers allow us to capture attention at a time scale that has more information content than the other measures like interface event logs, dialogues or gestures. In a controlled experiment (Lord & Levy, 2008) the duration of eye-fixations are of the order of 100 milliseconds, which gives them a place at the lower end of cognitive behavioral band (Newell, 1994). Cognitive behavioral bands have complex actions (e.g., reading or gestures) at the higher end. Anderson (2002) identifies cognitive modeling as bridging across the behavioral bands by taking the lower level bands into account. We will reuse the levels by Anderson (2002) to refer to the Task (where we usually measure understanding), Unit task (where we usually code dialogues) and Operations (where we usually collect raw data).

Through this contribution, we address both the social and temporal mismatch of current gaze methodology with the study of collaborative interaction. We propose a method to detect interaction episodes based on both gaze togetherness and stability and show that these measures are related to the level of understanding that a pair achieves at the end of the task. To support our proposal, we present a dual eye-tracking study in a remote pair program comprehension scenario.

The remainder of the paper is organized as follows: the second section gives the related work for the present study. The third section describes the main features of the study and the research questions. The fourth section describes the experiment and the various variables. The fifth section presents the analysis results. Finally the sixth section discusses the results and concludes the paper.

Related Work

Adaptive Support for CSCL

Adaptive CSCL has been around for about 10 years. Jermann et al. (2001) proposed a feedback model for collaborative interaction regulation. The regulation is based on the collection of collaborative indicators that are assessed by the system or by the human learners and teachers. More recently, Magnisalis (2011) propose that

web 2.0 and artificial intelligence are increasingly used to design reactive systems and that learners benefit from the adaptation of the systems.

Scaling Up The Social Unit

There are different gaze-based measures of collaboration given by Richardson & Dale (2005), Cherubini et al. (2008) and Pietinen et al. (2010). Richardson & Dale (2005) used “gaze togetherness” as a notion of gaze cross recurrence (how much the participants are looking at the same object at the same time). Cherubini et al. (2008) used eye tracking in a remote collaborative problem solving setup to detect the misunderstanding (distance between the referrer’s and the partner’s gaze points) between the collaborating (through chat) partners. Pietinen et al. (2010) gave a new metric to measure joint visual attention in a co-located pair programming setup, using the number of overlapping fixations and the fixation duration of overlapping fixation for assessing the quality of collaboration. The problem of these measures is that they characterize togetherness on a global temporal level or on an arbitrarily defined timespan (one could partition the interaction into “n” parts but these would not reflect the underlying interactive dynamics).

Linking Gaze and Speech

At the level of operations, there are studies about gaze and speech coupling (Mayer et al., 1998; Griffin & Bock, 2000; Zelsinky & Murphy, 2000). There are different notions for eye-voice span given in different studies but all the notions point towards a strong coupling between speaker’s gaze and speech. Allopenna et al. (1998) showed that the mean delay between hearing a verbal reference and looking at the object of reference (the listeners’ voice-eye span) is between 500 and 1000 milliseconds. The combination of eye-voice and voice-eye coupling is that the gaze of speakers and listeners are coupled with a lag of about 2000 milliseconds. This short term coupling between speaker and listener is at the operation level only and does not inform about the relationship of gaze and dialogue in longer episodes. This is problematic when one is interested in knowledge building episodes that usually consist of several utterances.

Linking Dialogue and Understanding

Concerning the relationship between dialogues and understanding, there is a long-standing tradition of research in CSSL. For example, the elaborated explanations (Cohen, 1994; Webb, 1989) were shown to be beneficial for learning. In the field of tutoring, research has shown that dialogue moves of tutors depend on their assessment of the tutee (Eugenio et al., 2009; Chi et al., 2008; Chi & Roy, 2010) and that they can predict better understanding of the tutee (D’Mello et al., 2010). What is missing is a gaze indicator at the same temporal level as dialogues.

The Domain: Program Comprehension and Eye Tracking

There have been studies (in the past) concerning eye-tracking and programming. Romero et al. (2002) compared the use of different program representation modalities (propositional and diagrammatic) in a debugging study where experts had a balanced shift of focus among the different modalities. Sharif et al. (2012) emphasized the importance on code scan time in a debugging task and conclude that experts perform better and have shorter code scan time than novices. Bednarik & Tukiainen (2006) examined coordination of different program representations in a program understanding task where experts concentrated more on the source code rather than looking at the other representations. Hejmady & Narayanan (2012) compared the gaze shift between different Areas of Interest (AOI) in a debugging Integrated Development Environment (IDE) and concluded that good debuggers were switching between code and the expression evaluation and the variable window rather than code and control structure and the data structure window.

Present Study and Questions

The present dual eye-tracking study examines the relationship between gaze, speech and performance in spatially distributed (remote) pair programming. We chose remote pair programming so that we can have two synchronized streams of eye-tracking data, which is difficult in the co-located pair programming (both programmers looking at the same screen). Baheti & Williams (2002) have shown that pair programming can be conducted remotely without negative effects on performance. We use two synchronized eye-trackers to study the gaze of two persons who have to read, understand, and explain functionality of a JAVA program.

Methodological Question

The present study identifies different time scales to characterize interaction. Our working hypothesis is that it is necessary to define a gaze measure at each level to reflect corresponding cognitive processes. Indeed, measuring gaze at a global task level does not inform about dynamics of interaction and measuring gaze at the operations

level reflects perception more than collaboration, elaboration and dialogue. Hence, our methodological question is what gaze measure reflects the dynamics of dialogue?

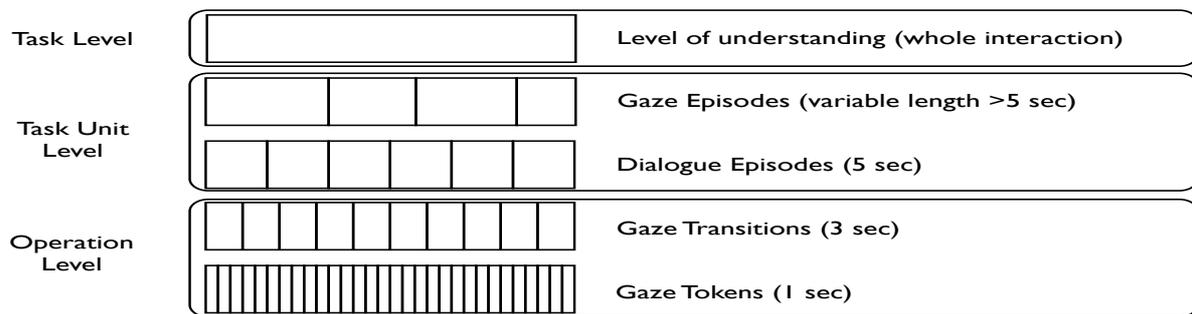


Figure 1. Interaction of the pair divided into different levels of time granularities.

We define gaze measures on two levels.

- On the task unit level, gaze episodes correspond to moments characterized by a stable gaze togetherness and gaze focus. For example, in a focused/together episode, programmers look together at a limited set of objects. These episodes typically last from 5 seconds up to 100 seconds.
- On the operations level we use gaze transitions among different set of objects. The transitions are based on a segmentation of gaze into 1-second slots and last for 3 seconds.

We define cognitive measures on two levels:

- On the task level we rate the level of understanding based on the explanations that are provided by the participants.
- On the task unit level we categorize the dialogues of participants depending on whether they are task related (describing the program) or whether they are about managing the task.

Research Questions

The answer to the methodological challenge allows for new research questions to be asked about the relationships between two consecutive levels of time granularity:

Question 1: task level and task unit level: How does the level of understanding relate to the prevalence of different gaze episodes?

Question 2: task unit level: How do the types of gaze episodes relate to the types of dialogue episodes?

Question 3: task unit level and operation level: How do different dialogue episodes relate to the different gaze transitions?

Experiment

In the experiment, pairs of subjects had to solve two types of pair programming tasks. The task consisted of describing the rules of a game implemented as a Java program. The experimental data used for this paper is the same as used in Nüssli (2011) and Jermann & Nüssli (2012), however the questions and analysis presented hereafter are completely different. 32 students participated in the study. The participants were typical bachelor and master students aged from 18 to 29 years old with a median of 23 years old. The participants were paired into 16 pairs without further consideration of their level of expertise, gender, age or familiarity. The subjects did a pretest that consisted of individually answering thirteen short programming multiple choice questions and then collaboratively solved the ten program understanding tasks which overall lasted approximately 45 minutes. Gaze was recorded with two synchronized Tobii 1750 eye-trackers that record the position of gaze at 50Hz in screen coordinates (see Figure 2). The interested reader can find technical details about the setting in Nüssli (2011).

Gaze Tokens

The JAVA program is composed of tokens (see Figure 2, bottom-left). For example, a line of code “location = array [c] ;” contains 13 tokens (‘location’, ‘c’, ‘=’, ‘array’, ‘;’, 2 brackets and 6 spaces). Fixations on the individual tokens are detected using a probabilistic model (for details see Nüssli (2011)). We categorized the program tokens into 3 gaze tokens: **Expression**, **Structural** and **Identifier**. Each second of the interaction is categorized as one of the gaze tokens (based on the maximum probability).

Gaze Transitions

We aggregated three consecutive gaze tokens into the following three categories (see Sharma et al., (2012)):

Expression: if all the three gaze tokens are expressions.

Data Flow: if there is a permutation of expressions and identifiers.

Read: if there is a permutation of all the three gaze token categories.

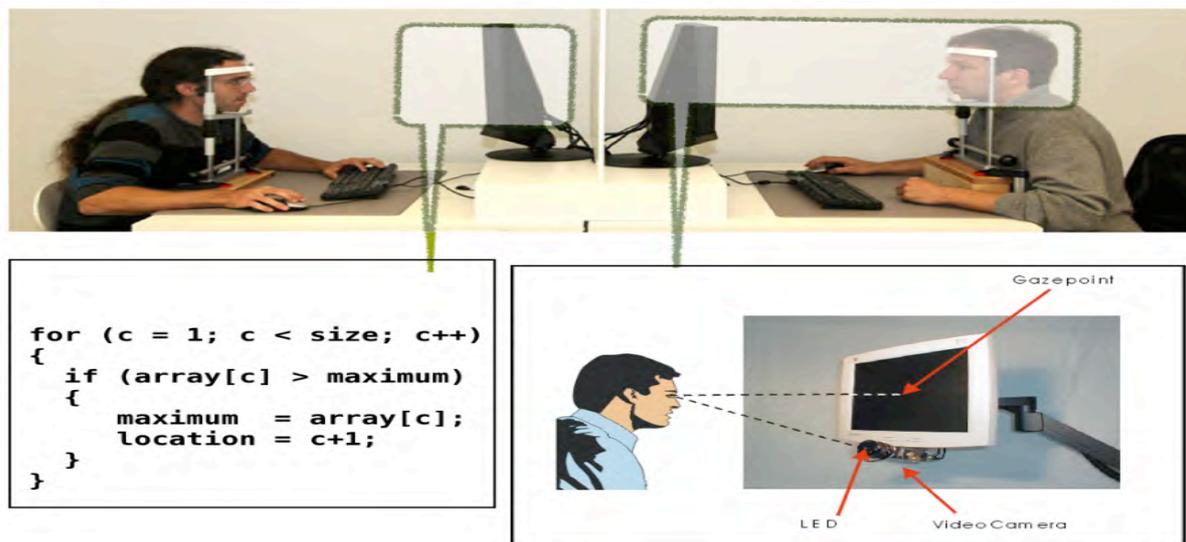


Figure 2. Setup used for the experiment. Upper half shows the laboratory setup for the experiment. The left bubble depicts the stimulus and the right bubble depicts the eye tracking setup.

Dialogue Episodes

We divided the dialogues into 2 major categories according to the content of dialogues. The first category comprises the dialogues containing the description of program functionality; and the second category contains task management utterances, for example, when participants talk about how to proceed, as well as about the controls of interface or where they should look next. Accordingly, we named the two categories as “**description**” and “**management**” respectively.

Gaze Episodes

The gaze episodes are identified based on two parameters: the visual focus of gaze of the participants and the similarity of their gaze. In order to characterize the visual focus of one subject, we compute the object density vector over a given time window. This density vector contains the probability of looking at the different objects of the stimulus. In order to compute this vector, we aggregate gaze data over a 1-second time window and we compute for each object the amount of gaze time that was accumulated inside the object.

We then define the visual focus size as the numbers of objects that are looked at during a 1-second time frame. The rationale is to distinguish between moments where subjects look essentially at few objects versus moments where they look more or less uniformly at several objects. In order to get a quantitative indicator of this focus size, we compute the entropy of the density vector. Entropy measures the level of uncertainty of a random variable, which is, in our case, the objects looked at by the subjects. Hence, high entropy indicates that the subjects looked at many objects (*not focused gaze*), while low entropy indicates that they mostly looked at few objects (*focused gaze*).

For each 1-second timeframe, we define the visual focus coupling as the similarity between the objects looked by one subject and the objects looked by the second subject. We quantify this coupling by computing the cosine between the gaze density vector of one subject and the gaze density vector of the other subject.

Episodes are obtained by combining focus size and similarity. An episode lasts as long as the focus size and similarity stay constant. Technically, a run length encoding procedure applied on the 1-second indicators for visual focus and similarity obtains this. When both subjects are focused and similar we define “**focused together**” gaze episodes. Similarly, we define three other types of gaze episodes that are: “**not focused together**”, “**focused not together**” and “**not focused not together**”. Since we are mostly interested in “*what happens during moments of high collaboration?*” we report only what happens in “**together**” episodes (i.e., “**focused together**” and “**not focused together**”). Typically, a “**focused together**” episode translates in terms of behavior as putting joint efforts to understand code while a “**not focused together**” episode translates as an effort to search some piece of code.

Level of Understanding

We distinguish between two levels of understanding based on how well they performed the description task. Pairs with **high level of understanding** are able to describe the rules of the game along with initial situation, valid moves and winning conditions. Pairs with **low level of understanding** only describe partial aspects of the game structure, and often give algorithmic descriptions of the program and try to guess the detailed rules from the method names; but they failed to get the winning condition.

Results

Question 1: Understanding and Gaze Episodes

The first question concerns the relation between the level of understanding attained by the pair and proportion of time spent by the pair in different gaze episodes. Table 1 shows the ANOVA results for gaze episodes “focused together” and “not focused together” across the two levels of understanding. Pairs with high level of understanding spend more time in gaze episode “focused together” than the pairs with low level of understanding ($F [1,16]=8.70, p=0.01$). Figure 3 shows the difference interval for the two types of gaze episodes across the levels of understanding.

Table 1: ANOVA results for different gaze episodes across two levels of understanding.

Episode Type	Df1	Df2	Sum Sq.	F-value	p-value
Focused Together	1	15	0.09	8.70	0.01
Not Focused Together	1	15	0.06	10.60	0.005

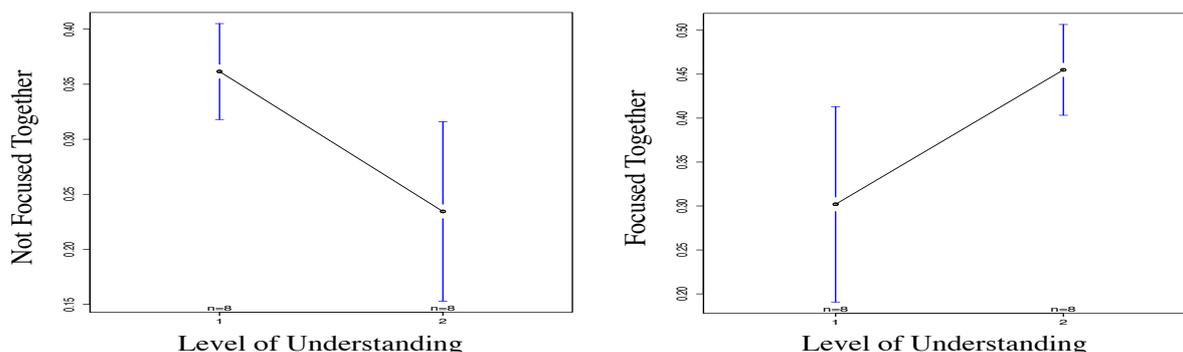


Figure 3. Difference Margin for focused together and not focused together gaze episodes for different levels of understanding (1=Low level of understanding; 2=High level of understanding).

Question 2: Gaze Episodes and Dialogue Episodes

The second question addresses the relationship between the gaze episodes and the dialogue episodes. Table 4 shows the mixed effect model for the two types of dialogue episodes with the factors level of understanding and gaze episodes. There is no significant difference between the proportion of total time spent in dialogue episodes and the gaze episodes, but, there is a significant interaction effect of level of understanding and gaze episodes on the proportion of total time spent on the different dialogue episodes ($F [1,61]=7.60, p=0.01$, Figure 4).

Table 2: Mixed effect model for dialogue episodes with factors level of understanding (UND) and gaze episodes (EPGAZE) (NS= Not Significant).

	Dialogue Episodes							
	Description Episodes				Management Episodes			
Model	Df	Sum Sq.	F-value	p-value	Df	Sum Sq.	F-value	p-value
UND	1	0.05	2.46	NS	1	0.01	1.56	NS
EPGAZE	1	0.04	1.71	NS	1	0.01	0.52	NS
UND * EPGAZE	1	0.17	7.80	0.009	1	0.07	7.60	0.01

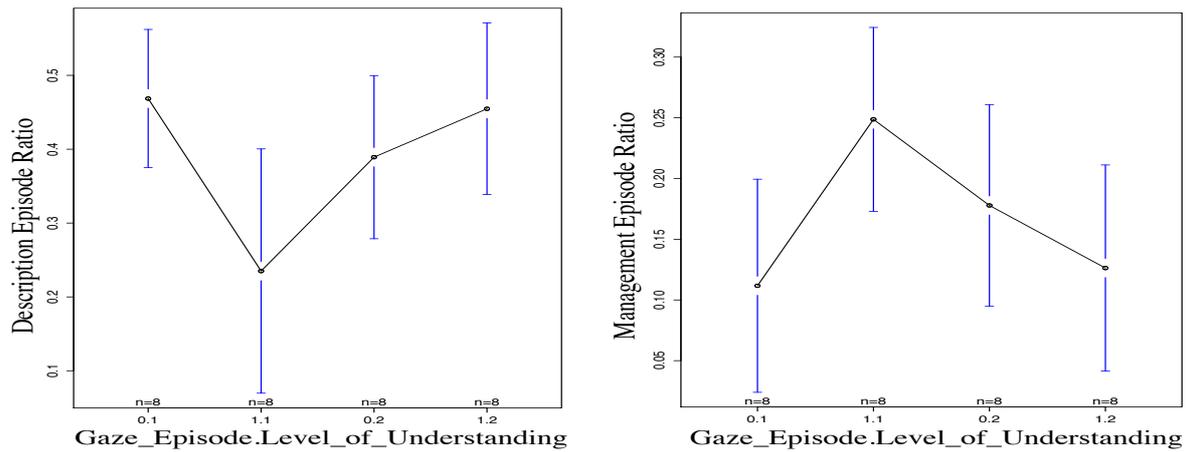


Figure 4. Interaction effect on DESC and MGMT dialogues in focused together and not focused together gaze episodes (0=focused together; 1=not focused together) for different levels of understanding (1=Low level of understanding; 2=High level of understanding).

The pairs with high level of understanding spend more time in “description” dialogue episodes when they are in a “focused together” gaze episode. On the other hand, pairs with low level of understanding spend more time on “management” dialogue episodes when they are in a “focused together” gaze episode. Figure 5 shows the dialogue snippets for pairs with different levels of understanding during different gaze episodes.

Question 3: Dialogue Episodes and Gaze Transitions

The third question considers the relation between the dialogue episodes and the gaze transitions. Table 3 shows the ANOVA results for different gaze transitions across different dialogue episodes. “Description” dialogue episodes have more gaze transitions as “expressions” than the “management” dialogue episodes. Moreover, “management” dialogue episodes have more gaze transitions as “read” than the “description” dialogue episodes. The differences are irrespective of the level of understanding or the type of gaze episodes. Figure 6 shows the difference intervals for the two gaze transition categories across the dialogue episodes.

Table 3: ANOVA (repeated measures) results for different gaze transitions against dialogue episodes.

Transition Type	Df1	Df2	Sum Sq.	F-value	p-value
Expressions	1	63	0.51	8.79	0.004
Read	1	63	0.45	8.31	0.005

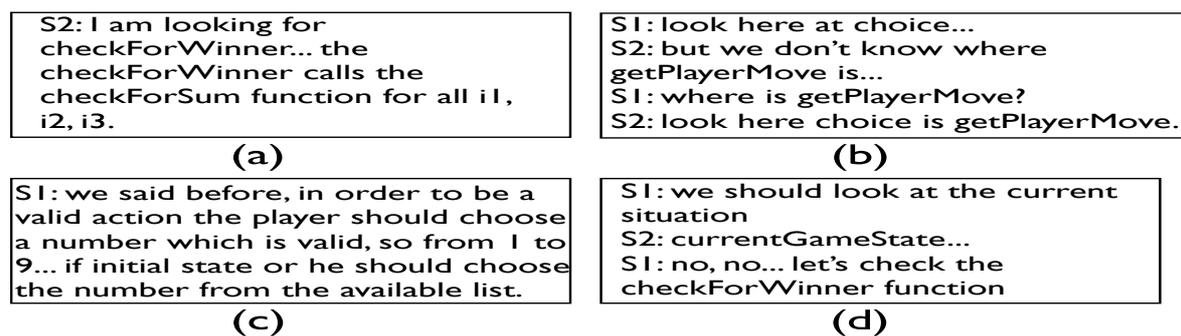


Figure 5. Dialogue snippets for pairs having different levels of understanding during different gaze episodes to show the differences between verbal communications among the participants in the pairs. (a) Low level of understanding and focused together. (b) Low level of understanding and not focused together. (c) High level of understanding and focused together. (d) High level of understanding and not focused together.

Discussion and Conclusion

We conducted the present study with a two-fold motivation. First, identifying gaze and dialogue indicators at different time scales in a pair program comprehension task. Second, bridging different levels of time scales to demonstrate the relationship between gaze and group cognition.

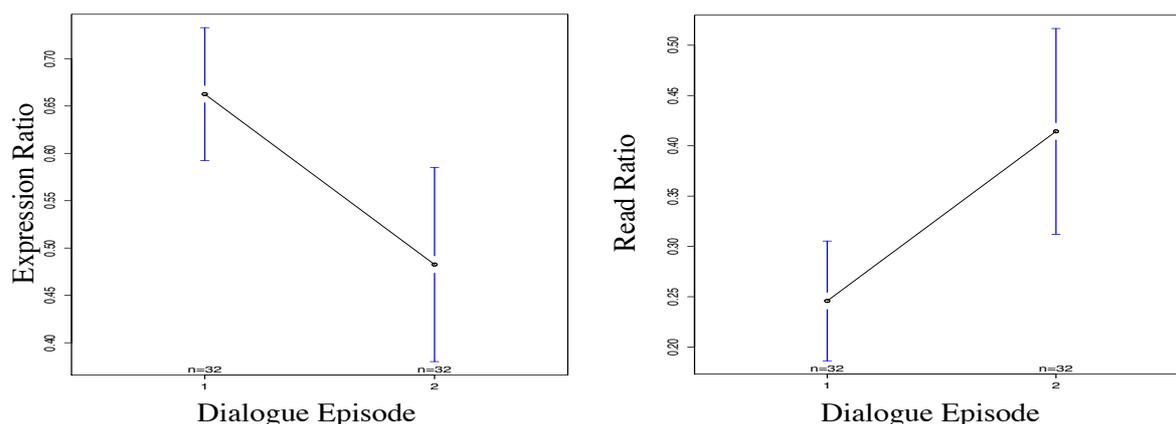


Figure 6. Difference Margin for expression and read gaze transitions for different dialogue episodes (1=Description dialogues; 2=Management Dialogues).

Concerning the methodological challenge, we have proposed gaze episodes as a description of the gaze of a pair on a task unit level. This measure is task independent and can be applied in a wide range of situations. For example, it could be used to describe the focus and similarity of gaze in a concept-mapping task, or in any text reading task. The level of detail for focus and similarity can be varied depending on the accuracy of the eye-tracker and depending on the task. With low-end eye-trackers, one could measure paragraph level, whereas with high-end machines, similarity can be measured at the word base.

Concerning the bridge between two consecutive time scales, we analyze each pair of time scales (see section “Present Study and Questions” and “Results”). We observed that the pairs with high level of understanding spend more time being “focused together” (see subsection “Understanding and Gaze Episodes”) and while they are “focused together” the participants in the pair explain the functionality of the program to each other (Figure 5 (c)). When the pairs with high level of understanding are “not focused together” they talk about their next steps in the task (e.g., they talk about where to look next, Figure 5 (d)). On the other hand, pairs with low level of understanding exhibit the opposite behavior as they spend more time being “not focused together” (see subsection “Understanding and Gaze Episodes”). Moreover, while the pairs with low level of understanding are “focused together” they talk about managing their focus and when they are “not focused together” the participants explain to each other a small part of the functionality of program to maintain a shared focus. Based on our observations, we think that this reflects different ways to understand the program. The “focused” way consists of explaining in depth the functionality of the program, whereas the “unfocused” way consists of describing the code to the partner and to “traverse” the code together.

One important observation is the interaction effect of level of understanding and gaze episodes on the type of dialogues (see subsection “Gaze Episodes and Dialogue Episodes”). There is no direct relation between the gaze episodes and dialogue episodes. However, we see a direct relation between gaze indicators at the level of operations and dialogues. Irrespective of the level of understanding, the pairs have a higher proportion of “expressions” gaze transitions within “description” episodes. Moreover, the pairs have a higher proportion of “read” gaze transitions within “management” episodes. A possible explanation to this observation is that within a “description” episode the participants are more concerned with “what the program does?” This piece of information is contained in expressions within the programming constructs and hence the participants spend their time on understanding the expressions. On the other hand, within a “management” episode participants are talking about where to go next of they are searching a particular piece of code hence the gaze of participants is as if they are scanning the code like English text.

In a nutshell, we showed that there is a relationship between gaze and dialogue indicators at different time scales. These relations help us understand the cognition that underlies program comprehension as well as the collaboration that underlies pair programming. The results are interesting enough to pursue further research in the same direction to find the causality between processes at different time scales.

References

- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, 38(4), 419-439.
- Anderson, J.R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science*, 26(1), 85-112.
- Baheti, P., & Williams, L., (2002). Exploring pair programming in distributed object-oriented team projects. In *Proceedings of XP/Agile Universe*.

- Bednarik, R., & Tukiainen, M., (2006). An eye-tracking methodology for characterizing program comprehension processes. In Proceedings of ETRA'06.
- Cherubini, M., Nüssli, M.-A., & Dillenbourg, P., (2008). Deixis and gaze in collaborative work at a distance (over a shared map): a computational model to detect misunderstandings. In Proceedings of ETRA'08.
- Chi, M., & Roy, M., (2010). How adaptive is an expert human tutor? In Intelligent Tutoring Systems, 2010.
- Chi, M., Roy, M., & Hausmann, R., (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, 32(2).
- Cohen, E. G., (1994). Restructuring the classroom: Conditions for productive small groups. *Review of educational research*, 64(1), 1–35.
- D'Mello, S., Olney, A., & Person, N., (2010). Mining collaborative patterns in tutorial dialogues. *Journal of Educational Data Mining*, 2(1), 1–37.
- Eugenio B. D., Fossati, D., Ohlsson, S., & Cosejo, D., (2009). Towards explaining effective tutorial dialogues. In Annual Meeting of the Cognitive Science Society.
- Griffin, Z.M., & Bock, K., (2000). What the eyes say about speaking. *Psychological science*, 11(4).
- Hejmady, P., & Narayanan, N. H., (2012). Visual attention patterns during program debugging with an IDE. In Proceedings of ETRA'12.
- Jermann, P., Soller, A., & Muehlenbrock, M., (2001). From mirroring to guiding: A review of the state of art technology for supporting collaborative learning. In Proceedings of EuroCSCL-2001, 324-331.
- Jermann, P., & Nüssli, M.-A., (2012). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. In Proceedings of CSCW '12, 1125–1134.
- Lord, R. G., & Levy, P. E., (2008). Moving from cognition to action: A control theory perspective. *Applied Psychology*, 43(3), 335–367.
- Magnisalis, I., Demetriadis, S., & Karakostas, A., (2011). Adaptive and intelligent systems for collaborative learning support: A review of the field. *IEEE Transactions on Learning Technologies*, 4(1), 5–20.
- Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M., (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, 66(2).
- Newell, A., (1994). *Unified theories of cognition*, volume 187. Harvard University Press.
- Nüssli, M.-A., (2011). *Dual-Eye Tracking Methods for the Study of Remote Collaborative Problem Solving*. PhD thesis, Ecole Polytechnique Federale de Lausanne.
- Pietinen, S., Bednarik, R., Glotova, T., Tenhunen, V., & Tukiainen, M., (2008). A method to study visual attention aspects of collaboration: eye-tracking pair programmers simultaneously. In Proceedings of ETRA'08.
- Pietinen, S., Bednarik, R., & Tukiainen, M., (2010). Shared visual attention in collaborative programming: a descriptive analysis. In Proceedings of the ICSE Workshop on Cooperative and Human Aspects of Software Engineering.
- 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).
- Richardson, D. C., Dale, R., & Kirkham, N. Z., (2005). The art of conversation is coordination. *Psychological Science*, 18(5).
- Romero, P., Lutz, R., Cox, R., & Boulay, B., (2002). Co-ordination of multiple external representations during java program debugging. In *Human Centric Computing Languages and Environments, 2002. Proceedings. IEEE 2002 Symposia on, 2002*.
- Sharif, B., & Maletic, J.I., (2012). An eye tracking study on camel case and under score identifier styles. In Proceedings of 18th International Conference on Program Comprehension.
- Sharif, B., Michael Falcone, & Maletic, J.I., (2012). An eye-tracking study on the role of scan time in finding source code defects. In Proceedings of ETRA'12.
- Sharma, K., Jermann, P., Nüssli, M.-A., & Dillenbourg, P., (2012). Gaze Evidence for different activities in program understanding. 24th Psychology of Programming Workshop.
- Stahl, G., (2000). A model of collaborative knowledge-building. In Proceedings of 4th International conference of the learning sciences, 70–77, 2000.
- Webb, N. M., (1989). Peer interaction and learning in small groups. *International journal of Educational research*, 13(1), 21–39.
- Zelinsky, G. L., & Murphy, G. L., (2000). Synchronizing visual and language processing: An effect of object name length on eye movements. *Psychological Science*, 11(2).

Acknowledgment

The work reported in this paper is part of the projects Matching Gaze Patterns and Interaction Patterns in Collaborative Tasks (II) (CR12I1_132996) and Multimodal Interaction Modeling and Regulation of Collaborative Problem-Solving (PZ00P2_126611), both funded by the Swiss National Science Foundation (SNSF).

Effects of Robots' Revoicing on Preparation for Future Learning

Hajime Shirouzu, NIER, 3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo, shirouzu@nier.go.jp
Naomi Miyake, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, nmiyake@p.u-tokyo.ac.jp

Abstract: In the method of *preparation for future learning*, learners often engage in *constructive interaction* with expressing, listening to and integrating their own multiple voices. In order to identify a specific discourse structure underlying successful collaboration, we used a remotely controllable robot as a member of a small discussion group of college students who solved a challenging physics problem. For the robot to act as a listener who solicits voices from students in the group, we manipulated its ways of "revoicing": it performed *minimum revoicing* of students' keywords without evaluative comments in one condition and *guiding revoicing* towards scientific models in the other condition. Comparing these two conditions in addition to a human-only condition, we found that the robot's minimum revoicing fostered students' agency and reflection on their mental models, which prepared them to learn from a lecture and solve a transfer problem. The role of listener for PFL was discussed.

Introduction

In this paper, we examine if there is a specific structure of discourse, i.e., *multi-vocal* structure, in successful collaborative learning using a remotely operate-able robot. In successful classrooms, learners often expand their understanding adaptively by revising their folk knowledge into scientific understandings. This process often requires collaborative forms in which learners express their own multiple thoughts (voices), exchange and integrate them in *constructive interaction* (Miyake, 1986; Shirouzu, Miyake & Masukawa, 2002). As one of forms, we use the learning method of *preparation for future learning* (PFL) (Schwartz & Martin, 2004) or *productive failure* (Kapur & Bielaczyc, 2012) because it is now widely practiced and yields positive results in classrooms. With this method, researchers set up a small-group "groping" activity (e.g., ill-defined problem-solving) prior to a lecture or whole-class discussion, and examine how such an activity prepares students for future learning from instructional resources. Their analyses indicated that the groping fosters students' epistemic agency and drives their extensive search in the problem space. They also implied that constructive interaction took place there. We can identify constructive interaction in Schwartz and Martin's conversation excerpt of ninth graders who tried to invent a measure of reliability for pitching machines. One student first proposed a solution, and the other two students monitored it, commenting on its limitations and flaws, which triggered the first one's reflection. In this way, they took turns acting as *task-doer* (*talker*) and *monitor* (*listener*), leading to the constructive proposal of newer and more general solutions, even if they did not reach canonical ones. In order to trigger such interaction rather than waiting for it to occur, we could put a listener into a group because the listener solicits others' active explanations and makes comments that foster talkers' reflection on their ideas.

Why do we try to put a listener instead of a leader who facilitates and demonstrates how to collaborate? It is because we believe all individuals have the potential to deepen their understanding through collaboration. Besides, observation of well-designed collaborative learning reveals that students who stay mostly silent during group work still listen to others' words and say simple but important words that provoke discussion and perspective shifting among the members. This concept is similar to Greeno and van de Sande's (2007) *constructive listening* and Engle and Conant's (2002) *problematizing*. However, such observation depends on "emergent" group conversation, which we do not know how to replicate systematically.

Revoicing is a good way both to demonstrate active listening and to mark ideas worth reflecting. In this study, robots react to paired students' discussions with different ways of revoicing. Revoicing originally refers to a teacher's repeating or rephrasing students' utterances. Many studies have focused on the pedagogical features of revoicing (e.g., guiding students' attention to critical contents, giving ownership to them, and positioning them to each other) (O'Conner & Michaels, 1996). In order to make the robot a listener, however, we must use student-like rather than teacher-like revoicing. Thus, our robots perform *minimum, stingy revoicing* that simply repeats students' utterances of keywords, regardless of their correctness. We compare this "minimum revoicing condition" with a "guiding revoicing condition" in which the robot adds evaluative phrases to keywords in order to guide students' attention to critical contents. At a glance, the latter condition seems to outperform the former, but we aim to demonstrate that the opposite is the case, since teacher, leader-like revoicing makes students passive and dependent on the robot, rather than promoting their own problem-solving.

If revoicing also serves to mark important ideas, it may provoke conflicts among contrasting ideas of students who engage in PFL tasks. The remaining question about PFL is how collaborative groping leads to learning from a lecture: specifically, whether (1) the groping only has to raise students' agency, and the quality of the search does not need to be high, or (2) the search must cover critical points and raise their awareness closest to the threshold for receiving the lecture. By using a challenging physics problem as a groping task that solicits conflicting ideas (*mental models*), we analyze the collaborative process as search, proposal, question,

and criticism of mental models, and the effects of a robot's revoicing on this process. We also prepare a "human-only condition" of groups of three students and compare it with the minimum revoicing condition to see if a robot that does not express its own ideas but just listens and re-utters helps students' construction of mental models by problematizing the differences among their ideas and keeping their multi-vocality.

Method

Participants

Forty-five undergraduates of a Japanese university participated. We assigned six pairs (12 students) each to the minimum revoicing condition and the guiding revoicing condition, and seven groups of three students (a total of 21 students) to the control (human-only) condition. Members in a pair or group knew each other well. None of the students knew the tasks, and there were no remarkable differences in their physics ability among the groups. We used the desktop robot "Robovie-W" for the experiment. It was 30cm tall and had 17 degrees of freedom with a built-in camera, speaker, and microphone. Robot utterances and actions can be generated and adjusted with a remote control by an operator, and utterances are achieved by voice-synthesizing software (XIMERA).

Tasks

Bobbin Problem

We chose the "bobbin problem" from Anzai and Yokoyama (1984) as the main task. The problem is multiple-choice, to predict the direction of movement of a bobbin as illustrated in Fig. 1. It is quite difficult to determine the correct answer as (1). More than 90% of university students as physics beginners select incorrect answers (2) and (3). According to Anzai and Yokoyama's analysis, beginners tend to make mistakes by drawing on their experience that the bobbin rolls clockwise when the string is pulled from the "fixed" center of the bobbin.

"The centers of two circular frames are interconnected by an axle, and a string is wound around it like a bobbin, as illustrated in the figure below. What will happen if you pull the string as shown in the figure? The discs may roll, but never slide. Mark the number that you think is correct. Let's discuss your reason for selecting it."

- (1) The bobbin rolls to the left (counterclockwise).
- (2) The bobbin rolls to the right (clockwise).
- (3) The bobbin does not move.
- (4) Other. (Write your answer concretely.)

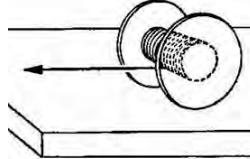


Fig. 1. Bobbin problem.

Transfer Problem: Toilet Paper Problem

We selected a transfer problem from Anzai (1991) (hereafter "the toilet paper problem"). This problem questions the movement of an object when the positions of the fulcrum and the power point are changed from those depicted in Fig. 1 and pulled as shown in Fig. 2. The correct answer is (2), but the participant who simply assumes that the object always rolls to the left when the string is pulled to the left will select the wrong answer.

"An axle was passed through the center of an object and a string was wound on it, as depicted in the figure below. What will happen if you pull the string as shown in the figure? The object may roll, but never slide. Mark the number you think correct, and discuss your reason for selecting it."

- (1) The object rolls to the left (counterclockwise).
- (2) The object rolls to the right (clockwise).
- (3) The object does not move.
- (4) Other. (Write your answer concretely.)

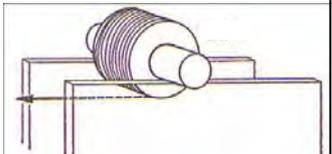


Fig. 2. Transfer problem: Toilet paper problem.

Lecture

For the bobbin problem, we delivered a lecture, the key points of which are as follows: "First, let's regard the contact point between the bobbin and desk as the fulcrum, and the point from where the string comes as the power point (Fig. 3a). Next, consider the line segment that passes both of these points as the axis of the bobbin, and you may understand more easily that the line segment falls to the left around the fulcrum (Fig. 3b). However, the bobbin is circular, not being composed of a single line segment. Therefore, if you assume the circle contains innumerable line segments, you can see the axes will fall to the left, one after another. As a result, the circle rolls to the left (Fig. 3c)." In order to provide exactly the same explanation for every condition, we recorded the lecture beforehand, cut it, and attached voice files with a total of 13 power-point slides.

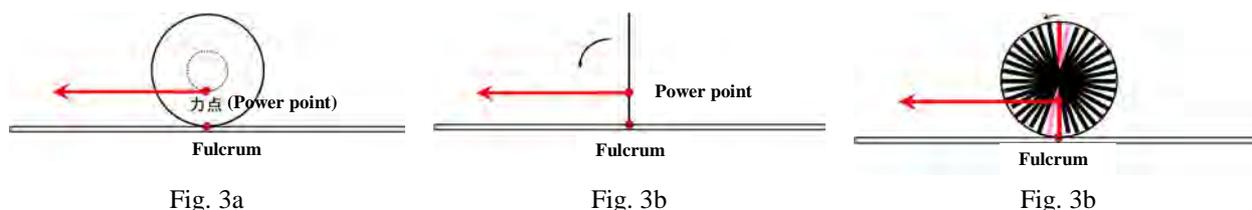


Fig. 3. Explanation slides for the bobbin problem.

Procedure

Minimum Revoicing Condition

First, each pair of students had 5min to interact freely with the robot. For ice-breaking, the robot asked the names of students and called them by their names, and they exchanged information about their origins. Other operations were left to the discretion of the operator of the remote control. After a brief explanation of the experiment, we distributed sheets of paper to the students describing the bobbin problem in Fig. 1. The students attempted to solve the problem for 10min while talking freely. The robot first asked students to read the problem aloud, in order to enhance their relationship. In the prediction phase, the robot revoiced the students' discussion, the details of which are presented in the next section. In the discussion phase, we distributed sewing machine bobbins. After conducting the experiment and confirming the result, the students discussed reasons for 6min (Fig. 4a). The robot also performed revoicing. An atmosphere was created so that the robot observed the experiment as the students discussed while pulling the bobbin and participated in their discussions. Thereafter, the students listened to the lecture for 5min, as displayed on the screen. The robot also turned its face to the screen to create the appearance of viewing it with the students (Fig. 4b). Upon completion of the explanations, the robot turned its face to the students again. If the students did not engage in discussion, it questioned, "Did you understand the explanations? Why did it go to the left?" If they were discussing the explanations, the robot waited a while and then asked these questions. After receiving answers from the students, the robot questioned, "What are the key points or keywords?" When only one student answered the first question, the robot asked the member who did not answer. Lastly, the robot asked students for summarization, "The lecturer mentioned the fulcrum, power point, and circle, didn't he? How did they relate to this problem?" Five minutes was allowed for the above discussion. Upon completion of discussion, power to the robot was turned off.

Next, we distributed sheets of paper describing the toilet-paper problem in Fig. 2. The students were told to talk freely for 5min, and to write down the answer and the reason for it. If the group did not reach a consensus, they were allowed different answers. These answers provide with primary data for deciding their performance of the transfer problem. After the teacher demonstrated the answer using toilet paper in front of the classroom, students spent another 5min discussing the reason and then answered the question again. Lastly, questionnaires and interviews about the impression of the robot were conducted for 10min. The questionnaires contained six items (e.g., "Do you want to learn together with Robovie again?" "Did you think that Robovie knew the answer to the bobbin problem?" and "What partner do you think Robovie would be if it were human?"). All these processes were recorded on video, ICR, and the log of the remote-control system. The groups were separated far enough from each other for clear recording and avoidance of contamination.



Fig. 4a. Discussion



Fig. 4b. Listening to lecture

Fig. 4. Humans and the robot on the bobbin problem.

Guiding Revoicing Condition

The procedures for the guiding revoicing condition were the same as those for the minimum revoicing condition, except for the method of revoicing, as described in the next section.

Control (Human-Only) Condition

The procedures for the control condition were the same as for the two conditions above, except that three students formed a group and the prompts for discussion after the lecture were given on a piece of paper. Questions about "the reason for rolling to the left," "key points/words," and "relationships among the fulcrum, power point, and circle" were printed, and students were instructed to discuss them at their own pace.

Experiment Manipulation: Robot's Revoicing

The robot merely repeated the keywords that students mentioned under the minimum revoicing condition; however, it reacted affirmatively to correct keywords and negatively to incorrect keywords as terms in physics under the guiding revoicing condition. For example, under the minimum revoicing condition if students said “It rolls to the right” (wrong prediction), the robot said only “Rolls to the right”; however, under the guiding revoicing condition it suggested a negative evaluation by adding a phrase like “Does it roll to the right?”

We implemented revoicing in two ways: the operator pressed the button for a preloaded input prompt, or input the prompt by text. To load the keywords that would appear frequently in students' discussions into the operation system, we observed 24 extra juniors as they discussed the bobbin problem prior to this experiment in a similar experiment procedure. We identified frequently used words both in prediction and discussion phases. These “hot” words included correct physical terms, incorrect ones, and unrelated ones. The left-hand column of Table 1 lists candidate keywords for preloading in the prediction phase under the minimum revoicing condition. Under the guiding revoicing condition, we added phrases (in red, underlined letters in the right-hand column of Table 1) to these keywords (most of these phrases were expressed by Japanese sentence-ending particles such as *yo*, *ne*, *yone*, or *kana*, and were difficult to translate). We added three utterances of “evasion” to avoid questions from students in order to position the students, rather than the robots, as task-doers, along with three utterances for “the answer” and 15 utterances for “reasons.” Eight of these utterances were manipulated across the conditions. For the discussion phase, we prepared 20 preliminary utterances for the minimum revoicing condition (1 for “answer” (“Rolls to the left!”), 12 for “reasons,” 3 for “evasion,” and 4 for “emotional expression”). Utterances of emotion were expression of surprise at the experiment result (e.g., “Oh, why?”). Because they were unnatural as utterances of the guiding-revoicing robot, two of the four utterances were deleted. Therefore, 18 utterances were used for the guiding revoicing condition. Among these, seven were manipulated across the conditions. We manipulated a total of 15 of 41 utterances (36.6%) in sum.

Table 1. Examples of revoicing utterances in the prediction phase of the bobbin problem.

	Minimum revoicing condition			Guiding-revoicing condition		
Ans -wer	Rolls to the left			Rolls to the left		
	Rolls to the right			<u>Does it</u> roll to the right?		
	Does not move			<u>It may</u> move (only when students said “move”)		
Rea- son	Pull to the left			<u>Yes, we</u> pull it to the left		
	It goes forward			It goes forward, <u>doesn't it?</u>		
	The string goes out		Pull it upward	<u>Does</u> the string go out?	<u>Do you</u> pull it upward?	
	Friction		Force	<u>Is it</u> friction?	<u>Is it</u> force?	
	Rotates		It will come here	Rotates	It will come here	
	Does not slide		If it does not slide	Does not slide	If it does not slide	
	Pull the string	Pull	When pulled	Pull the string	Pull	When pulled
	The string comes from the bottom			The string comes from the bottom		
	Pull in parallel with the desk			Pull in parallel with the desk		
	Eva- sion	Well, I can't understand. Please think together.			Well, I can't understand. Please think together.	
Well, how does it move?			Well, how does it move?			
I haven't decided the answer. Let me think more.			I haven't decided the answer. Let me think more.			

For free utterances, the minimum revoicing condition allowed free utterances (inputs) on the principle that “when a cluster of utterances is heard and the two students become silent, the robot will speak the keywords of each cluster.” Under the guiding revoicing condition, the robot also performed utterances on the principle that “it may guide students to the correct answer by adding an evaluating phrase to the keyword, but never tell the answer.” We appointed two postgraduates and four undergraduates who were familiar with the bobbin problem and well-trained as remote-control operators. To minimize the differences between the conditions, we had them take charge of the remote control under one condition as many times as under the other condition. Actually, no large difference in either the numbers or the ratios of preliminarily input sentences and freely input sentences was observed between the conditions. Revoicings were made 8.8 times in the minimum revoicing condition and 7.5 times in the guiding revoicing condition on average per pair in the experiment.

Results

In this section, we first confirm students' performance of the transfer task, then examine their feeling of agency through their perception of the robot, and finally analyze the collaborative process and the robots' effect on it.

Performance of the Transfer Problem

We examined the ratios of students who were able to correctly predict the transfer (toilet paper) problem and those who were able to explain the reason, based on the relationship between the fulcrum and the power point as described in the Lecture section (e.g., “The axis falls to the right because the power point is located under the fulcrum.”). Fig. 5 indicates that the correct answer ratio was the highest under the guiding revoicing condition, followed by the minimum revoicing condition and then the human-only condition, but the difference is not significant. However, the ratio of description of the correct reason under the minimum revoicing condition exceeded those under other conditions. The chi square test indicated significant differences among conditions ($\chi^2(2) = 6.69, p < .05$), with residual analysis indicating that more students in the minimum revoicing condition gave correct explanations than expected. Considering that no students in the minimum revoicing condition correctly predicted the bobbin problem, they had no superior prior knowledge, but learned from the discussions and lecture, and transferred their knowledge appropriately to the subsequent, toilet-paper problem.

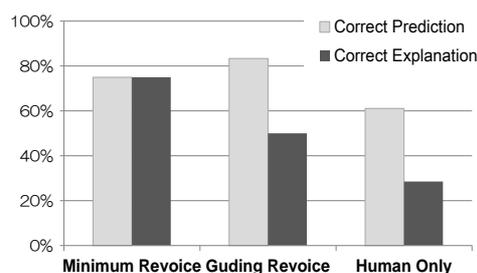


Fig. 5. Results by conditions on the transfer problem.

Perceived Roles of the Robot

Next, we examined the results of post-questionnaires and interviews to determine the students' perception of the robot. First, the ratio of students who felt that the robot knew the answer was higher under the guiding revoicing condition (75%) than under the minimum revoicing condition (58%). We classified their perceptions of the robot into five categories using their literal expressions (Table 2). While 58% indicated that the robot was like their friend under the minimum revoicing condition, nobody answered this way under the guiding revoicing condition. Instead, the robot was perceived as a heterogeneous, obtrusive being, such as a teacher or facilitator who gave guidance or as a child “who always asks ‘why?’” Corresponding to this result, 83% under the minimum revoicing condition indicated that they wanted to learn with the robot again, whereas only 50% did so under the guiding revoicing condition. In summary, the students in the minimum revoicing condition tended to perceive the robot as their thinking partner who did not know the answer. The students did not seem to rely on the robot but to think of themselves as epistemic agents.

Table 2. Perceived roles of the robot (Number of persons).

	Teacher	Facilitator	Student/Friend	Listener	Children
Minimum	1	3	7	1	0
Guiding	2	4	0	3	3

Collaborative Groping Processes

These results imply that the subtle difference in revoicing made all the difference in test performance and perception of the robot. In order to determine how the students deepened their understanding through interaction with the robot, we analyzed the process of collaborative groping before the lecture. Tables 3 presents excerpts of typical pairs under the minimum revoicing condition, and Table 4 presents those under the guiding revoicing condition. Before tracing them in detail, we explain *types of utterance* and *mental models*.

Types of Utterance: We roughly classified students' and robots' utterances into six types: the students' “questioning” of answers to the robot, their “reasoning” of answers and justification, “observation” of experiment results, “revoicing-back (re-uttering)” to the robot's revoicing (e.g., line 20 of Table 3), and the robot's “evasion” and “revoicing.” Using these types, we could determine the discourse pattern among the students and the robot (e.g., how the robot avoided students' questioning and let them act as reasoners).

Mental Models of the Bobbin Problem: We coded utterances and gestures of students' “reasoning” into five mental models (and their variations) in Fig. 6. These models are developed from Anzai and Yokoyama (1984) and our observations. The *tension-force model* (Fig. 6a) simply considers the tension force in the string that pulls the bobbin to the left (broad arrow in Fig. 6), resulting in the correct answer. Although this model has a weakness of not specifying how the string behaves as the bobbin moves, we often observed that it re-emerged

in students' discussions as if they could not ignore it to the end. The *rotation-only model* (Fig. 6b) exploits experiential knowledge of something rotating around a fixed axle (e.g., toilet paper). This model does not predict the direction of the bobbin movement. Expanding the rotation-only model, the *rotation-to-right model* (Fig. 6c) assumes that the rotation creates the force to move to the right, resulting in the wrong answer. The tension-force model and the rotate-to-right model predict opposite directions of the bobbin; however, the students referred to them naturally, as if they believed both. Thus, these models are *fragmental, multiple, and autonomous objects* (Williams, Holland & Stevens, 1983). However, the students sometimes struggled to integrate them, through which two other models developed. One is the *correct model* (Fig. 6d), which predicts that the bobbin rotates to the left with winding the string. Interestingly, students often referred to it while denying its possibility. In this sense, this model is not fully scientific but only phenomenological, yet closest to scientific models such as that delivered by the lecture. The other is the *force-balance model* (Fig. 6e), which claims that the tension force balanced with the rotation (or friction) force, resulting in no movement (wrong answer). Using this framework, we can analyze how widely the students explored the problem space and how closely they approached the scientific models.

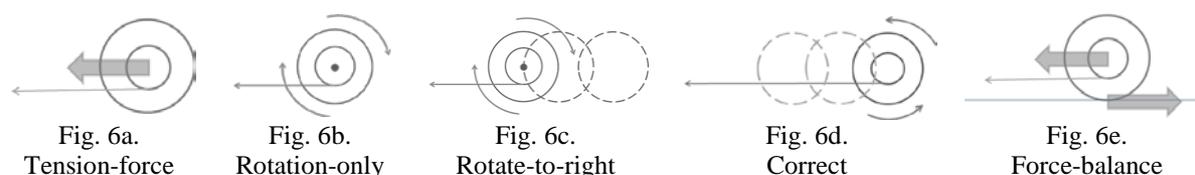


Fig. 6. Mental models of the bobbin problem.

Pair in the Minimum-Revoicing Condition: As indicated in Table 3, Pair A under the minimum revoicing condition, especially Student A1, first asked the robot for the answer, but the robot gave evasive responses (lines 1-6). The students then began to reason and express various models (lines 7-11), which caused conflicts between the models (italicized in lines 10-11). After the robot gave evasive responses again (lines 12-13), A1 developed his model into the rotate-to-right model, against which Student A2 protested from using the tension-force model (lines 14-15). Here, the robot first revoiced keywords, which were revoiced back by Student A1 and integrated into his explanation (underlined in lines 18-20). That is, A1 used the robot's subjective mode (if-clause) at line 19, and continued reasoning about the importance of the rotation (since preloaded revoicings are not always exactly the same as students' expressions, some students utilized such *new* forms of expressions). As a result, Student A1 leaned to the rotate-to-right model; in contrast, A2 seemed to focus on the tension force (line 15) and simulated the leftward movement again (lines 21-22), resulting in externalization of the correct model. Even though he might not have believed this model, it is important that the students were exposed to it. As a result, they focused on the movement of the string as soon as they observed the experiment results (line 102), the utterance of which was also revoiced by the robot and revoiced back by themselves (lines 103-104) and developed into another model (line 105). In summary, the robot's revoicing neither happened often nor guided students to scientific models (line 20); instead, it indirectly problematized the difference and conflict among the models and pushed the students towards integration.

Pair in the Guiding-Revoicing Condition: In contrast, Pair B under the guiding revoicing condition first talked about reasons (Table 4, lines 1-2); however, with the robot's guiding revoicing at line 6, Student B1 noticed the sentence ending particle of the question ("kana") and interpreted that the robot knew the answer (line 9). They then engaged in obtaining the answer from the robot.

In sum, the interactional patterns of both pairs rapidly formed through interactions in the early stage of the prediction phase, where participants neither knew the correct answer nor knew if the robot knew it, and might expect to gain it from the robot. In that stage, guiding revoicing tended to deprive the students of agency.

Mental Models and Subsequent Learning

We analyzed the quantity and quality of mental models to which the students referred. As indicated in Table 3, Pair A referred to more kinds of models more times (seven kinds, ten times) than Pair B in Table 4 (one kind, two times). Qualitatively, Pair A questioned the rotate-to-right model and referred to the correct model, which was not observed in Pair B. Table 5 summarizes all the results of the three conditions during the prediction and discussion phases, broken into correct and incorrect explainers of the transfer problem. The number indicates the average per student, and that in parentheses indicates standard deviation per student. "Referring to Correct model or/and Questioning Rotate-to-right Model" means percentages of the pairs/groups who referred to the correct model or questioned the rotate-to-right model in the total of pairs/groups in each cell. Because such references were infrequent and we assumed that exposure to them was important, we took the pair or group as a unit of analysis and examined whether at least one member made such references. We also grouped a pair or group as "correct" if at least one member correctly explained the transfer problem.

Table 3. Collaborative groping by a pair under the minimum revoicing condition.

Line	Talker	Utterance	Utterance type	Mental model
1	A1	What do you think?	Questioning	
2	R	Well, I can't understand. Please think together.	Evasion	
3	A1	Do you think if it rolls to the left?	Questioning	
4	R	Well, how does it move?	Evasion	
5	A1	Does it roll to the right?	Questioning	
6	R	Well, how does it move?	Evasion	
7	A1	If it doesn't rotate, it comes... [gesture]	Reasoning	Tension-force
8	A2	The problem says "it never slides." [gesturing the leftward-move]	Reasoning	Tension-force
9	A2	When we pull the string this way, the bobbin rotates like this...	Reasoning	Rotation-only
10	A1	And the string pulls the bobbin this way?	Reasoning	Tension-force?
11	A2	Coming with rotating like this? [gesturing the bobbin's leftward movement with the string coming untied]	Reasoning	Tension-force + Rotation-only?
12	A1	Do you think the bobbin remains still?	Questioning	
13	R	Well, how does it move?	Evasion	
14	A1	The bobbin rotates like this to that way. [gesture]	Reasoning	Rotate-to-right
15	A2	But the pulling force works this way. [gesture]	Reasoning	Tension-force
18	A1	Because the problem says " <u>it never slides</u> ," so...	Reasoning	
19	R	If <u>it does not slide</u>	Revoicing	
20	A1	<u>If it does not slide</u> , the point is the rotation of the axle	Revoicing-back	
21	A2	If we pull the string, it rotates like this,	Reasoning	Rotation-only
22	A2	unless the bobbin rotates like this to the left [gesturing the bobbin's leftward movement with winding the string]	Reasoning	Correct
<i><At the discussion phase (after the bobbin delivered)></i>				
102	A1/A2	Oh, it (the string) winds!	Observation	
103	R	The string gets wound.	Revoicing	
104	A1/A2	Yes, it gets wound...	Revoicing-back	
105	A2	There are only two forces leftwards and downwards, so it comes...	Reasoning	To Correct one

Table 4. Collaborative groping by a pair under the guiding revoicing condition.

Line	Talker	Utterance	Utterance type	Mental model
1	B1	If it does not slide, it goes to the right.	Reasoning	Rotate-to-right
2	B2	Because the string is wound clockwise, when we pull it, the bobbin goes that way.	Reasoning	Rotate-to-right
3	B1	The answer may be remains still or moves to the right. Do you think if it remains still?	Questioning	
4	R	Well, how does it move?	Evasion	
5	B1	Or, <u>rolls to the right</u> .	Questioning	
6	R	Does it <u>roll to the right</u> ? (Migi-ni-korogaru-kana)	Guiding revoicing	
7	B1	Does it? (kana)	Questioning	
8	R	Well, how does it move?	Evasion	
9	B1	You said the answer is the right!	Blaming	

Table 5. Mental models referred by correct and incorrect explainer of each condition.

Condition	Kinds	Times	Referring to Correct model and/or Questioning Rotate-to-right Model
Minimum: Correct (n = 9)	3.78 (1.87)	6.22 (3.22)	80% (4 out of 5 pairs)
Minimum: Incorrect (n = 3)	2.67 (1.25)	3.33 (2.05)	0% (0 out of 1 pair)
Guiding: Correct (n = 6)	2.33 (1.11)	4.00 (2.74)	75% (3 out of 4 pairs)
Guiding: Incorrect (n = 6)	2.00 (1.00)	1.83 (1.07)	50% (1 out of 2 pairs)
Human: Correct (n = 6)	1.83 (0.69)	2.50 (0.76)	100% (3 out of 3 groups)
Human: Incorrect (n = 15)	1.60 (1.62)	2.00 (1.90)	25% (1 out of 4 groups)

As shown in Table 5, first, the numbers of models depended on the conditions, but the numbers of times of reference were constantly high for the correct explainers, regardless of the conditions (shaded in Table 5). Thus, successful members repeatedly referred to the particular models. Second, the successful pairs/groups

tended to question the wrong, rotate-to-right model and/or verbalize the correct model (also shaded in the table). These results imply that the quality of the collaborative groping should be as high as possible for receiving the lecture. In addition, the minimum revoicing condition in total contributed to raising this quality (see “*n*” of Table 5). Space prevents us from describing how the revoicing functioned in all the pairs, but we often observed “revoicing-back.” For example, in one pair, when one member referred to the rotate-to-right model as “If we pull it this way, it produces the force for the bobbin to rotate to that way,” the robot just revoiced “Force” and the other member said “Force, pull, force, pull... So if we pull it this way, the force works this way too.” The latter member, having the tension-force model in mind, seemed to take advantage of the robot’s revoicing and put that model on the table again, which kept their multi-vocality and sustained the discussion.

Discussion

This study included a robot in a discussion group of collaborative learning, and let it perform revoicing to the other members in a minimum, stingy way or in a facilitative, guiding way. We found that minimum-revoicing promoted students’ performance of the transfer problem and egalitarian perception of the robot, which implies students’ agency as *task-doers* of problem solving. Process analyses indicated that guiding-revoicing made the students perceive the robot as a *knower*; whereas minimum revoicing caused them to perceive it as a *co-solver*. Even though the minimum-revoicing robot provided no new information, it listened to students’ words heedfully and revoiced them. Reciprocally, the students did not ignore the robot but listened to it and took advantage of the revoiced words. Since the task provoked multiple mental models, the students often confronted conflicts with one another. The robot’s revoicing contributed to making explicit such conflicts, letting role-exchange happen in the students, and sustaining the discussion, which forced students to integrate these models.

This paper provides PFL researchers with the finding that collaborative groping should prepare students at the highest levels of understanding for receiving a lecture. Schwartz and Bransford (1998) once reported that groping activity creates “a time for telling” for teachers; however, this study implies that such a time may be very limited. A time for telling comes diversely, depending on students’ readiness, and thus repeated chances to access instructional resources are needed.

The present study has many limitations in addition to the small sample size. The experiment situation in which the students first worked with the robot might cause them to focus too much on the robot’s words (the revoicings were often made but ignored in the human-only condition), to which the second author is now conducting experiments that include successive opportunities for children to interact with the robot. Some operators reported difficulty in remote-controlling (e.g., when to avoid, when to revoice, and how to revoice). With more findings of HRI and HRL, we should design a longer collaborative learning environment.

References

- Anzai, Y. (1991). Learning and use of representations for physics expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a General Theory of Expertise* (pp. 64-92). New York: Cambridge University Press.
- Anzai, Y., & Yokoyama, T. (1984). Internal models in physics problem solving. *Cognition and Instruction*, 1, 397-450.
- Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement. *Cognition and Instruction*, 20, 399-483.
- Greeno, J., & van de Sande, C. (2007). Perspectival understanding of conceptions and conceptual growth in interaction. *Educational Psychologist*, 42, 9-23.
- Kapur, M., & Bielaczyc, K. (2012). Designing for Productive Failure. *The Journal of the Learning Sciences*, 21, 45-83.
- Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive Science*, 10, 151-177.
- O’Conner, C., & Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussion. In D. Hicks (Ed.), *Discourse, learning and schooling* (pp. 63-103). New York: Cambridge University Press.
- Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for learning: The hidden efficiency of original student production in statistics instruction. *Cognition & Instruction*, 22, 129-184.
- Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science*, 26, 469-501.
- Williams, M. D., Hollan, J. D., & Stevens, A. L. (1983). Human reasoning about a simple physical system. In D. Gentner & A. L. Stevens, (Eds.), *Mental models* (pp.131-153). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Acknowledgments

This study is supported by MEXT Japan, Grant-in-Aid for Scientific Research on Innovative Areas No. 4101-21118001 and 21118007.

Examining Dynamics of Implementing Flexible Group Discourse in a Principle-based CSCL Environment

Tuya Siqin, Jan van Aalst, Samuel Kai Wah Chu, The University of Hong Kong, Pokfulam Road, Hong Kong
Email: siqintuya2010@hku.hk, vanaalst@hku.hk, samchu@hku.hk

Abstract: This study aimed to examine the effect of flexible group collaboration on students' constructive discourse in a principle-based CSCL environment. The participants included 27 Chinese undergraduate students taking a 16-week introductory research methods course. The online discourse adopted a flexible collaboration structure: fixed small group collaboration in the first eight weeks, and opportunistic collaboration in the second eight weeks. The data were collected from students' online discourse notes and assignment tasks. Group differences in small group collaboration and differences between two collaboration structures were examined. Findings indicated that flexible collaboration design helped students to work towards constructive discourse progressively. The study also provided evidence of how students new to principle-based approach engaged in online discourse for conceptual understanding in the Chinese context.

Introduction

Most studies in CSCL area use the small group as the collaboration structure in classrooms. Empirical studies have identified that there can be different patterns of group interaction and discourse in a same learning environment. For instance, Hmelo-Silver (2003) uncovered possible group differences owing to the nature of knowledge that students created. Muukkonen and Lakkala (2009) stated that different groups may hold different epistemic goals in their online inquiry. In another study, different patterns of online discourse were identified: some groups focused on knowledge sharing while others concentrated mostly on knowledge construction or knowledge creation discourse (van Aalst, 2009). In addition, the importance of collaboration at the whole class level has also been pointed out. For instance, Sawyer (2003) argued that most social collaboration in real life is improvisational and that it is necessary to account for this characteristic in classroom collaboration. Online discourse can better cope with emergent learning needs and ongoing inquiry progress as students opportunistically choose collaborators. Zhang et al. (2009) studied three social configurations in the context of encouraging "collective cognitive responsibility" (i.e., knowledge building): fixed groups, interacting groups, and opportunistic groups. During opportunistic collaboration, a number of informal groups formed, disbanded, and recombined to pursue both individual and collective understanding. The examination of the participatory patterns as well as students' knowledge gains indicated that opportunistic groups facilitated knowledge advancement. But little is known whether this design can be generalized to different classroom settings.

In online discourse, students are not only expected to complete group tasks and share ideas; they are also expected to share responsibility to steering toward constructive knowledge work (Muukkonen, Lakkala, & Hakkarainen, 2005; Olson, 2003; Scardamalia & Bereiter, 2003). In other words, students should take the roles of being active agents in the process of externalizing cognition at collective levels along with social collaboration, so that knowledge can be discussed, co-constructed, and advanced. One innovative pedagogical approach that encourages and supports students as agents of their collaborative learning is principle-based, and provides principles rather than procedures to guide their work (Hong & Sullivan, 2009; Scardamalia & Bereiter, 2003). Previous studies (e.g., Lee et al., 2006; Zhang, 2011) posited that the application of such an approach in secondary and primary classrooms can facilitate students' collaboration and knowledge building. Despite much progress, principle-based approach remains rarely implemented, particularly in undergraduate classrooms. Different collaboration structures can be adopted only if they are in accordance with principles (Brown & Campione, 1996). To some extent, the small group is suitable for students who are novices in computer discourse, since there are limited numbers of notes to read and respond to, and it is easy to follow learning progress. By contrast, the whole-class discourse is more demanding as students should regulate their own learning continuously, in order to contribute to collective knowledge and benefit from opportunistic collaboration.

Therefore, this study proposed that the adoption of a flexible collaboration structure with a combination of fixed small group collaboration and opportunistic collaboration may be a more appropriate way of helping students engage in constructive discourse in a principle-based CSCL environment. Specifically, the study aimed to partially replicate Zhang et al.'s (2009) design implemented in a Western primary classroom to a one-semester undergraduate course in the Chinese context. The multi-faceted analysis was carried out to reveal possible group dynamics characterizing the nature of online discourse and their relation with conceptual understanding. The following three questions were addressed: (1) How did the different groups collaborate during fixed small group discourse? (2) Did the students go beyond fixed small group collaboration towards

opportunistic collaboration productively? (3) What was the relation between group discourse dynamics and individuals' conceptual understanding?

Methods and Design

Participants and Instructional Design

The participants in this study were one class of 27 undergraduate students majoring in educational technology at a university in mainland China. A one-semester course entitled "Basic Research Methods" was divided into two 8-week phases. In the first phase, the students collaborated in fixed small groups, and in the second phase the whole class collaborated. There was one-hour, student-centered synchronous discourse in an online platform-Knowledge Forum (Scardamalia & Bereiter, 2003) arranged after regular two-hour lectures every week. The students had no previous experience of participating in online constructive discourse before, while the course teacher who had five years' experience of using online platforms to supporting students' constructive discourse. The course materials were divided into several discussion themes covering key concepts included in the textbook. The classroom activities included creating knowledge products (i.e. questionnaires, group reports) through designing and implementing small research projects, along with concept-learning inquiries for strengthening students' understanding. The main goals of group collaboration were: to develop students' responsibility for collaborative learning via conducting small group projects and concept-based discussion, and to help students to obtain deeper understanding of the nature of knowledge in the context of learning about research methods.

A principle-based CSCL environment was constructed based on the twelve knowledge building principles (Scardamalia, 2002). This study adopted four principles: (1) idea-centered progressive discourse; (2) community knowledge, collective awareness; (3) constructive use of information; and (4) monitoring and regulating discourse. These four principles were acceptable for new learners to understand practical meanings of using them (Lee et al., 2006; Hakkarainen, 2009). They could also avoid conceptual overlap and map out core features of the twelve knowledge building principles (Chan & Chan, 2011; van Aalst & Chan, 2007). The students were encouraged to use the above four principles as discussion norms, to work in small groups and then dig deeper to facilitate learning as a whole class. At the same time, scaffolds in Knowledge Forum such as "I need to understand", "Information", "My theory" were intended to support students' cognitive processing corresponding to basic ideas of the knowledge building principles. The students used these scaffolds when they created notes, raised questions, and replied to notes. In addition, the course teacher organized offline activities to facilitate students' online discourse. For instance, the teacher asked students to draw concept maps to frame and plan their online discourse, evaluate progression and constrains of ongoing collaboration based on the four principles.

Data sources and analysis

The data sources in this study included students' discourse notes posted to Knowledge Forum and assignment tasks arranged at the end of the course. Group discourse dynamics were delineated through analyzing social and cognitive indicators, which were measured through social network analysis (SNA) and content analysis. SNA is a quantitative method that reveals features of social structures formulated in a community (Haythornthwaite, 1996). In contrast, qualitative content analysis can uncover the nature of knowledge distributed over a particular network (Gunawardena, Lowe, & Anderson, 1997; Hmelo-Silver, 2003). The combination of these two methods, therefore, enables complement measurements and provides fruitful information about online discourse (De Laat et al., 2007; Lipponen et al., 2003).

The process of data analysis followed four steps. First of all, SNA was conducted to capture the general picture of collaborative networks generated in Knowledge Forum. Two measures: density and betweenness centrality were employed to evaluate unidirectional note reading and note responding interactivity. Density measured the intensity of interconnection among participants; while Freeman's betweenness centrality measured the extent to which this network showed equal and distributed interactions rather than dominated by a few participants (Scott 1991).

Secondly, content analysis was carried out to discern indications of the four principles and characteristics of knowledge distributed in group discourse. For this purpose, a coding scheme was refined through both theory-driven and data-driven approaches, using a note as the unit of analysis. As demonstrated in Table1, four main categories (question, idea, metacognition, and reference) identified were basically in line with the knowledge building principles. The first author rated all online discussion notes, and a second rater re-coded 30% of the notes independently. The inter-rater reliability measured by Pearson Correlation was .83. The proportion of each category of knowledge distributed in group discourse was then calculated, followed by a Chi-square test performed to examine possible group difference.

Thirdly, we used an inquiry thread as the unit of analysis to assess the patterns of group discourse and advances of collective knowledge. An inquiry thread was a number of notes that address the same principal

problem, thus forming a conceptual stream plotted against a timeline (Zhang et al., 2007). Using this method, all the notes were reorganized into inquiry threads in terms of discourse themes being investigated. Then notes included in each thread were sequenced along the timeline of contribution. To trace the processes of constructive discourse and knowledge advances, we further classified two main categories: question and idea into several subcategories. According to Hakkarainen (2003), progressive constructive discourse can be characterized as the iterative process of questioning and explanation, with the shift from fact-oriented to explanatory-oriented knowledge. In van Aalst's (2009) study, questions were subcategorized as seeking facts, clarifications, or explanations; while ideas were classified into seven subcategories: fact, concept, elaboration, explanation, conjecture, opinion, and rise above. Based on these two coding schemes, top-down and bottom-up processes were performed to code all notes relating to questions and ideas. We then obtained the coding subcategories at four levels (from low to high): fact-oriented, clarification-oriented, elaboration-oriented, explanation-oriented question or idea, respectively. The possible knowledge advancement was then examined by assessing the changes of mean levels of questions and ideas produced in discourse threads. The first author rated all online discussion notes, and another researcher re-rated 30% of notes independently. The inter-rater reliabilities were calculated based on Pearson Correlation, to be .83 for question subcategory and .80 for idea subcategory, respectively.

Table 1: Operational definitions of coding categories in online discourse.

Category		Definition
Question	Fact-oriented	Ask for the definition of a concept or factual information
	Clarification-oriented	Ask for clarifying relevant elements or characteristics of a concept, or different opinion
	Elaboration-oriented	Ask for interpretation on relation, difference, practical meaning of certain opinion, claim, or theory
	Explanation-oriented	Ask for providing explanation on a particular theory or strategy of implementing a concept, theory, or claim
Idea	Fact-oriented	Point out a concept or factual information simply
	Clarification-oriented	State conceptual difference, similarity, characteristic, personal opinion or experience
	Elaboration-oriented	Elaborate a theory, claim, or opinion with specific statement
	Explanation-oriented	Explain a concept and theory with the support of relevant information, and example
Metacognition		Monitor, regulate or evaluate ongoing inquiry process and group collaboration progress
Reference		Introduce reference and information from an outside source without any additional interpretation

Finally, the students were asked to complete assignment tasks at the end of the course, so as to examine their understanding of core concepts relating to research methods that had been discussed on the phases of small group collaboration and opportunistic collaboration. These assignments were scored on a 4-point scale to evaluate individuals' conceptual understanding, following the scheme developed by van Aalst (2009) with the consideration of the degree of misunderstanding on key concepts and discourse themes being investigated. Details of this rating scale are shown in Table 2. Two raters scored all the assignments independently and the inter-reliability was .79 in terms of Pearson Correlation. A Pearson's correlation analysis was performed to test the relation between the indicators characterizing social and cognitive dynamics of group collaboration and individuals' conceptual understanding.

Table 2: The rating scale of evaluating students' assignment tasks.

Scale	Definition
1	Strong evidence of misunderstanding, without specific explanation of core concepts being investigated
2	Little evidence of misunderstanding, with vague and unclear explanation of core concepts being investigated
3	No misunderstanding, with explanation but lack of coherence and linking to related core concepts being investigated
4	No misunderstanding, with explanation having coherence and linking to related core concepts being investigated

Results

How did the different groups collaborate during fixed small group discourse?

Small group discourse was examined using social network analysis and content analysis, so that possible differences of group collaboration were disclosed. Overall, the students generated 660 notes in Knowledge Forum. In fixed group phase, the students were randomly assigned to 5 small groups to pursue conceptual learning and group project design for eight weeks. The degrees of group participation and patterns of interaction were calculated (see Table 3). There was no variation among five groups on note reading interactions, as shown by the highest reading density (100%) and lowest betweenness centrality (0.0%). This meant that each participant read their group members' notes actively, resulting in equal and distributed reading interactions occurred in the social networks. By contrast, some variations have been observed in note responding interactions. More interestingly, Group One and Group Two's measurements in interaction patterns showed the same results, with the highest density (100%) and lowest betweenness centrality (0.0%) values on both note reading and responding. It was obvious that all students in these two groups worked intensively with one another. To further uncover possible group differences, the subsequent analyses were carried out to test cognitive processing using quantifying content analysis (Chi, 1997).

Table 3: Patterns of social networks in small group collaboration (with Standard Deviations in Parentheses).

	Group 1 n=5	Group 2 n=5	Group 3 n=5	Group 4 n=5	Group 5 n=6
Mean number of notes	16.6 (4.3)	13.8 (3.7)	9.6 (2.1)	12.4 (2.3)	12.3 (3.8)
Note reading density	100%	100%	100%	100%	100%
Betweenness centrality of note reading	0.0%	0.0%	0.0%	0.0%	0.0%
Note responding density	100%	100%	90.0%	80.0%	93.3%
Betweenness centrality of note responding	0.0%	0.0%	2.8%	12.5%	1.00%

Note: One student was excluded from analysis as he started to attend the course near the middle of the semester.

Figure 1 shows the percentages of knowledge classified by four main categories: question, idea, metacognition, and reference. Apparently, the majority of notes produced by five groups referred to idea category. Each group also generated relative higher proportion of knowledge categorized as metacognition. There were noticeable variations on the proportions of question and reference categories among five groups. A Chi-square analysis confirmed that the distribution of knowledge differed significantly between groups across four categories ($\chi^2 = 26.2, df = 12, p < .05$). In particular, there was a substantial difference ($\chi^2 = 23.6, df = 3, p < .001$) between Group One and Group Two on the distribution of knowledge in terms of four categories, even though they displayed the same patterns of collaborative networks. It was apparent that Group One contributed higher percentage of question compared with other four groups. For Group Two, however, knowledge distribution was mostly dominated by idea, but there were fewer questions than other groups. Group One and Group Two were further selected to examine group discourse patterns and possible knowledge advances emerged in discourse threads.

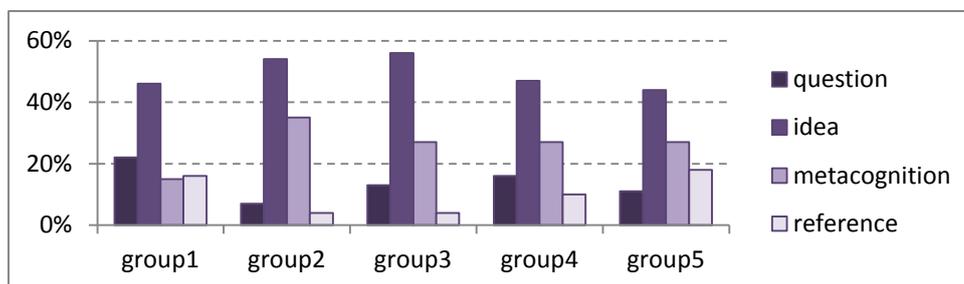


Figure 1. Percentage of different categories of knowledge in small group collaboration.

Overall, both groups formulated five discussion threads covering concepts related to research methods: research question, variables, sampling, interview and questionnaire. On average, there were 16.4 (SD = 12.9) and 13.4 (SD = 6.2) notes in each thread for Group One and Group Two, respectively. We evaluated all discourse threads produced in two groups in terms of the discourse patterns identified by van Aalst (2009). In Group One, two out of five discourse threads revealed knowledge construction, while rests of others remained at knowledge sharing; but for Group Two, only one discussion thread revealed knowledge construction. In order to

uncover to what extent knowledge was advanced, all notes were reordered following the timeline of contributions in each discussion thread. We then divided notes in each thread into two periods with equal proportion of notes. The levels of questions and ideas in discourse threads across two periods were rated on a 4-point scale. As shown in Table 4, Group One generated relatively higher levels of questions and ideas than Group Two, and the mean levels of questions and ideas increased slightly through discourse as well. In Group Two, however, there was no increase observed in the means levels of questions and ideas across two periods. Results indicated that two groups performed differently in sustaining online discourse over time. Group One seemed to work productively towards constructive discourse compared to the counterpart. No statistical test was conducted due to small sample sizes ($n = 5$).

Table 4: Mean levels of questions and ideas for period 1 and period 2 in small group collaboration (with Standard Deviations in Parentheses).

	Question		Idea	
	Period 1	Period 2	Period 1	Period 2
Group 1	2.6 (1.1)	2.7 (1.2)	2.7 (1.0)	2.8 (1.8)
Group 2	2.3 (0.6)	2.3 (0.6)	2.4 (0.9)	2.3 (0.8)

Did the students go beyond fixed small group collaboration towards opportunistic collaboration productively?

To assess whether the flexible collaboration structure promoted students' productive discourse, we used the same data analysis procedures employed in addressing the first research question to evaluate possible group dynamics across two social configurations as well. On average, the students contributed 12.9 (SD = 3.8) and 12.0 (SD = 4.9) notes during small group and opportunistic collaboration, with relative consistent note contributions across two phases. Table 5 shows the patterns of social collaborative networks in two types of collaboration structures. Not surprisingly, the value (68.0%) of note reading density measurement of small group collaboration at the whole class level was much less than the values measured at the five small groups (100%) separately. In general, the rates of note reading and note responding densities increased from fixed small group collaboration to opportunistic collaboration, which indicated that the classroom interactivity had spread to more participants. Simultaneously, betweenness centrality of social collaborative network was calculated considering both note reading and note responding activities. The decreasing trend in this indicator implied that opportunistic collaboration made a broader scope of collaboration possible, resulting in relatively distributed and evenly social network occurred in the classroom.

Table 5: Patterns of social networks in two types of collaboration structures.

	Density of note reading	Density of note responding	Betweenness centrality of collaborative network
Small group collaboration	68.0%	16.6%	4.4%
Opportunistic collaboration	90.3%	21.9%	0.2%

The examination of different characteristics of knowledge (see Figure 2) created by the students found a significance difference on the distribution of knowledge between small group and opportunistic collaboration ($\chi^2 = 38.9$, $df = 3$, $p < .001$). It was apparent that the highest proportion of knowledge contributed by the students during group discourse was idea. Moreover, the students' engagement showed increasing proportions of questions and ideas from fixed small group collaboration to opportunistic collaboration. The subsequent analysis intended to uncover the possible knowledge advancement emerged in discourse threads.

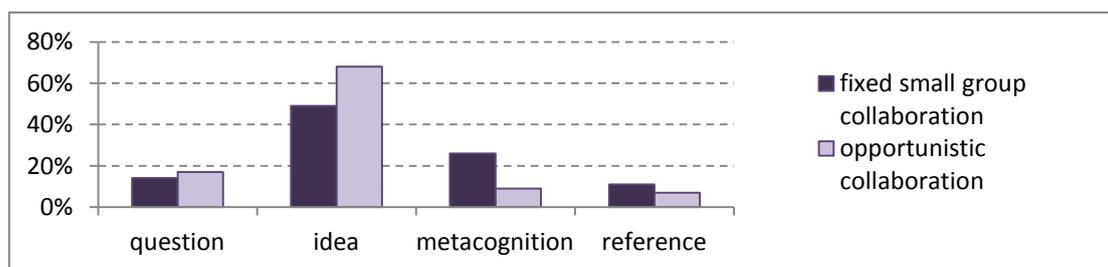


Figure 2. Percentage of different characteristics of knowledge in two types of collaboration structures.

Altogether, the 27 participants were involved in 33 inquiry threads during the semester. We assessed each thread according to the discourse patterns identified by van Aalst (2009). Five of twenty-three inquiry threads involved knowledge construction during small group collaboration, while six of ten threads revealed knowledge construction during opportunistic collaboration. All other threads were assessed as knowledge sharing discourse. In addition, the mean lengths of threads were 14.3 (SD = 8.3) notes for small group collaboration and 31.0 (SD = 22.2) notes for opportunistic collaboration. As found in previous analysis, knowledge construction discourse was examined to be more evident in promoting knowledge advances than knowledge sharing discourse. Five threads rated as knowledge construction discourse in each collaboration structure were selected to further investigate the degrees of knowledge advancement. As shown in Table 6, the mean levels of questions and ideas raised in the process of small group and opportunistic collaboration. It was noticeable that the mean level of questions in opportunistic collaboration was higher than that in small group collaboration. Also, the advancement of knowledge in opportunistic collaboration was more evident than in fixed groups. In addition, there was very little increase observed in the mean level of ideas contributed in two collaboration structures in the discussion threads. The results suggested that the students attempted to engage in constructive discourse by going beyond small group collaboration towards opportunistic collaboration progressively. During this process, the students made progress in generating higher level of questions. However, there was little improvement in contributing higher level of ideas. No statistical test was performed due to limited number of threads ($n = 5$).

Table 6: Mean levels of questions and ideas for period 1 and period 2 in two types of collaboration structures (with Standard Deviations in Parentheses).

	Question		Idea	
	Period 1	Period 2	Period 1	Period 2
Small group collaboration	2.4 (0.9)	2.8 (1.1)	2.6 (0.8)	2.7 (0.8)
Opportunistic collaboration	2.9 (1.2)	3.2 (1.0)	2.6 (0.7)	2.8 (0.7)

What was the relation between group discourse dynamics and individuals' conceptual understanding?

The above two research questions addressed discourse dynamics within and across groups by measuring social and cognitive indicators. The third research question moved to understanding how group collaborative dynamics impacted students' conceptual understanding by performing a Pearson correlation analysis. In general, the students' assignment tasks covered 18 inquiry threads discussed in Knowledge Forum. The assignments were scored to evaluate the degree of conceptual understanding for each individual student. At the same time, social and cognitive indicators corresponding to those 18 discussion threads were calculated. Table 7 shows the relationships among the indicators of group discourse dynamics and individuals' conceptual understanding.

Table 7: Correlations among indicators characterizing group discourse dynamics and conceptual understanding.

	note read	note respond	question	idea	metacognition	reference	level of questions	level of ideas
note respond	.63**							
question	.76**	.81**						
idea	.96**	.72**	.83**					
metacognition	.34	.33	.06	.25				
reference	.40	.34	.16	.36	.00			
level of questions	.65**	.72**	.61**	.74**	.18	.40		
level of ideas	.28	.40	.17	.34	.20	.19	.55*	
conceptual understanding	.64**	.53*	.50*	.64**	.34	.27	.59**	.31

Note: * $p < .05$, ** $p < .01$

Results found that the numbers of note reading and note responding indicators in each discussion thread were significantly correlated to students' conceptual understanding ($r = .64, p < .01$; $r = .53, p < .05$). Of four indicators characterizing different types of knowledge contributed in online discourse, question and idea were correlated significantly to conceptual understanding ($r = .50, p < .05$; $r = .64, p < .01$). However, no significant correlation was found among the numbers of metacognition, reference contributed in the discourse threads and conceptual understanding. The statistical result demonstrates that the mean level of questions contributed to conceptual understanding positively ($r = .59, p < .01$). In addition, the mean level of questions significantly correlated to the level of ideas ($r = .55, p < .05$). Yet, there was no significant relation found between the mean level of ideas and conceptual understanding. The results suggested that contributing larger numbers of ideas and questions could facilitate students' conceptual understanding. In particular, questioning played an important role in deepening understanding.

Discussion and Conclusion

The purpose of this study was to investigate the effect of combining fixed small group and opportunistic collaboration to facilitate Chinese undergraduate students' constructive discourse in a principle-based CSCL environment. Departing from the earlier research that adopted opportunistic collaboration in a Western primary classroom, the current study extended the design in a Chinese context by integrating online discourse into a regular undergraduate course.

The present study did not simply replicate methods of the earlier study that primarily compared social structures and processes at a class level across three successful years. Rather, the study targeted at another angle by examining group collaboration dynamics at three aspects: patterns of social interaction, characteristics of knowledge distributed within groups, and knowledge advances emerged along discourse threads. This kind of multi-faceted analysis not only validated the findings against the previous research, but also uncovered fine-grained dynamics of online group discourse. The examination of group collaboration in fixed small groups found that all five groups showed intensive collaboration, but they differed substantially on contributing different characteristics of knowledge. In particular, two small groups with the same interaction patterns showed large variations in the knowledge distribution and knowledge advancement. It implied that intensive group interactions might not necessarily lead to higher cognitive processing. Moreover, social network analysis revealed that social interactivity spread to more participants from fixed small group collaboration to opportunistic collaboration and this extended scope of group collaboration promoted relative equal interactions in the class, which was compatible with the findings from the previous study (Zhang et al., 2009). Apart from examining the positive impact of changing group configuration on individuals' knowledge gains that has been reported in Zhang et al.'s (2009) study, the present study also tested associations between the indicators characterizing social and cognitive dynamics and students' conceptual understanding. Consistent with Hakkarainen's (2003) study that uncovered the important roles of questions in deepening understanding, data analysis found that both the number and level of questioning were positively correlated to individuals' conceptual understanding. It was noteworthy that the majority of knowledge produced by the students referred to idea. However, the level of ideas did not change greatly in discussion threads and it was also not significantly correlated to conceptual understanding, whereas the number of ideas was positively correlated with conceptual understanding. One reason may be because the students encountered difficulties in improving ideas collectively. Another possible reason is that discourse itself encompasses the interplay between different levels of knowledge, which caused the advancement of ideas along inquiry threads was slower than expected.

To conclude, this study supported the advantage in facilitating students' constructive knowledge work through flexible collaboration design in a principle-based CSCL environment. The examination of students' online discourse and assignment tasks revealed how Chinese undergraduate students new to principle-based approach and constructive discourse could go beyond small group collaboration towards opportunistic collaboration in advancing individual and collective understanding progressively. Additionally, the study provided an example of integrating social network analysis and content analysis to uncover interdependent roles of social and cognitive dynamics in online discourse. Owing to the limitation on small sample size, further studies are needed to validate the findings in various classroom settings. It would also be valuable to trace how opportunistic groups emerge and how individual students benefit from and contribute to group collaboration.

References

- 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.
- Chan, C. K. K., & Chan, Y.-Y. (2011). Students' views of collaboration and online participation in Knowledge Forum *Computers & Education*, 57(1), 1445-1457.
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, 6(3), 271-315.

- De Laat, M., Lally, V., Lipponen, L., & Simons, R.-J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: a role for social network analysis. *International Journal of Computer-supported Collaborative Learning*, 2(1), 87-103.
- 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.
- Hakkarainen, K. (2003). Emergence of progressive-inquiry culture in computer-supported collaborative learning. *Learning Environments Research*, 6(2), 199-220.
- Hakkarainen, K. (2009). A knowledge-practice perspective on technology-mediated learning. *International Journal of Computer-Supported Collaborative Learning*, 4(2), 213-231.
- Haythornthwaite, C. (1996). Social Network Analysis: An Approach and Technique for the Study of Information Exchange. *Library and Information Science Research*, 18(4), 323-342.
- Hmelo-Silver, C. E. (2003). Analyzing collaborative knowledge construction: multiple methods for integrated understanding. *Computers & Education*, 41(4), 397-420.
- Hong, H.-Y., & Sullivan, F. R. (2009). Towards an idea-centered, principle-based design approach to support learning as knowledge creation. *Educational technology research and development*, 57(5), 613-627.
- Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction*, 13, 487-509.
- Lee, E. Y. C., Chan, C. K. K., & van Aalst, J. (2006). Students assessing their own collaborative knowledge building. *International Journal of Computer-Supported Collaborative Learning*, 1(1), 57-87.
- Muukkonen, H., & Lakkala, M. (2009). Exploring metaskills of knowledge-creating inquiry in higher education. *International journal of Computer-Supported Collaborative Learning*, 4(2), 187-211.
- Muukkonen, H., Lakkala, M., & Hakkarainen, K. (2005). Technology-mediation and tutoring: How do they shape progressive inquiry discourse? *Journal of the Learning Sciences*, 14(4), 527-565.
- Olson, D. R. (2003). *Psychological theory and educational reform: How school remakes mind and society*. New York: Cambridge University Press.
- Sawyer, R. K. (2003). Emergence in creativity and development. In K. Sawyer, V. John-Steiner, S. Moran, S. Sternberg, D. H. Feldman, J. Wakamura, et al. (Eds.), *Creativity and development* (pp.12-60). Oxford, England: Oxford University Press.
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal education in a knowledge society* (pp. 76-98). Chicago: Open Court.
- Scardamalia, M., & Bereiter, C. (2003). Knowledge Building. In J. W. Guthrie (Ed.), *Encyclopedia of Education* (2nd edition ed.) (Vol. 4). New York: Macmillan Reference, USA.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York: Cambridge University Press.
- Scott, J (1991). *Social Network Analysis: a Handbook*. Sage Publications, London.
- van Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge construction, and knowledge-creation discourses. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 259-287.
- van Aalst, J., & Chan, C. K. K. (2007). Student-directed assessment of knowledge building using electronic portfolios. *Journal of the Learning Sciences*, 16(2), 175-260.
- Zhang, J. (2011). Sustaining knowledge building as a principle-based innovation at an elementary school. *Journal of the Learning Sciences*, 20(2), 262-307.
- Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of 9- and 10-year-olds. *Education 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.

Resources for Connecting Levels of Learning

Gerry Stahl, Drexel University, Philadelphia, USA, Gerry@GerryStahl.net
Diler Öner, Boğaziçi University, Istanbul, Turkey, Diler.Oner@boun.edu.tr

Abstract: CSCL research typically investigates processes at the individual, small-group and community units of analysis. However, CSCL analyses generally focus on only one of these units, even in multi-method approaches. Moreover, there is little data-based analysis of how the three levels are connected. This paper proposes that the levels of individual learning, group cognition and community knowledge building are connected by *interactional resources*, which can mediate between the levels. A theory of the connection of the levels is sketched. Then examples of such connections by interactional resources are presented from logs of several CSCL experiments. Finally, a curriculum for gradually providing math teachers and math students with a complex of resources relevant to dynamic geometry is described as an example of how to support the connection of small-group interaction with individual understanding and with cultural practices in a CSCL adaptation of geometry education.

The Problem of Connecting Levels

Learning, knowledge building and cognition can be analyzed at multiple units of analysis. For instance, analyses of CSCL are often conducted on one of three levels: individual learning, small-group cognition or community knowledge building. This tri-partite distinction is grounded in the nature of CSCL. With its focus on collaborative learning, CSCL naturally emphasizes providing support for dyads and small groups working together. In practice, CSCL small-group activities are often orchestrated within a classroom context by providing some initial time for individual activities (such as background reading or homework practice) followed by the small-group work and then culminating in whole-class sharing of group findings. Thus, the typical classroom practices tend to create three distinguishable levels of activity. Often, the teacher sees the group work as a warm-up or stimulation and preparation for the whole-class discussion, facilitated directly by the teacher. Conversely, the importance of testing individual performance and valuing individual learning posits the group work as a training ground for the individual participants, who are then assessed on their own, outside of the collaborative context. In both of these ways, group cognition is treated as secondary to either individual or community goals. By contrast, the role of intersubjective learning is foundational in Vygotsky (1930/1978), the seminal theoretical source for CSCL. Regardless of which is taken as primary, the three levels are actualized in CSCL practice, and the matter of their relative roles and connections becomes subsequently problematic (Dillenbourg et al., 1996; Rogoff, 1995; Stahl, 2006).

While these different units, levels, dimensions or planes are intimately intertwined, research efforts generally focus on only one of them, and current analytic methodologies are designed for only one (Stahl, 2013b; Suthers et al., 2013). Furthermore—and most importantly for this paper—there is little theoretical understanding of how the different levels are connected. To the extent that researchers discuss the connections among levels, they rely upon commonsensical notions of socialization and enculturation, popularizations of traditional social science. There are no explicit empirical analyses of the connections, and it is even hard to imagine where one would find data that would lend itself to conducting such analyses (Stahl et al., 2012).

The individual unit of analysis is the traditional default. It is supported by widespread training of researchers in the methods of psychology and education. In the era of cognitive science, analysis made heavy usage of mental models and representations (Gardner, 1985). With the “turn to practice” (Lave & Wenger, 1991; Schatzki, Knorr Cetina & Savigny, 2001), the focus shifted to communities-of-practice. Group cognition lies in the less-well-charted middle ground (Stahl, 2006). It involves the semantics, syntactics and pragmatics of natural language, gestures, inscriptions, etc. These meaning-making processes involve inputs from individuals, based on their interpretation of the on-going context (Stahl, 2006, esp. Ch. 16). They also take into account the larger social/historical/cultural/linguistic context, which they can reproduce and modify (Stahl, 2013b).

This paper will argue that the connections between the individual, group and community planes take place through the mediation of *interactional resources* (section on the theory)¹. To provide specificity and to ground the presentation in empirical data, the paper then considers the resources that appear in recorded examples of mathematical work (section on the analysis). Applying this problematic to the learning of mathematics, the paper adopts a discourse-centered view of mathematical understanding as the ability to engage

¹ While the problem of connecting levels has recently been raised within the CSCL community—e.g., in the workshop at ICLS 2012 (Stahl et al., 2012) and in editorials in *ijCSCL* (Stahl, 2012b; 2013a)—this paper goes beyond those efforts to propose a central role for interactional resources and to review supporting analysis of empirical CSCL data. For further exploration since this paper was written, see (Stahl, 2013c).

in significant mathematical discussion (Sfard, 2008; Stahl, 2008). Here, “discourse” includes gesture, inscription, representation and symbol, as well as speech and text; these multiple modes are often closely interwoven in effective interaction (Çakir & Stahl, 2013; Çakir, Zemel & Stahl, 2009).

Computer technologies play a central role in mediating the multi-level, intertwined problem-solving, learning and knowledge-building processes that take place in CSCL settings. From a CSCL perspective, emergent technologies should be designed to support this mediation. This involves considering within the design process of collaboration environments how to prepare groups, individuals and communities to take advantage of the designed functionality and to promote mathematical thinking at all levels. This paper reports on the design of a curriculum in dynamic geometry to support group cognition, individual learning and community practices in a coordinated way, based on how interactional resources are visibly used in analyzed excerpts of pilot case studies of the use of dynamic-geometry software (section on the pedagogy). The curriculum addresses both communication issues—such as effective collaboration practices—and mathematical issues—such as focusing on dependencies among math objects—as well as technological issues of software usage.

The Theory of Connecting Levels

The idea of viewing interactional resources as central to mathematical discourse around dynamic geometry was proposed by Öner (2013). This paper cited a number of distinctions drawn in the CSCL literature for contrasting social/collaborative/relational resources with content-related resources:

- Text chat versus shared-whiteboard graphics (Çakir, Zemel & Stahl, 2009);
- Building a joint problem space (JPS) versus solving a problem (Roschelle & Teasley, 1995);
- A relational space versus a content space (Barron, 2000);
- Diachronic content versus temporal dimensions of the JPS (Sarmiento & Stahl, 2008);
- Project discourse versus mathematical discourse (Evans et al., 2011);
- Spatio-graphical observation (SG) versus technical reflection (T) (Laborde, 2004).

Öner then generated some data to explore the interaction of the contrasting dimensions by having two people work together face-to-face in front of a shared computer on a particular dynamic-geometry problem whose solution required a mix of spatio-graphical observation and technical reflection involving mathematical theory—a mix of SG and T resources, to use the distinction she adopted from Laborde.

Inspired by Öner’s experiment, Stahl (2013c) presented the same dynamic-geometry problem to two groups of people collaborating online in a CSCL system. We will review the sorts of resources that occur in the data generated in these two experiments after first considering the theoretical notion of resources as connections between levels.

Consider highway ramps or bridges used as resources for connecting road levels or landmasses. While we are more interested in linguistic interactional resources in this paper, it may be helpful to first consider the more intuitive physical case. A ramp or bridge often creates a possibility that did not otherwise exist for going from one level to another at a given point. To go from a local road to a limited-access superhighway, one must first find an available on-ramp. To cross a river from one side to the other, one may need a bridge. This is the individual driver’s view. From a different vantage point—the perspective of the resource itself—the creation of a ramp or the building of a bridge “affords” connecting the levels (Dohn, 2009).

By “affords,” we do not simply mean that the connecting is a happy characteristic or accidental attribute of the bridge, but that the bridge, by its very nature and design, “opens up” a connection, which connects the banks of the river it spans. In his early work, Heidegger (1927/1996) analyzed how the meaning of a tool was determined by the utility of the tool to the human user, within the network of meaning associated with that person’s life and world. In his later writings, Heidegger (1935/2003) shifted perspective to focus on things like bridges, paintings, sculptures, pitchers and temples in terms of how they themselves opened up new worlds, in which people could then dwell—opening new opportunities or possibilities for living. In considering the intersubjective world in which collaboration takes place on multiple connected levels, we might say that the work of resources like bridges is to contribute the spanning of shores within the way that the world through which we travel together is opened up as a shared landscape of resources for discourse and action.

This transformation of perspective away from a human-centered or individual-mind-centered approach became characteristic for innovative theories in the second half of the 20th Century. It is a shift away from the individualistic, psychological view to a concern with how language, tools and other resources of our social life work. It is a post-cognitive move since it rejects the central role of mental models, representations and computations. The things themselves have effective affordances; it is not just a matter of how humans manipulate models in which the things are re-presented to the mind. In phenomenology, Husserl (1929/1960) called for a return to “the things themselves” (*die Sache selbst*) and Heidegger (1950/1967) analyzed “the thing” (*das Ding*) separate from our representation of it. In ethnomethodology, Garfinkel and Sacks (1970) followed Wittgenstein’s (1953) linguistic turn to focus on the language games of words and the use of conversational resources (Koschmann, Stahl & Zemel, 2004). In distributed cognition, Hutchins (1996) analyzed the

encapsulation of historical cognition in cultural artifacts. In actor-network theory, Latour (2007) uncovered the agency of various kinds of objects in how they move across levels in enacting social transformations. Our use of the term “resources” in the 21st Century is intended to carry forward these groundbreaking approaches into the study of how the various planes of human interaction are connected. Vygotsky (1930/1978) used the term “artifact” to refer to both tools and language as mediators of human cognition; we prefer to use the broader term “resource” as it is frequently used in sociocultural analysis (Furberg, Kluge & Ludvigsen, 2013; Linell, 2001) for references brought into discourse. Like artifacts, resources are identifiable units of the physical or linguistic world that are involved in meaning-making practices—spanning the classical mind/body divide.

A central research issue for CSCL is how collaborative knowledge building takes place. The main problem seems to be to understand the role of individual cognition and of societal institutions in the small-group meaning-making processes. At ICLS 2000, Stahl (2000) presented a diagram that was intended more to raise this question than to answer it. In this diagram, “cultural artifact” served to connect the three planes of meaning-making processes. The diagram was based on an eclectic combination of major theories influential in CSCL. It is now time to conduct empirical investigations of the connections suggested by these theories.

In recent years, the Virtual Math Teams (VMT) Project (Stahl, 2009; Stahl, Mantoan & Weimar, 2013) has conducted case studies of small-group interaction. In doing so, it has tried to focus exclusively on the small-group unit of analysis. It has done so based upon three observations:

1. That most CSCL studies have focused either on the individual (cognitive) plane or on the community (practices) plane. For instance, they code utterances of individuals (Strijbos & Stahl, 2007) and reduce interaction to contributions of individuals or else they view interaction as participation in community processes and institutions.
2. That the small-group unit is fundamental to learning; as Vygotsky (1930/1978) said, one learns most human skills in social interaction first, only then being able to do so individually.
3. That the multiple levels are so complexly intertwined that it is hard to imagine studying them all together without first understanding much of what takes place at each level, temporarily taken on its own.

A number of studies have recently analyzed the problem-solving activities of virtual math teams (Stahl, 2009). In these studies, the interaction of students is analyzed at the small-group unit of analysis as a sequential progression. The collaborative knowledge-building activity that takes place there is mediated by a variety of interactional resources.

The theory sketched in this paper is not meant to reify different levels or processes, but to suggest some of the constraints between different phenomena and possible flows of influence. The distinctions between levels and the identification of typical processes at each level are intended to operationalize an infinitely complex and subtle matter for purposes of concrete analytic work by CSCL researchers. We propose the term “resource” to name the entities that are involved in mediating these connections.

In the work of small groups typical in CSCL, the sequential interaction brings in resources from the individual, small-group and community planes and involves them in procedures of shared meaning making. This interaction requires co-attention to the resources and thereby shares them among the participants. The process results in generating new or modified resources, which are then retained at the various planes. The resources that are brought in and those that are modified or generated often take the form of designed physical artifacts and sedimented elements of language. In other words, “small groups are the engines of knowledge building. The knowing that groups build up in manifold forms is what becomes internalized by their members as individual learning and externalized in their communities as certifiable knowledge” (Stahl, 2006, p. 16).

Each of the sentences of the preceding paragraph could be taken as a research question: a hypothesis about how levels are connected and an agenda for exploration. The following sections begin that undertaking. They present examples of interactional resources in small-group discussions of dynamic-mathematics problem solving and then describe some illustrative resources that are being prepared to help students engage in collaborative dynamic-mathematics problem solving.

The Analysis of Connecting Levels

An early attempt within CSCL to present an extended argument for the centrality of the small-group unit of analysis appeared in (Stahl, 2006), with a preliminary draft in (Stahl, 2004). These lengthy discussions were grounded in a half-minute interaction among four students working with a computer simulation of model rockets. The excerpt involved the students coming to understand how to interpret a textual resource: a table of rocket components arranged to facilitate comparisons among differently configured rockets. At first, none of the students could see the designed affordance of the table, but after the half-minute, they could all see the shared artifact as a resource for their scientific discourse. The interaction analysis of this excerpt showed how aspects of the table artifact were brought in as resources for the group discourse; as were shared and repeated words like “*same*” and “*different*.” The words of the dominant student, Chuck, were brought into the interaction by others in order to re-orient Chuck to a new, shared understanding of the co-attended-to table. The resource that

emerged for the group's subsequent practice was a sophisticated understanding of the organization of the table (Stahl, 2006, Ch. 12 & 13). This locally achieved understanding was congruent to a standard scientific understanding, which the instructor had assumed in designing the table and offering it as a resource for the group task. Here we can see the use of interactional resources connecting ideas from novice *individual* and scientific *community* planes in the small-group discourse, which led to a significant advance in the *group's* meaning-making ability.

In the experiment reported by Öner (2013), two graduate students work on a dynamic-geometry task, using a shared computer running Geometer's Sketchpad software. The task was specifically selected because it tends to make visible a combination of exploring a figure to discover its dependencies and then duplicating the figure using those dependencies. Thus, it involved a combination of spatio-graphical (SG) observation and theoretical (T) mathematical construction. The task was to duplicate a given figure, consisting of an equilateral triangle inscribed within another equilateral triangle.

At the start of the group's work on this task, one of the students, Ayla, says, "*Are these equal, these distances?*" The group then points to and measures the short segments along the outer triangle up to the interior triangle, which look about the same length. They confirm that these line segments (EC, AD and FB) are always of equal length, even when the figure is dragged and the lengths change. The similar appearance of the three segments in the graphical view provides a perceptual resource, which Ayla brings into the discourse and points to both with her statement and with her finger on the computer screen, establishing co-attention to this resource.

Later, at the crucial point in the construction at which a second vertex of the inner triangle is to be specified, the earlier finding about the original figure is recalled as a resource for duplicating it. As Mete goes to position the second vertex on segment AB, Ayla points to segment EF and wonders quietly as if to herself, "*Hmm, the distance does not have to be always equal.*" Then she says aloud, "*Does it? Look, EC and AD and FB are always equal in length,*" while pointing at the three segments on the screen. Mete immediately responds, "*Ha. Then we'll do the thing; we'll measure that gap,*" and begins to do the corresponding construction. This is an instance of group memory (Sarmiento & Stahl, 2007), in which the group references a previous finding and re-situates it in the current interactional context, providing it as a shared resource for the current work. The subsequent 30 speech turns of the dyad are concerned with figuring out how to use the software tools to construct their equivalents of EC, AD and FB to be equal lengths. Geometer's Sketchpad provides a tool to do this simply in a couple of ways. However, the resourcefulness of the tool has to be reconstructed by the group interaction to be a usable and effective resource for the group effort. The reconstruction effort itself takes advantage of various interactional resources, such as the letters labeling the triangle vertices, which the group discusses in order to simplify the work of relating corresponding points between their duplicated figure and the original.

Examples of resources from the (Öner, 2013) analysis include those classified as theoretical (T)—such as the geometry problem, the software tools or the relevant concepts, definitions, axioms and theorems of geometry. There are also spatial-graphical (SG) resources—including various visual properties of the figure like segment lengths and point labels.

The experiment with reproducing the inscribed equilateral triangles was replicated within the VMT Project (Stahl, 2013c). Two teams (A and B) of three adults each spent about a half an hour in the online VMT collaboration environment including multi-user GeoGebra. The software supported text chat with graphical referencing and dynamic-geometry construction, providing a contrast to the face-to-face speech and finger pointing in the Öner scenario. The task was identical to Öner's, implying that many of the resources for group work were identical: concepts and theorems of geometry (to the extent that the participants had working knowledge of them) and the visual properties of the figure (as it was dragged in the dynamic-geometry software display).

Although Team A in the VMT experiment focused on observing the spatio-graphical behavior of the points under dynamic dragging, it took them a long time to make Ayla's key observation. Finally, Jan said, "*So I think F is CD units away from B on BC. Its not constructed as an equilateral triangle, it happens to be an equilateral triangle because of the construction.*" Here, the SG observation leads immediately to a T statement about the construction of the internal triangle, namely that it is not constructed by making its sides or angles equal, but rather their equality is a consequence of imposing a different dependency involving distances of the vertices of the interior triangle from those of the exterior triangle. Visual resources are turned into resources for construction and reflection.

Team B took even longer to arrive at the key observation for constructing the inscribed triangles. They pursued multiple strategies, such as using geometric theorems about centers of triangles and correspondences of similar triangles. Finally Lauren said, "*I abandoned the center, and worked with the lengths of the sides.*" Then she "*used the compass tool to measure the distance from D to C*" and constructed the circles around the two other vertices of the exterior triangle, each with radius equal to CD to locate the vertices for constructing the interior triangle.

The use of social conventions and other relationship-building resources in addition to the content-oriented phases of chats seem to play an important role in problem-solving interactions. As Mercer and Sams (2006, p. 517) put it, “while working in classroom groups, children use talk to do much more than engage in curriculum tasks: they form relationships, develop social identities, and pursue ‘off-task’ activities which may be more important to them than the tasks in which they officially engaged—and as Wegerif (2005) has argued, may be essential to the process of establishing good relationships so that effective ‘on-task’ activities result.” The use of social-discourse resources to build group cohesion may be even more pronounced, salient and varied in online interactions, which lack some of the social resources provided by physical presence.

Groups in CSCL contexts can be seen to be making considerable use of resources to accomplish their interactional work. Often, they bring in resources from their individual backgrounds or from a community plane (the classroom, the history of mathematics, the subculture of social texting, the practices common in society, the conventions of ordinary language). Frequently, they build local resources within the group, available for repeated use and for “internalization” into resources for the individuals or for “externalization” into disseminated resources for the larger community.

The resources must be shared—attended to by the group and similarly understood—for them to be effectively used. This may be achieved through pointing, questioning, explaining, drawing and illustrating (Stahl et al., 2011). In a problem-solving session, one of the first resources co-constructed by the group might be a formulation of the question that they will pursue, based perhaps on an assigned task, which they must understand and articulate collaboratively (Zemel & Koschmann, 2013). The use of resources can be accumulated in the sequentiality of interaction to produce larger group-cognitive accomplishments such as mathematical problem solving (Stahl, 2011). Across a somewhat larger time scale, resources can build on one another, much as Euclid’s proofs built upon previous proofs. Groups can use their earlier formulations of interactional resources to construct higher-level resources and to refine previous understandings, just as scientific knowledge advances by accumulation and revision (Kuhn, 1972). In each case, the group must enact the resource, coming to a shared understanding of it and situating it in the group-discourse context for it to function as a resource for them. In this sense, resources are emergent from the group interaction.

CSCL research can connect the levels in its research data by identifying the resources that are being enacted in collaborative interactions and by tracking how they are constituted, understood and applied in the meaning-making process. CSCL studies can contribute to our understanding of collaborative meaning-making processes by providing detailed analyses of the ways in which group discourses involve resources interactionally and how the resources are shared, interpreted, refined and preserved.

The Pedagogy of Connecting Levels

If resources play such an important role in collaborative learning, then how can CSCL designers support the use of resources? Clearly, it would be useful to make sure that students have access to relevant resources and that they understand how to use them. In situations where teachers play a central role in guiding the collaborative learning, it would similarly be important to ensure that the teachers have access to relevant resources and that they understand how to facilitate student use of them. Early attempts to support CSCL resources for teachers and students were proposed in (Stahl, Sumner & Owen, 1995) and (Stahl, Sumner & Reppenning, 1995).

In the Virtual Math Teams (VMT) Project, we have learned through pilot trials of the VMT-with-GeoGebra environment that this relatively complex system requires careful preparation and training for teachers, students, online groups and classes to use effectively without encountering frustration. In response to this, we have drafted a set of dynamic-geometry curricular activities, interspersed with tutorial tours of the technology features (Stahl, 2012a). These materials are designed for use both by teachers in professional-development contexts and by student teams in online-classroom or after-school settings.

The VMT curriculum activities have been designed to promote collaborative learning, particularly as it occurs in significant mathematical discourse about geometry. We do this by providing a carefully structured set of resources for use by teachers and students. These include the following:

1. Resources for engaging in significant mathematical *discourse*; to collaborate on and discuss mathematical activities in supportive small online groups. This includes suggested uses of linguistic and interactional resources for coordinating collaboration, as well as tutorials in using the communication tools of the VMT software.
2. Resources to collaboratively *explore* mathematical phenomena and dependencies; to make mathematical phenomena visual in multiple representations; and to vary their parameters. This includes scaffolded exercises in noticing visual characteristics of dynamic-geometric figures being dragged and in wondering in chat postings about their dependencies.
3. Resources for *constructing* mathematical diagrams—understanding and exploring their structural dependencies. This consists primarily of a semester-long sequence of construction activities, initially with step-by-step instructions and tutorials about GeoGebra tools.

4. Resources to notice, wonder about and form conjectures about mathematical relationships; to justify, explain and *prove* mathematical findings. This involves discussion prompts and situated examples of explanations or proofs.
5. Resources to understand core concepts, relationships, theorems and constructions of basic high-school *geometry*. The included materials and activities cover central conceptual and procedural resources from Euclid's first book of propositions and from the Common Core standards for beginning geometry.

The presentation of resources is organized developmentally, so that understanding of the resources presented first can be used to build understanding of resources presented subsequently. Concomitant with this is a progressive shift from scaffolded explanation of basic resources (like software tools) to open-ended inquiry of more complex resources (like mathematically interesting micro-worlds).

There is a theoretical basis for gradually increasing skill levels in terms of both geometric understanding and deductive proof. The van Hiele theory (see deVilliers, 2003, p. 11) specifies several levels in the development of students' understanding of geometry resources. The implication of van Hiele's theory is that students who are at a given level cannot properly grasp ideas presented at a higher level until they work up to that higher level. That means that unprepared groups will fail to enact available resources in a meaningful way. Thus, a developmental series of activities pegged to the increasing sequence of levels is necessary to effectively present the various resources of geometry, such as, eventually, the formal structure of deductive proof. Failure to lead students through this developmental process is likely to reinforce student feelings of inadequacy and consequent negative attitudes toward geometry.

A particularly important resource for understanding and working in dynamic geometry is the concept of *dependency*. GeoGebra allows one to construct systems of inter-dependent geometric objects. The dependencies built into dynamic-geometry constructions are intimately related to proofs illustrated by those constructions. Often, to understand a dependency and to be able to implement it in a construction is tantamount to being able to articulate a proof and to explore its validity dynamically (Stahl, 2013c). Students have to learn how to think in terms of these dependencies. They can learn through use of resources like visualizations, manipulations, constructions and verbal articulations. These can all be modeled by examples, and these resources can be provided gradually.

The VMT Project is now drafting and piloting versions of curricular activities designed to develop significant mathematical discourse focused on dependencies among geometric objects (Stahl, 2012a). Concomitantly, it is implementing software support for teachers and students to explore the dependencies and assembling materials for professional development to prepare teachers to enact this curriculum with their students (Stahl & Powell, 2012). The set of activities is designed to provide the most important basic geometry resources to math teachers and students, taking them from a possibly novice level to a more skilled level, at which they will have a sufficient portfolio of resources for engaging in significant mathematical discourse without continuing scaffolding. The resources of classical Euclidean geometry were decisive in the historical development of rational thinking by literate individuals and of scientific culture in the modern world (Netz, 1999; Stahl, 2013c). We hope to adapt these resources to the CSCL context, where they may enter into small-group collaborative online interactions and thereby influence both individual understanding and classroom practices.

In on-going experiments within the VMT Project and elsewhere, our colleagues and we will be logging the use of the resources by teachers and students in order to analyze how resources connect levels of learning in a CSCL setting. We will track individual and group performance in significant mathematical discourse as resources and practices from community levels are taken up in sequential small-group interaction. Perhaps we will witness the formation of local practices and group interactional resources, which can influence individual and community levels over time. In these ways, we will study resources for connecting levels of learning in CSCL. More generally, through analysis of the nature and work of resources in case studies of a broad variety of CSCL interactions, the CSCL research community can expect to reach a better understanding of the nature of different levels of analysis in CSCL research and how the levels may be connected in terms of their mediation by diverse resources. Gradually, we will discover how resources are enacted, understood, shared, designed, adapted and preserved—and how they mediate connections among levels of learning through social interaction.

References

- Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of The Learning Sciences*. 9(4), 403-436.
- Çakir, M. P., & Stahl, G. (2013). The integration of mathematics discourse, graphical reasoning and symbolic expression by a virtual math team. In D. Martinovic, V. Freiman & Z. Karadag (Eds.), *Visual mathematics and cyberlearning*. New York, NY: Springer. Web: <http://GerryStahl.net/pub/visualmath.pdf>.

- Çakir, M. P., Zemel, A., & Stahl, G. (2009). The joint organization of interaction within a multimodal CSCL medium. *International Journal of Computer-Supported Collaborative Learning*, 4(2), 115-149. Web: http://GerryStahl.net/pub/ijCSCL_4_2_1.pdf.
- deVilliers, M. (2003). *Rethinking proof with the Geometer's Sketchpad*. Emeryville, CA: Key Curriculum Press.
- Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines: Towards an interdisciplinary learning science*. (pp. 189-211). Oxford, UK: Elsevier.
- Dohn, N. B. (2009). Affordances revisited: Articulating a Merleau-Pontian view. *International Journal of Computer-Supported Collaborative Learning*, 4(2), 151-170.
- Evans, M. A., Feenstra, E., Ryon, E., & McNeill, D. (2011). A multimodal approach to coding discourse: Collaboration, distributed cognition, and geometric reasoning. *International Journal of Computer-Supported Collaborative Learning*, 6(2), 253-278.
- Furberg, A., Kluge, A., & Ludvigsen, S. (2013). Students' conceptual sense-making with and of science diagrams in computer-based inquiry settings. *International Journal of Computer-Supported Collaborative Learning*, 8(1)
- Gardner, H. (1985). *The mind's new science: A history of the cognitive revolution*. New York, NY: Basic Books.
- Garfinkel, H., & Sacks, H. (1970). On formal structures of practical actions. In J. Mckinney & E. Tiryakian (Eds.), *Theoretical sociology: Perspectives and developments*. (pp. 337-366). New York, NY: Appleton-Century-Crofts.
- Heidegger, M. (1927/1996). *Being and time: A translation of Sein und Zeit* (J. Stambaugh, Trans.). Albany, NY: SUNY Press.
- Heidegger, M. (1935/2003). Der ursprung des kunstwerkes. In M. Heidegger (Ed.), *Holzwege*. Frankfurt a. M., Germany: Klostermann.
- Heidegger, M. (1950/1967). Das ding. In *Vorträge und aufsätze ii*. (pp. 37-60). Pfullingen, Germany: Neske.
- Husserl, E. (1929/1960). *Cartesian meditations: An introduction to phenomenology* (D. Cairns, Trans.). The Hague, Netherlands: Martinus Nijhoff.
- Hutchins, E. (1996). *Cognition in the wild*. Cambridge, MA: MIT Press.
- Koschmann, T., Stahl, G., & Zemel, A. (2004). *The video analyst's manifesto (or the implications of Garfinkel's policies for the development of a program of video analytic research within the learning sciences)*. Paper presented at the International Conference of the Learning Sciences (ICLS 2004). Los Angeles, CA. Proceedings pp. 278-285. Web: <http://GerryStahl.net/pub/manifesto2004.pdf>.
- Kuhn, T. (1972). *The structure of scientific revolutions (2nd ed.)*. Chicago, IL: University of Chicago Press.
- Laborde, C. (2004). The hidden role of diagrams in pupils' construction of meaning in geometry. In C. H. J. Kilpatrick, & O. Skovsmose (Ed.), *Meaning in mathematics education*. (pp. 1-21). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Latour, B. (2007). *Reassembling the social: An introduction to actor-network-theory*. Cambridge, UK: Cambridge University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Linell, P. (2001). *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives*. New York, NY: Benjamins.
- Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve maths problems. *Language and Education*, 20(6), 507-528.
- Netz, R. (1999). *The shaping of deduction in Greek mathematics: A study in cognitive history*. Cambridge, UK: Cambridge University Press.
- Öner, D. (2013). Analyzing group coordination when solving geometry problems with dynamic geometry software. *International Journal of Computer-Supported Collaborative Learning*, 8(1)
- Rogoff, B. (1995). Sociocultural activity on three planes. In B. Rogoff, J. Wertsch, P. del Rio & A. Alvarez (Eds.), *Sociocultural studies of mind*. (pp. 139-164). Cambridge, UK: Cambridge University Press
- Roschelle, J., & Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer-supported collaborative learning*. (pp. 69-197). Berlin, Germany: Springer Verlag.
- Sarmiento, J., & Stahl, G. (2007). *Bridging and persistence in sustained, collaborative problem solving online*. Paper presented at the Hawaii International Conference on System Sciences (HICSS 2007). Hawaii, HI. Web: <http://GerryStahl.net/pub/hicss07>.
- Sarmiento, J., & Stahl, G. (2008). *Extending the joint problem space: Time and sequence as essential features of knowledge building*. Paper presented at the International Conference of the Learning Sciences (ICLS 2008). Utrecht, Netherlands. Web: <http://GerryStahl.net/pub/icls2008johann.pdf>.
- Schatzki, T. R., Knorr Cetina, K., & Savigny, E. v. (Eds.). (2001). *The practice turn in contemporary theory*. New York, NY: Routledge.

- Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses and mathematizing*. Cambridge, UK: Cambridge University Press.
- Stahl, G. (2000). *A model of collaborative knowledge-building*. Paper presented at the Fourth International Conference of the Learning Sciences (ICLS '00). Ann Arbor, MI. Proceedings pp. 70-77. Lawrence Erlbaum Associates. Web: <http://GerryStahl.net/pub/icls2000.pdf>.
- Stahl, G. (2004). Building collaborative knowing: Elements of a social theory of CSCL. In J.-W. Strijbos, P. Kirschner & R. Martens (Eds.), *What we know about CSCL: And implementing it in higher education*. (pp. 53-86). Boston, MA: Kluwer Academic Publishers. Web: <http://GerryStahl.net/cscl/papers/ch16.pdf>.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press. Web: <http://GerryStahl.net/mit/>.
- Stahl, G. (2008). Thinking as communicating: Human development, the growth of discourses and mathematizing. *International Journal of Computer-Supported Collaborative Learning*. 3(3), 361-368. Web: <http://GerryStahl.net/pub/Sfardreview.pdf>.
- Stahl, G. (2009). *Studying virtual math teams*. New York, NY: Springer. Web: <http://GerryStahl.net/vmt/book>.
- Stahl, G. (2011). *How a virtual math team structured its problem solving*. Paper presented at the Connecting computer-supported collaborative learning to policy and practice: CSCL 2011 conference proceedings. Lulu: ISLS. Proceedings pp. 256-263. Web: <http://GerryStahl.net/pub/cscl2011stahl.pdf>.
- Stahl, G. (2012a). *Dynamic-geometry activities with GeoGebra for virtual math teams*. Web: <http://GerryStahl.net/vmt/activities.pdf>.
- Stahl, G. (2012b). Traversing planes of learning. *International Journal of Computer-Supported Collaborative Learning*. 7(4), 467-473.
- Stahl, G. (2013a). Learning across levels. *International Journal of Computer-Supported Collaborative Learning*. 8(1), 1-12.
- Stahl, G. (2013b). Theories of collaborative cognition: Foundations for CSCL and CSCW together. In S. Goggins & I. Jahnke (Eds.), *CSCL@work*. (Vol. #13 Springer CSCL Book Series). New York, NY: Springer. Web: <http://GerryStahl.net/pub/collabcognition.pdf>.
- Stahl, G. (2013c). *Translating Euclid: Creating a human-centered mathematics*: Morgan & Claypool Publishers. Web: <http://GerryStahl.net/pub/translating.pdf>.
- Stahl, G., Jeong, H., Sawyer, R. K., & Suthers, D. D. (2012). Workshop: Analyzing collaborative learning at multiple levels. Presented at the International Conference of the Learning Sciences (ICLS 2012), Sydney, Australia. Web: <http://GerryStahl.net/pub/icls2012workshop.pdf>.
- Stahl, G., Mantoan, A., & Weimar, S. (2013). Demo: Collaborative dynamic mathematics in virtual math teams. Presented at the International Conference of Computer-Supported Collaborative Learning (CSCL 2013), Madison, WI. Web: <http://GerryStahl.net/pub/cscl2013demo.pdf>.
- Stahl, G., & Powell, A. B. (2012). *Multi-user GeoGebra for virtual math teams*. Paper presented at the GeoGebra ICME Pre-conference. Seoul, Korea. Web: <http://GerryStahl.net/pub/ggbkorea.pdf>.
- Stahl, G., Sumner, T., & Owen, R. (1995). Share globally, adapt locally: Software to create and distribute student-centered curriculum. *Computers and Education. Special Issue on Education and the Internet*. 24(3), 237-246. Web: <http://GerryStahl.net/cscl/papers/ch05.pdf>.
- Stahl, G., Sumner, T., & Repenning, A. (1995). *Internet repositories for collaborative learning: Supporting both students and teachers*. Paper presented at the International Conference on Computer Support for Collaborative Learning (CSCL '95). Bloomington, Indiana. Proceedings pp. 321-328. ACM Press. Web: <http://GerryStahl.net/cscl/papers/ch06.pdf>.
- Stahl, G., Zhou, N., Cakir, M. P., & Sarmiento-Klapper, J. W. (2011). Seeing what we mean: Co-experiencing a shared virtual world. In *Connecting computer-supported collaborative learning to policy and practice: CSCL 2011 conference proceedings*. (Vol. I, pp. 534-541). Lulu: ISLS. Web: <http://GerryStahl.net/pub/cscl2011.pdf>.
- Strijbos, J. W., & Stahl, G. (2007). Methodological issues in developing a multi-dimensional coding procedure for small group chat communication. *Learning & Instruction. Special issue on measurement challenges in collaborative learning research*. 17(4), 394-404. Web: <http://GerryStahl.net/vmtwiki/jw.pdf>.
- Suthers, D., Lund, K., Rosé, C. P., & Law, N. (2013). *Productive multivocality*. Cambridge, MA: MIT Press.
- Vygotsky, L. (1930/1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Wegerif, R. (2005). Reason and creativity in classroom dialogues. *Language and Education*. 19(3), 223-237.
- Wittgenstein, L. (1953). *Philosophical investigations*. New York, NY: Macmillan.
- Zemel, A., & Koschmann, T. (2013). Online math problem solving as a process of discovery in CSCL. *International Journal of Computer-Supported Collaborative Learning*. 8(1)

Learning with Collaborative Inquiry: a Science Learning Environment for Secondary School Students

Daner Sun, Chee-Kit Looi, Learning Sciences Laboratory, National Institute of Education, Singapore
daner.sun@nie.edu.sg, cheekit.looi@nie.edu.sg,
Evelyn Teo, School of Science and Technology, Singapore, evelyn_teo@sst.edu.sg

Abstract: The paper presents a study of science instruction using collaborative inquiry with the CSI (Collaborative Science Inquiry) system which combines multiple learning design features with CSCL design elements. The study reported here explores the educational value of the system on students' conceptual understanding, performance on inquiry phases and collaborative work. Promising results have been received from the comparison of pre-test and post-test achievements, the examination of artefacts, peer discussions, and interview transcripts. The results indicate that the collaborative inquiry facilitated by the CSI system can engage students in activities and promote their conceptual understanding in a progressive way.

Introduction

Computer Supported Collaborative Learning (CSCL) applications incorporating design elements such as shared (work) space, chat tools, collaborative text editor, and argumentative editor, have been identified as being capable of facilitating collaborative learning in pupils (Bouyias & Demetriadis, 2012; Gogoulou, et al., 2008). Concerning the benefits addressed with the use of CSCL approach, the study reported here aims to implement a web-based science-learning environment employing collaborative inquiry approach to facilitate science instruction and learning at secondary schools in Singapore. The system, known as Collaborative Science Inquiry (CSI) science learning environment, leverages both guided inquiry and modelling with multiple CSCL components. The conceptualization of the learning design for the system is influenced by existing design principles and relevant applications, such as WISE, CMapTools, Co-Lab, and ModelingSpace, which seek to integrate appropriate pedagogies with appropriate CSCL design elements.

In the CSI system, learning designs included guided inquiry, modelling and visualization, have been proved to be contributing positively to science learning (Buckley et al., 2004; Jackson, et al., 2008). The guided inquiry process is developed and modified from dominant model-based inquiry principles (Bell et al. 2010; White et al. 2002). It consists of eight phases: Contextualize, Questions and Hypothesis (Q&H), Pre-model, Plan, Investigate, Model, Reflect, and Apply. Modelling pertains to the construction of scientific models by the use of drawing-based, qualitative and quantitative modelling tools (Lerner, 2007). These modelling tools are embedded in both the Pre-model and Model phases. The system allows the display of various visualizations, such as images, videos, and dynamic simulations, to support virtual inquiry. Multiple CSCL design elements, including synchronous modelling and editing, shared workspace, peer review, chat tool, and social presence, are also integrated in each inquiry phase as required. Thus, the unique feature of the system is the tight coupling of relevant CSCL design at each inquiry phase, such that each phase can be utilised in a flexible way towards inquiry learning through modelling and visualization.

To validate the educational value of the system, a series of pilot studies, each with different educational aims, have been conducted (Sun & Looi, 2013). Due to the continuous efforts made in system development, we have further continued the exploration of the deployment methods of the system in various classroom settings. Hence in this study, we interpret a study of trial instruction of the CSI system in a biology class at Secondary 1 (Grade 7) level. The aims of the study are: 1) Demonstrating the effect on promoting conceptual understanding of abstract concepts associated with the system. 2) Exploring the progressive process of students' conceptual changes at every stage of the inquiry process. 3) Summarizing students' performance on peer discussion which contributes to the accomplishment of collaborative artefacts.

System Overview

General Structure

There are two major functional modules in the system: teacher module and student module. The teacher module consists of six sections: Profile, Subject Management, Project Management, Simulation Library, Solutions Review, and Mailbox. An authoring tool is available for teachers to design lessons when they enter the various sections. Project Management enables teachers to set up the inquiry project, and the stages by creating the project and tasks, posing guiding questions and configuring student groups. Simulation Library allows for importing visualizations (e.g. Java applets, videos, flash applications) for projects. Solutions Review facilitates reviewing and evaluating students' artefacts (e.g. answers, models, reflections) and their chat logs. The student module consists of four sections: Profile, My Project, Group Management, and Mailbox. As the core

component, My Project allows students to access the assigned project to conduct inquiry activities and complete a series of tasks with their group members. The tasks may include reading and discussing textual information, proposing and negotiating solutions, manipulating and observing simulations, responding to guiding questions, constructing models, and writing reflections at the assigned inquiry phases. The system supports the inquiry either in a linear or non-linear manner. Students can switch between phases easily by clicking tabs on the tool bar. The project work session at student module is presented in Figure 1. The CSCL design elements are annotated (see the following introduction of CSCL design features).

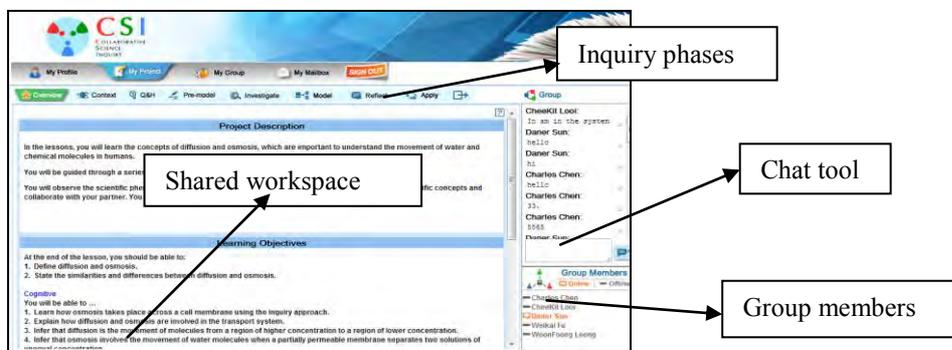


Figure 1. General Interface of the Project Work Session in Student Module

CSCL Design Features

The system employs multiple CSCL design elements in inquiry phases to facilitate students' various forms of collaboration as mentioned above (Pozzi, 2009). Figure 2 depicts a map representing how CSCL design elements are combined and integrated in the system. In Overview and Contextualize, all online members can share the text information. Besides peer review and mutual editing, students are allowed to edit and revise their answers synchronously in the private editing box in Q&H, Plan, and Reflect. Furthermore, the shared workspace in Pre-model and Model is able to receive inputs from multiple devices to permit concurrent multi-user operations (Yang & Lin, 2010), such as co-constructing models, mutually reviewing and revising models in real time. The design intends to encourage students to pursue the common goal of creating joint models through a collaborative and interactional process. The system allows peer review of the individual models within the private modelling space. Coupled with a chat tool, each phase supports students' peer discussion synchronously.

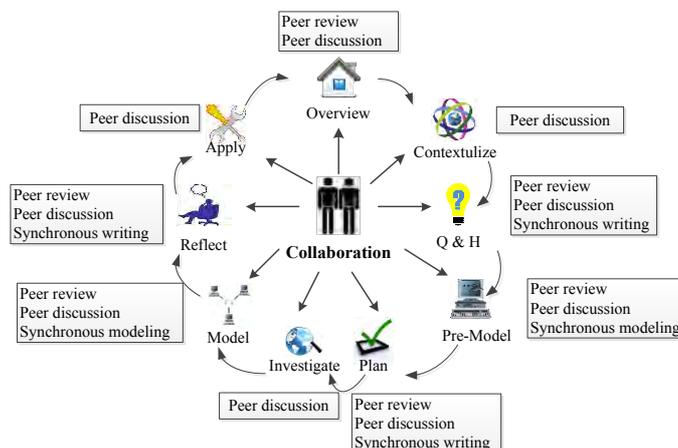


Figure 2. CSCL Design Elements Integrated in the Inquiry Phases within the CSI System

The system also integrates other CSCL design elements to facilitate ease of use. The task checklist in Overview facilitates students within the same group to keep track of their progress in carrying out activities in real time. The online member window can display students' social presence to facilitate coordination and collaboration between students from different spatial locations (Janssen, et al., 2007). An email box available in both the teacher and the student modules is used for exchanging of ideas, written materials and other information.

Methods

Participants

The trial instruction was conducted by four science teachers from a junior secondary school in Singapore. These teachers had rich and extensive teaching experiences and had attended most of regular meetings of CSI project. Hence, they had some understanding of the system development and its underlying pedagogy. A total of 201 secondary 1 (Grade 7) students from 9 classes participated in this study. During the instruction, students were mostly organized in pairs ($N_{\text{pair}}=96$), with only three groups working in triads. The school had excellent computer facilities, with each student owning and utilising a MacBook for daily lessons in the various subjects.

Lesson Design

The class studied the topic of “Diffusion and Osmosis”. Science teachers, researchers and collaborators co-designed and finalized the lessons. The study was conducted as two 50-min consecutive sessions for each class. The lesson sequence was executed in the following order: Overview → Contextualize → Q&H → Pre-model → Investigate → Reflect → Apply. After students reviewed the textual information (e.g. brief description of the project, learning objectives, tasks) in Overview, they were introduced to a story related to the topic in Contextualize. In Q&H, the students were engaged in forming their answers to the questions. In Pre-model, the students performed individual and collaborative modelling after observing two videos (a. the diffusion of red ink in the water; b. the changes of egg in different solutions). In Investigate, students interacted with three simulations and answered guiding questions. The simulations were: a) the diffusion of sugar in the water; b) the movement of water molecules in osmosis; c) a dynamic simulation for observing the results of osmosis. Finally, in Reflect and Apply, the students reflected, refined and validated their conceptual understanding. Except for the Contextualize and Apply, the activities in other phases were conducted in collaboration.

Data Sources and Data Analysis

The data sources included pre- and post-tests, field notes, observation sheets, interview transcripts, on-site videos and audio transcripts, learning artefacts, and chatting information (system log). The use of different data sources provided complementary information and enabled a more thorough and reliable understanding of students’ performance observed in CSI lessons. In data analysis, a paired-samples t-test was conducted to identify the difference between pre-test and post-test scores; an item-by-item analysis of the test responses was carried out to further expose misconceptions amongst students. Furthermore, students’ responses to Q&H and Apply, pre-models, and reflections were scrutinized to uncover the conceptual transformation process by the use of coding methods. The chatting were further analysed to probe students’ performance on collaborative work. Only data from the participating students in all sessions and activities was used in the analysis.

Pre-test and post-test

In the study, identical pre- and post-test instruments were used at the beginning and concluding stages of the lessons (10 minutes for each). The 10-paired questions in tests were built on the previously validated two-tier “Diffusion and Osmosis Diagnostic Test” (Odom & Barrow, 1995). In all question sets, the A questions asked for students’ direct answers to a given scenario (the “what” questions), while the B questions focused on students’ explanations to the answer in A question (the “why” questions). The questions covered all the content at the appropriate difficulty levels that the teachers expected the students to learn in the topic of “diffusion and osmosis”. The tests can be retrieved from: <https://sites.google.com/a/wimvt.info/wimvt/teacher-pedagogical-resources/Diffusion-and-Osmosis-Test>.

Q&H and Apply answers

A coding method was employed to assess the understanding levels of conceptions through categorizing answers in Q&H and Apply into five categories. The categorization was built on the knowledge integration scoring rubric (Linn & Eylon, 2011), which is an appropriate and effective way to obtain how students developed the existing ideas to more normative and coherent understanding (Liu et al., 2008). The categories were refined and modified as follows: L_1 : students have irrelevant ideas and make incorrect links between context and their explanations (incorrect answers); L_2 : students have relevant ideas and make partial correct links between context and their simple explanations (partial correct answers with simple explanations); L_3 : students have relevant ideas and make correct links between context and their simple explanations (correct answers with simple explanations); L_4 : students have relevant ideas and make links between context and their elaborated explanations (correct answers with elaborated explanations); L_5 : students have complete relevant ideas and make links between context and their elaborated explanations, as well as related contexts (correct answers with extended elaborated explanations). Frequency of each category were calculated and analysed through this coding approach.

Pre – models

To evaluate students’ modelling performance, we classified the quality of models into three levels based on a literature review (Ergazaki, et al., 2005; Grosslight, et al., 1991; Halloun, 1997): A. High Quality Models (H) refers to models containing accurate description of science concepts that involve objects with basic properties, and reflect interaction between objects; B. Medium Quality Models (M) refers to models with partially exact description of science concepts, which represent some of model components and the possible relations. C. Low

Quality Models (L) refers to models containing inaccurate description of all model components. If the models are built at the macroscopic level, they are marked as sublevel “1”, while models built at the particulate level are marked as sublevel “2”. For example, if the students draw high quality models at the particulate level, the models are coded as H₂. Furthermore, work completion is considered as another indicator for assessing students’ modelling performance.

Reflections

We coded the responses to Reflect into four categories: verification, explanation, improvement and critical reflection. The method was modified from the principle of reflective thinking (Kember, et al., 2000). “Verification” refers to the reflection with simple confirmation of the artefacts. “Explanation” focused on interpreting the definitions of the concepts, but without commenting on how to improve the artefacts. “Improvement” means the reflection expressing students’ ideas on how to improve their artefacts. “Critical reflection” pertains to those reflections that involve the critiques, and the proposals of improvement, as well as further explanation of the artefacts. Reflection from low-level to high-level thinking is ranked progressively from “verification” to “critical reflection”. The ranking of students’ reflection responses enabled researchers to probe the degree of students’ thinking and understanding of their work in the inquiry.

Peer discussions

We had extracted and analysed available peer discussions (taking one sentence as a unit) generated in the chat box to explore students’ performance on involvement and collaboration in each phase and to observe the process of conceptual understanding transformation. The method was developed and refined based on the principles of good feedback (Nicol & Macfarlane-Dick, 2006). Aligning with these principles, the peer discussions were classified into A. task-oriented, B. knowledge-oriented, C. strategy-oriented, D. assessment-oriented discourse and E. affection-oriented discourse. Category A clarifies the task specificities, such as procedures, duration, and work division. Category B provides necessary information relative to the key concepts, such as definition, explanations, and reasoning. Category C provides strategic methods to complete the task. Category D provides constructive comments on the work. Category E provides comments with intentions to improve motivation of group members.

Video and audio data were transcript and analysed to reveal students’ interactions in the class, as well as their learning performances in the activities. Furthermore, interviews with the teachers (n=4) and randomly selected students (n=16) were administrated to collect feedback on the system implementation. The data were coded and analysed independently by the first author and another researcher. The inter-rater reliability coefficient for these coding was $r = 0.93$.

Findings and Discussions

Pre-test and post-test achievements

In this section, 139 valid tests were received. A one-way ANOVA was conducted to compare the initial levels of the classes. The results indicates that students’ prior knowledge varied very little among the classes, as $F(7, 132) = 2.773$, $p = 0.01$ (the priori alpha level was set at .01). This implies that the students started the lessons with the equivalent cognition levels. Paired-samples t-test analysis indicates that the post-test scores were significantly higher than the pre-test, which received $M = 12.97$, $SD = 2.774$; $t(36) = -4.299$, $p = 0.000$, compared to the pre-test ($M = 10.62$, $SD = 2.792$). It suggests, in general, CSI lessons enhanced the understanding of diffusion and osmosis in most of the pupils.

The Item-by-item analysis shows major conceptual changes concentrated on responses on the reasoning of diffusion, the judgment of solution’s concentration, the identification of osmosis and the effect of osmosis (item 2B, 4A, 5B, 7A, 9A, and 9B). Specifically, a significant finding is that students attained great understanding of the mechanisms of the scientific phenomena after the lessons. In the pre-test, we found that approximately 40.8% of students failed to answer correctly the “why” questions while they succeeded in answering the “what” questions. It suggests that students often could predict the correct outcome but had relatively little understanding about the underlying mechanisms. This may further demonstrates that students’ knowledge on the key conceptions remained at the superficial level before the lessons. After the lessons, only 15.3% of the students continued to have difficulties in responding to “why” questions.

The high rate of failures to answer the questions on most distracters exposes the prevalent misconceptions in both pre-and post-tests. Before the lessons, the high rate of failures was caused by distracters distributed on most of the items (14 misconceptions were identified). After the lessons, the rate of failures due to distracters reduced dramatically, with only two items (item 3B and 4B) continuing to receive high rate of failure due to present distracters. Most of the students wrongly believed that the particles would cease moving when equilibrium concentrations had been reached. Hence, it reveals that students still had some degree of confusion of how the process of diffusion influences the particles’ movement, and vice versa.

Students’ performance on the inquiry

Q & H

In Q&H, Question 1 (Q_1) asked students to propose a reason for the smell of the cooked fishes from a distance (diffusion). Question 2 (Q_2) asked students to propose a reason on why drinking seawater killed the sailors faster than expected compared to not drinking any water at all (osmosis). In sum, the responses to Q_1 were usually more correct and more complete compared to responses to Q_2 . Q_1 received 38.6% of L_2 response and 33.3% of L_3 response, compared to Q_2 which received 56.1% of L_1 responses. The results seem to confirm our findings in pre-test that students had difficulty in reasoning scientifically and deeply about the basic processes of diffusion, because they provided (partially) relevant answers with simple explanations on the reasons for Q_1 . On the other, the result is also consistent with the initial findings which had already suggested that a large percentage of students struggled to comprehend osmosis and its mechanism. As osmosis contained more invisible attributes and processes, students failed to connect the macroscopic observations with the mechanism of osmosis at the microscopic level. Only a fraction of students (Q_1 with 14% of L_4 and Q_2 with 3.5% of L_4) managed to answer them correctly with elaborated explanations. An interesting finding was observed that although the general performance on responses to Q_1 was better than Q_2 , a fraction of students (3.5%) performed better in Q_2 compared to Q_1 . These cases mostly existed among groups in which the pairs interacted with each other frequently as we observed.

Pre-model

In this section, students generally responded to the individual modelling tasks positively with the high percentage of work completion (80%). However, 70% of the students failed to finish the collaborative models. Meanwhile, the division of labour on collaborative modelling was not equally distributed between group members. There are three possible reasons to these observations: 1) limited class time affected students' group modelling. 2) few opportunities were offered for students to participate in synchronously collaborative activities in previous lessons. 3) collaborative scripts were not provided in time by the teacher to guide and structure students' collaboration.

Data analysis of their resultant models suggests that most students constructed the individual models at the particulate level but varied on the model quality. It reveals the different representations that individual students had about the concept of diffusion and osmosis as found in Q&H. Positively, more than half of the students drew the middle quality of diffusion models ($M_2=54.8\%$). However, some issues were found: 1) the models lacked the necessary annotations of each model component; 2) most of the models described the result of diffusion with water and ink particles in a container, but failed to represent the process of how particles scattered over time; 3) the particles drawn were placed in an orderly arrangement in the container, which should not be the case.

Variation of model quality also existed among the osmosis models, with M_1 and M_2 , taking up 40%, and L_2 taking up 35%. The significant proportion of M_1 and M_2 models may indicate that a significant part of students, who had viewed and observed the videos, had acquired a more appropriate perception of the process of osmosis and developed more accurate knowledge about osmosis. However, some students failed to distinguish the functions of model components (e.g. the egg yolk membrane as the partially permeable membrane, and the net movement of water molecules through the partially permeable membrane) although they knew that osmosis would happen when the egg was immersed in the corn syrup, as inferred from the significant proportion of L_2 models. We also notice that students' active engagement in peer review and discussion of models resulted in the improvement of their prior knowledge of osmosis.

Reflect

Responses in the Reflect varied, consisting 30.28% of "verification", 23.33% of "explanation", 18.33% of "improvement" and 28.06% of "critical reflection". Although 30.28% of students reflected upon their artifacts through "verification", the rest of the reflections indicated some deep thinking of the artifacts. Students' "explanations" were mainly concentrated on: 1) providing supplementary comments on interpreting the process of diffusion and osmosis at particulate level. 2) writing the definition of diffusion and osmosis, in order to show their current understanding. 3) explaining the effects of diffusion and osmosis correctly. Students that gave the "explanation" reflections achieved better understanding of diffusion and osmosis, especially with respect to the understanding of the definitions, the movement of the particles as the physical basis, and the results of diffusion and osmosis. They could link the key concepts or terminologies with the context after receiving more knowledge from Investigate. Students that gave "improvement" reflection generally thought that they should revise and improve their artefacts, it means they realised the misconceptions they held in their prior knowledge. Students that gave "critical reflection" indicated that they succeeded in developing a more sophisticated understanding and presentation of the target concepts.

Apply

Three questions, Q_1 , Q_2 , and Q_3 , were provided for students to answer in the Apply phase. Q_1 concerned the possible outcome of placing the microorganism *Elodea* into the ocean. Q_2 was about the reasons on vegetables becoming soggy when salad dressing was applied. Q_3 asked for the explanations on killing weeds using salt

water. The results show that Q_1 and Q_3 received more L_2 (Q_1 : 30.3%, Q_3 : 30.5%) and L_3 answers (Q_1 : 40.5%, Q_3 : 34.6%), than L_1 answers (Q_1 : 8.9%, Q_3 : 15.3%). Furthermore, a number of students attained L_4 (Q_1 : 20.3%, Q_3 : 19.6%). The presence of L_4 answers means that these students managed to apply their knowledge learnt from the lessons in the new context. However, students seemed to have difficulties in understanding the nature of the dressing relative to the vegetables in Q_2 , as most of the students defined the dressing as the hypotonic solution compared to the cellular content in the vegetables (56.3% L_1 , 12.5% L_2 , 18.8% L_3 , 12.5% L_4). This issue indicates that it may not be advisable to assess students' new acquired understanding within a new context that has little connection with the information in Investigate. In this case, it is easier for students to compare between liquid solutions (e.g. ocean, salt water) with hypotonic, isotonic or hypertonic relationships. However, students were hesitant in linking their new knowledge between salad dressing (colloidal mixture), and vegetables (cellular matrix).

Students' performance on the collaborative work

Peer discussions by students were mainly categorised as task-oriented, knowledge-oriented, strategy-oriented, and assessment-related discourse based on our data analysis. The distribution in these categories in the specific inquiry phases depicted in Table 1.

Table 1. The percentage of the peer discussion in the inquiry phase

Category	Q& H /%	Pre-model/%	Investigate/%	Reflect/%
Task	34.2	50.5	41.2	48.3
Knowledge	35.6	9.0	42.0	32.2
Strategy	23.2	33.5	11.9	11.9
Assessment	7.0	7.0	4.9	13.1

Among these discourse, task-oriented discourse was rated highest (42%). This indicates that students were primarily concerned with establishing work procedures to complete different learning related tasks. Furthermore, the proportions of task-oriented discourse were found to be the highest in Pre-model (50.5%), followed by Reflect (48.3%), Investigate (41.2%) and Q&H (34.2%). The discourses were found to focus on managing the division of labour and the procedures to complete the tasks. Students engaged with task-oriented discourses were generally observed to perform better in time management, with most of them posing artefacts of higher quality. Therefore, students' frequent engagement in task-oriented discourse is associated with resultant higher work quality and efficiency.

The proportions of knowledge-oriented discourse were found to be the highest in Investigate (42%), followed by Q&H (35.6%) and Reflect (32.2%). These discourses were mostly associated with knowledge sharing and the construction of the new knowledge between team members (van Aalst, 2009). It was found that students who had limited prior knowledge of diffusion and osmosis tended to discuss and share their existing knowledge with their team members at the initial stage of inquiry. For example, in Q&H phase, students knew the term "Diffusion" but lacked adequate knowledge on its mechanism. Thus, questions were raised, such as "what kind of particles diffused into the air", and "what is the trigger for particles in diffusion". Their knowledge-oriented discourse also concentrated on negotiating the answer to Q_2 . Most students developed their naïve understanding of osmosis based on their prior knowledge. The terms like "saltwater", "organs", "dehydrate", "absorb" appeared in their chatting. In Investigate phase, it was noticed that while most of the students used terminologies to respond to guiding questions, they were generally observed to work together to synthesize the understanding of new abstract concepts. They tend to discuss on the best definition and mechanism of diffusion and osmosis, identification of solution concentrations, movement of particles and the energy associated with the molecules' movements. Hence, the frequent discussions on interpreting new knowledge and extending the prior knowledge have improved their understanding of diffusion, with greater gains in understanding of osmosis. In Reflect phase, the discourses indicated that students started relating their new knowledge with reference to the previous work. Through reflection, they improved their new understanding and they could apply new knowledge in revising the previous work into better quality.

The strategy-oriented discourse took place mostly at the stage of Q&H (23.2%), Pre-model (33.5%) and Investigate (11.9%). The discourse provided partners with resources or the methods to complete the activities. In Q&H, most of the strategy-oriented discourse emphasized the ways to search for relevant information. However, some students were observed to refer either to the textbooks, Internet or the teacher when the team member(s) did not reach a consensus on their ideas. In Pre-model, students discussed with their team member(s) on choosing the appropriate drawing tools, and how to conduct individual drawings and group drawings. Generally, these discourses directed them to find the best way to obtain the answers and to confirm the quality of modelling work. Particularly in Investigate, it was observed that when some of the students did not gain knowledge about the new concepts, the strategy-oriented discourse would focus on obtaining the answers through available resources. This strategy however, resulted in a fraction of students turn to the Internet to get the answers directly rather do any further reasoning and deduction on their own.

Students' assessment-oriented discourses were distributed with the lowest rate at each stage (7.0 % at Q&H, 7.0 % at Pre-model, 4.9% at Investigate and 13.1 % at Reflect). The comparative higher rate appeared in Reflect, as students were allowed to review each other's reflections and the students generally had more confidence of their conceptual understanding. We can conclude that the rate of assessment comments would increase with the improvement of their conceptual understanding. At the beginning of the lesson, students might not have a good understanding of the concepts and concerned about potentially providing incorrect assessments on their partners' work. When they obtaining know knowledge in Investigate, they were observed to be spending most of time to check or revise their prior knowledge. Therefore, students' assessment-oriented discourse received the lowest coding at the stage of Investigate. Another reason we can infer is that the guiding questions came in the form of multiple-choice. Hence, it is neither appropriate nor possible for students to comment on their partners' ideas and reasons of their choice without reviewing the explanations.

Teachers and Students Attitudes towards CSI lessons

The interview transcripts indicates that both teachers and students conveyed overall positive attitudes toward the CSI implementation in the science class. After conducting CSI lessons, the teachers recognize that the lessons have its unique educational value on fostering conceptual understanding and developing collaborative learning skills in pupils: 1) CSI can be an active and innovative pedagogy for science instruction; 2) the lessons can improve students' involvement in the activities, and lead them towards new conceptual understanding through various forms of manipulating data and information, collaborative work, peer review and discussion; 3) students are given more opportunities to exhibit their thoughts about topic (e.g. do peer discussion with a partner instead of in front of an entire class). Overall, the teachers are interested in continuing to use the system for other scientific topics to probe more about potential value of the system in science education. The students comment that they are more motivated when they were participating in the activities that were facilitated by real-time chatting, modelling and visualization. In the lessons, their learning can be guided by both instruction and questions within the system. Meanwhile, peer discussion, review and social presence also help them to keep at pace with their team member(s), which improve their time management skills and collaborative learning skills. Furthermore, they benefit from the activities in Pre-model and Investigate. The Pre-model has been identified as the most interesting and engaging stage for students, as it provides an avenue for them to explore and discuss in great depth what they observe and construct. The simulations in Investigate provide important information for them to learn new ideas such that when they were answering the guided questions, they can manipulate the simulations to review the results and then check whether the answer agrees with the simulation result.

Conclusions and Further Work

In summary, the CSI system has been demonstrated as a valuable application for enhancing students' conceptual understanding if the lessons are well designed and implemented. During the collaborative inquiry process with the system, students were provided with more opportunities to deal with the complex problems. Students' conceptual understanding are mutually improved and elaborated which are the outcomes associated with the use of multiple CSCL design elements. Besides the great improvement on the post-test achievement, students' conceptual changes were revealed in the inquiry process progressively. It can be depicted as: eliciting and applying prior knowledge in Contextualize and Q&H (exposing misconceptions) → transferring knowledge from the view of macroscopic level to particulate level in Pre-model (exposing misconceptions and establishing the microscopic view of representing the scientific phenomena) → obtaining new knowledge at particulate level through Investigate (acquiring normative ideas of the scientific phenomena) → revising and improving prior knowledge by reflection (Revising prior knowledge) → reinforcing new knowledge through their applications in the new problematized context (elaborating new understanding). The presented method of learning also enables teachers to trace students' progress in inquiry activities, their status of conceptual understanding, as well as to identify their learning difficulties in particular phases (e.g. group modelling, responses to Q₂ in Apply) as we discussed in the above section. In lessons, students have more opportunities to participate in various forms of collaboration as they become more engaged in CSI activities. Students were particularly more active with respect to collaborating with their team member(s) in Pre-model, Investigate and Reflect phases. In conclusion, the multiple CSCL design elements that are integrated at the inquiry phases in a flexible way supplement the different demands in the collaborative inquiry.

Our future research will focus on the instruction of biochemistry topic "photosynthesis". More emphasis will be paid on teachers' influence in the lessons through comparisons on the system implementation and instruction among different teachers. The studies intend to investigate the relations between students' performance and teachers' teaching methods, strategies, as well as their teaching belief on CSCL learning environment.

References

- Bell, T., Urhahne, D., Schanze, S., & Ploetzner, R. (2010). Collaborative inquiry learning: models, tools, and challenges. *International Journal of Science Education*, 32(3), 349–377.
- Bouyias, Y., & Demetriadis, S. (2012). Peer-monitoring vs. micro-script fading for enhancing knowledge acquisition when learning in computer-supported argumentation environments. *Computers & Education*, 59(2), 236-249.
- Buckley, B.C., Gobert, J.D., & Kinfield, A.C., Willett, J. (2004). Model-based teaching and learning with BioLogica™: What do they learn? How do they learn? How do we know? *Journal of Science Education and Technology*, 13(1), 23-41.
- Ergazaki, M., Komis, V., & Zogza, V. (2005). High-school students' reasoning while constructing plant growth models in a computer-supported educational environment. *International Journal of Science Education*, 27(8), 909-933.
- Gogoulou, A., Gouli, E., & Grigoriadou, M. (2008). Adapting and personalizing the communication in a synchronous communication tool. *Journal of Computer Assisted Learning*, 24(3), 203-216.
- Grosslight, L., Unger, C., Jay, E., & Smith, C., L. (1991). Understanding models and their use in science conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, 28(9), 799-822.
- Halloun, I. (1997). Schematic concepts for schematic models of the real world: The Newtonian concept of Force. *Science Education*, 82(2), 239-263.
- Jackson, J., Dukerich, L., & Hestenes. (2008). Modeling instruction: An effective model for science education. *Science Educator*, 17(1), 10-17.
- Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning?. *Computers & Education*, 49(4), 1037-1065.
- Kember, D., & Leung, Y.P., Jones, A., . . . Yeung, E. (2000). Development of a questionnaire to measure the level of reflective thinking. *Assessment & Evaluation in Higher Education*, 25(4), 381-395.
- Kilinc, A. (2007). The opinions of Turkish high school pupils on inquiry based laboratory Activities. *The Turkish Online Journal of Educational Technology*, 6(4).
- Lerner, N. (2007). Drawing to learn science: Legacies of Agassiz. *Journal of Technical Writing and Communication*, 37(4), 379-394.
- Linn, M. C. & Eylon, B.-S. (2011). *Science learning and instruction: Taking Advantage of Technology to Promote Knowledge Integration*. New York: Routledge.
- Liu, O. L., Lee, H.-S., Hofstetter, C., & Linn, M. C. (2008). Assessing knowledge integration in science: Construct, measures, and evidence. *Educational Assessment*, 13(1), 33-55.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Teaching*, 32(1), 45-61.
- Pozzi, F. (2009). Using collaborative techniques in virtual learning communities. *Lecture Notes in Computer Science*, 5794, 670-675.
- Sun, D., & Looi, Chee-Kit. (2013). Designing a web-based science learning environment for model-based collaborative inquiry. *Journal of Science Education and Technology*, 22(1), 73-89.
- van Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge construction, and knowledge-creation discourses. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 259-287.
- White, B.Y., Frederiksen, J., Frederiksen, T., Eslinger, E., Loper, S., & Collins, A. (2002). Inquiry Island: affordances of a multi-agent environment for scientific inquiry and reflective learning. In: Proceedings of the 5th international conference of the learning sciences (ICLS), Erlbaum, Mahwah, NJ.
- Yang, J. C., & Lin, Y. L. (2010). Development and evaluation of an interactive mobile learning environment with shared display groupware. *Educational Technology & Society*, 13(1), 195-207.

Acknowledgments

This research is funded by National research Foundation in Singapore (Project #: NRF2009-IDM001-MOE-019, IDM SST Future School-Science project). We would like to thank Baohui Zhang for his contributions in conceptualizing this project, and the CSI team members and our collaborators: Weikai Fu, Chaohai Chen, Jean Phua Yin Chiun, Shan Gao, Karel Mous, Kwai Yin Loh, Charles Low Soo Peng, Kassandra Lim Lay Han, Justin Ke Kaijie and their students for working with us.

Experts Learn More (than Newcomers): An Exploratory Study of Argumentation in an Online Help Forum

Hon Jie Teo, Aditya Johri, Virginia Tech, USA
Email: hjteo@vt.edu, ajohri@vt.edu

Abstract: Online help forums have emerged as a powerful form of knowledge sharing with significant implications for CSCL. To understand knowledge construction and learning within online help forums, this study examined the interactional dynamics amongst help-givers (experts) and help-seekers (newcomers) in an open, voluntary help forum for Java computer programming language. A purposive sample of active threads was analyzed using a coding scheme adapted from the argumentative knowledge construction (AKC) framework. The study found that while help-givers actively engaged in advanced argumentative and social moves, help-seekers seldom did. Furthermore, the number of advanced epistemic messages was lower than expected and both help-givers and help-seekers often veered off-topic towards socialization and conflict-oriented interactions. The study questions the feasibility of online help forums for facilitating acquisition of domain-specific knowledge by help-seekers but demonstrates the viability of the platform as a mechanism for experts to expand their knowledge. Our findings suggest that those who know more are able to learn more given their ability to engage in argumentation.

Introduction

With the increase in access and use of digital media among youth (Johri, Teo, Lo, Dufour, & Schram, 2013), research on online learning has gained increased significance. Within CSCL, contexts for research have included online virtual math chat groups (Stahl, 2006, 2011), online mathematics help forums (van de Sande, 2010; van der Sande & Leinhardt, 2007), forum for mathematics teachers (Renninger & Shumar, 2002), online distance learning classes (Johri, 2005), free non-course-related mathematics forum (Chen, Chiu & Wang, 2010), and K12 programming communities such as Scratch (Resnick et al., 2009) and MOOSE Perl (Bruckmann, 2006). Researchers are interested in online contexts because of their ability to augment formal learning by supporting knowledge construction (Scardamalia & Bereiter, 2006; Law, Yuen, Wong & Leng, 2011), sense-making (Kirschner, Buckingham Shum & Carr, 2003) and facilitating complex problem solving (Munneke, Andriessen, Kanselaar & Kirschner, 2007). Furthermore, the online discussion can be enhanced through the use of mechanisms such as 'scripts' which allow learners to engage in discourse of high epistemic quality (Stegman, Wecker, Weinberger & Fischer, 2007). The viability of asynchronous online discussions in fostering knowledge construction and acquisition has been well-documented. Analytical frameworks (Suthers et al., 2010; Weinberger & Fischer, 2006) and content analysis instruments (Clark & Sampson, 2007; Gunawardena et al., 1997; Pena-Shaff & Nicholls, 2004) have been used to understand learning in online settings. In this study, we adopt the argumentative knowledge construction (AKC) framework proposed by Weinberger and Fischer (2006) to obtain an in-depth picture of the interactions between help-givers and help-seekers in an unscripted online forum (1). Our interest in this study and the framework emerged from two primary research questions about online help forums – what does learning look like and who learns by participating in these communities? As we discuss in detail below, our use of the Weinberger and Fischer (2006) framework is that it lends itself well to our goal – an examination of the quality of arguments and the degree of social interaction in an unscripted environment where contributions are voluntary and expertise levels may vary amongst learners.

Argumentative Knowledge Construction

Knowledge construction is a useful framework for analyzing collaborative online activity as it emphasizes the role of social activity in negotiating meanings relevant to the learning task (Dillenbourg, 1999; Stahl, Koschmann & Suthers, 2006). Through dialogue, participants are able to create a coordinated activity where they can maintain a shared conception of a problem (Dillenbourg, 1999; Scardamalia & Bereiter, 2006; Pena-shaff & Nicholls, 2004; Hmelo-Silver, 2003). Within the knowledge construction framework, researchers interested in the design of learning environments have examined learner interaction and discourse to identify mechanisms that make knowledge construction viable and have found that argumentation activities support help students learn about argumentative structures and constructing explanations (Andriessen, 2007; Berland & Resier, 2008; Toulmin, 1958). Argumentative knowledge construction (AKC) has emerged as a useful construct for framing and understanding learning tasks. It emphasizes problem discussion by making argumentative elaborations with the goal to contribute towards multiple perspectives, produce joint-solutions, and acquire domain-specific knowledge (Astleitner, Brunken & Leutner, 2003; Jamaludin, Chee & Ho, 2009; Stegmann, Weinberger & Fischer 2006). To achieve AKC, learners collaborate and exchange their knowledge through discourse on an

open and complex learning task through the construction of arguments, balance of arguments and supporting their arguments with evidences, and counter arguments (Weinberger, 2003; Weinberger & Fischer 2006). The examination of argumentative discourse produces further insights into how learners engage in negotiation of different perspectives through theoretical conceptualization and application (Weinberger & Fischer, 2006). For instance, Stegmann and colleagues (2007) studied a problem-based learning in a higher education setting and found that argumentative collaboration scripts can foster the quality of argumentations, such that there are more arguments contain verifiable claims and qualifying statements, and help learners acquire knowledge about argumentation. However, they reported that there are no learning gains when it comes to the acquisition of content-specific knowledge. Overall, the ability of the knowledge construction framework to illuminate joint activity where learners improve upon initial ideas through dialogue makes is pertinent for studying online help forums. In these forums interaction takes place through text based dialogue which is stored and available for analysis and this is the only form of interaction that takes place among the participants thereby increasing the usefulness of the data in terms of interpretation.

Research Setting – Java Newcomer Forum

Online help forums are as old as electronic communication itself and their continued viability and growth over the years is an indicator of their usefulness to users. In recent years, one form of online communities – help forums – have flourished in membership and have become a powerful complement to informal learning opportunities available to people and both the level of participation as well as the diversity of topics covered by help forums is astounding (Singh, 2008). In the seven websites she sampled, Singh (2008) found approximately 3.8 million registered users excluding the guests who post and the visitors who lurk on the forums as invisible visitors and never post or register. One arena in which help forums have had a persistence and successful presence is computer programming. The usefulness of online help-forums for learning to program is easy to see. It is easy to represent software code online and relatively straightforward for the help giver to ‘run’ the code and see if it works. For our study, we targeted a popular programming language with millions of user worldwide and significant online support – Java™. The popularity of Java comes from its usefulness for developing Internet related applications and the fact that it is open-sourced. Open sourcing of any software artifact – language, program or product – results in significant growth in the supporting online community (Crowston & Howison, 2003) and Java is no exception (Johri, Nov & Mitra, 2011a, 2011b). Studies have found that users in online communities that are formed around open source software actively engage in collaboration, learning and socialization (Lakhani & Wolf, 2005) and leverage tools such as code repositories, forum discussion, blogs and chat clients to facilitate interaction (Crowston, Wei, Howison & Wiggins, 2012). In these communities help-seeking and help-giving are common and Lakhani and Von Hippel (2003) found that help-givers spent 2% of the total spent time on the online community answering questions and 98% of their time reading questions as well as crafting potential answers to the questions. Singh (2008) found that the process of help-giving in these communities is many-to-many (many help-seekers and many help-givers are involved) and that the process of help-giving is seldom dyadic and often iterative (Singh & Twidale, 2008; Singh, Kathuria & Johri, 2012).

In contrast to formal and highly structured instruction, the site we selected for this study (the Java community) has an informal structure and both participation and collaboration is voluntary and not mandated by coursework. Help is mainly provided by a core group of volunteers and the discussion task structure deviates from the common set-ups of common wrapper/starter roles and open-ended class discussions without pre-designated roles. This setting is problem driven as help-seeker starts a discussion soliciting help from voluntary help-givers to assist them with their learning needs and that challenges emerge through the engagement of newcomers with a programming task. Figure 1 shows a screenshot of the forum and Figure 2 represents the structure of a sample discussion. The forum explicitly prohibited off-task discussion – such as discussion of Java as a language or its open sourcing – in the forum. The terms of participation made this explicit.

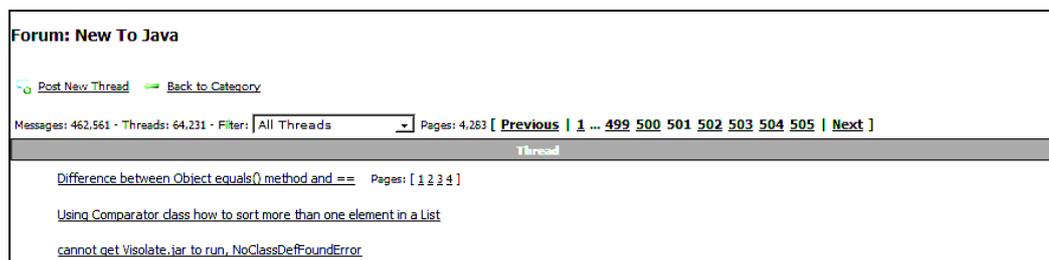


Figure 1. A page listing discussion topics in the ‘New to Java’ discussion forum

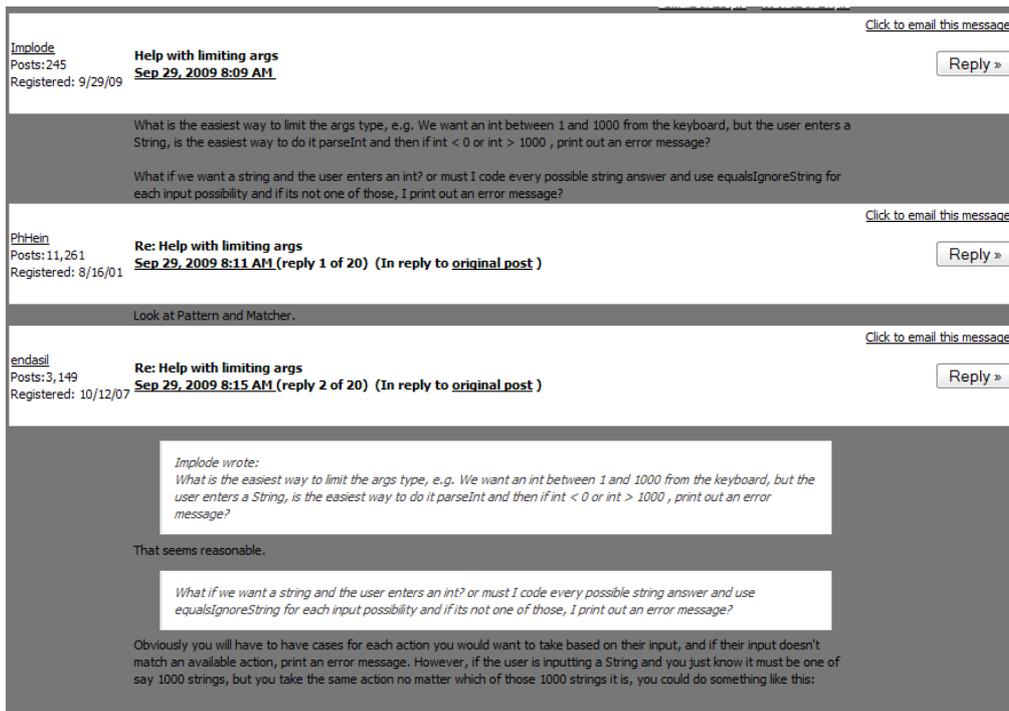


Figure 2. Example of a help request and follow-up discussion

Methodology, Data Collection, and Data Analysis

The discussion forum data that we collected consisted of a total of 37,472 discussion topics. Our data collection complied with both internal institutional IRB protocol and the ‘terms of services’ of the forum. We limited data collection to public data that that did not require creating an account and logging in to the site to be able to read the forum messages. The data ranged from a time period between 2006 and 2010. User post counts (see Table 1) indicates that most users have less than 10 posts in this discussion forum with only 326 users having post counts more than 500. Therefore, similar to other help forums, the majority of users participate infrequently in contrast to a much smaller number of expert helpers who contribute much more frequently to the forums. For the study reported in this paper, data were collected from the most active discussion topics created in the month of September in year 2009 and which accounted for 9.4% of all topics started in the month. By only examining the most active discussion topics, we are able to focus on the entirety of discussion to account for a wide range of domain-specific content, argumentative and interactional moves. In total, 1119 messages from 47 discussion threads amounting to a total of 108 pages of conversations were analyzed. Further details of the research site and data collection are available in Mitra (2011).

Table 1: Users activity in the New to Java discussion forums

Post Count	Number of Users
0 to 10	14114
11 to 50	5076
51 to 500	1922
> 500	326

Our coding scheme is based on the three independent dimensions: epistemic, argumentative and social modes (See Table 2) as described by Weinberger & Fischer (2006). The epistemic dimension allows us to investigate if learners are on-task and if they are able to convey theoretical concepts to explain concepts or apply concepts to provide a solution to the case at hand. The argumentative dimension is focused on the argumentation sequences and whether learners make arguments, counter arguments, and interactive arguments in addition to drawing from prior knowledge. The social modes dimension is concerned with the extent to which the learners externalize their thoughts, question others, accept the contributions of others and disagree or agree with the perspectives of others. While the analytical framework (Weinberger & Fischer, 2006) is useful for guiding analysis at both the micro and macro levels, a macro-approach was chosen such that the complete message was chosen as the unit of analysis for the coding. This is in line with the suggestion by Rourke et al. (2001), who

argued that taking the complete message as the unit of analysis is the most objective approach as data analysis is performed under the entirety of interaction. The inter-rater agreement for this study is 0.78 (Cohen, 1960).

Table 2: Coding scheme with categories in epistemic, argumentative and social dimensions (adapted from Weinberger & Fischer, 2006)

Code	Category	Description
EPI1	Non-epistemic activities	Learners discussing off-topic or digressing off-topic
EPI2	Construction of problem space	Learners convey contextual information within the argumentation space with the aim of fostering understanding of the topic
EPI3	Construction of conceptual space	Learners convey theoretical concepts and explain theoretical principles to foster theoretical understanding of the topic
EPI4	Construction of adequate relations between conceptual and problem space	Learners convey theoretical concepts of case information and apply relevant theoretical concepts adequately to provide a solution for the topic problem
EPI5	Construction of relations between prior knowledge and problem space	Applying concepts that stem from prior knowledge rather than the new theoretical concepts that are to be learned
ARG1	Non-argumentative moves	Questions, coordinating moves, and meta-statements on argumentation
ARG2	Argument	Statement put forward in favor of a specific proposition
ARG3	Counterargument	An argument opposing a preceding argument, favoring an opposite proposition
ARG4	Integration	Statement that aims to balance and to advance a preceding argument and counterargument
SOC1	Externalization	Externalizing or articulating thoughts to the group
SOC2	Elicitation	Eliciting a response, questioning the learning partner, or provoking a reaction from group
SOC3	Quick consensus building	Accepting the contributions of group members in order to move on with the task
SOC4	Integration-oriented consensus building	Taking over, integrating and applying the perspectives of group members
SOC5	Conflict-oriented consensus building	Disagreeing, modifying or replacing the perspectives of group members

Findings

Our analysis (see left most bar in Fig. 2) indicates that non-epistemic (EPI1) messages comprised 42.2% of all messages corresponding to the epistemic dimension followed by epistemic activities of higher order such as conveying of contextual information (EPI2) at 32.6% and conveying of theoretical concepts (EPI4) at 17.6%. We, however, also found that both help-givers and help-seekers seldom explicitly drew from any prior knowledge (EPI5) at 1.6%. Chi-square analysis showed that help-givers are more likely to engage in non-epistemic messages ($\chi^2(18, N = 94) = 43.5, p < 0.001$). Together with the previous finding, that there are a large number of non-epistemic messages, this significant difference suggests that as help-seekers engage in help-oriented discussions they do not at the same time actively participate in any off-task activities, such as engaging in socialization moves. On the other hand, experts are more likely to engage in higher level epistemic activities from explanation of theoretical concepts ($\chi^2(5, N = 94) = 38.5, p < 0.001$) to application of theoretical concepts ($\chi^2(10, N = 94) = 60.6, p < 0.001$). This finding reflects the nature of help-oriented discussion where help-givers have more expertise and more likely to share their theoretical and conceptual knowledge to provide a solution to the case at hand or to explain theoretical concepts.

In the argumentative dimension, we found that non-argumentative moves (ARG1) constitute 90.6% of all argumentative messages and made up the majority of messages in the argumentative dimension (see middle bar in Fig. 2). This is not surprising since topics started by help-seekers may not be entirely conducive for AKC. Amongst the argumentative moves, the occurrences of ARG2, ARG3 and ARG4 make up 37.5, 41.3 and 21.2% of the total moves. Our analysis also highlight that help-seekers seldom engaged in any of these argumentative moves at 99%. Chi-square analysis showed that that help-givers are more likely to engage in argumentative moves ARG3 and ARG4 which are putting forth arguments ($\chi^2(4, N = 94) = 16.1, p = 0.0013$) and counterarguments ($\chi^2(5, N = 94) = 13.8, p = 0.0058$) and integrative arguments ($\chi^2(3, N = 94) = 16.5, p < 0.001$). This discussion platform, being a help-oriented online discussion forum, is made up of help-givers dominantly engaged in putting forth their arguments when the topic at hand requires it. This seems to suggest that help-seekers are not actively participating in the argumentative processes as they may lack the knowledge needed to engage in any arguments. It is therefore suspect that the multiple perspectives that originate from help-givers' arguments have any positive impact on help-seekers' learning experiences, especially when experts

touch on advanced domain-specific content without explicit explanation of the content. The limitation of this finding is that we are not able to detect any form of lurking behavior and therefore cannot account for help-seekers who are observing yet do not make a contribution to the discussion.

In the social dimension, the use of questions or provocation to elicit responses (SOC2) constitutes 32.0% of all social messages whereas externalization of thoughts occurred at a frequency of 22.7%. On the other hand, quick consensus building and conflict-oriented consensus building messages (SOC3 and SOC4) constitutes 14.0% and 26.4% of all social messages respectively (see right most bar in Fig. 2). The relatively higher number of SOC2 moves can be explained by the nature of the help-oriented discussions where information or prompts, rather than solutions, was provided to enable help-seekers to meeting their learning needs (Lakhani & von Hippel, 2003). It is also observed that both help-givers and helps-seekers seldom participate in moves that show integration of others' perspectives into their viewpoints (SOC5) and these moves only feature in 4.9% of all social messages. Through chi-square analysis, our analysis suggests while help-seekers are more likely to use questions and provocation to solicit responses, help-givers are more likely to take over the perspectives of others ($\chi^2 (4, N = 94) = 9.5, p < 0.028$) and are more likely to engage in conflict-oriented consensus building social moves ($\chi^2 (13, N = 94) = 44.7, p < 0.001$). There are no significant differences in our examination of the frequencies of SOC2 ($\chi^2 (7, N = 94) = 8.6, p < 0.001$) and this suggests that both help-seekers and help-givers are also engaged in questioning and provocation, which is a positive aspect of the online help discussion. To complement the discussion of our findings, we also present Table 3 which shows an excerpt for a help-discussion featuring acts from the three dimensions.

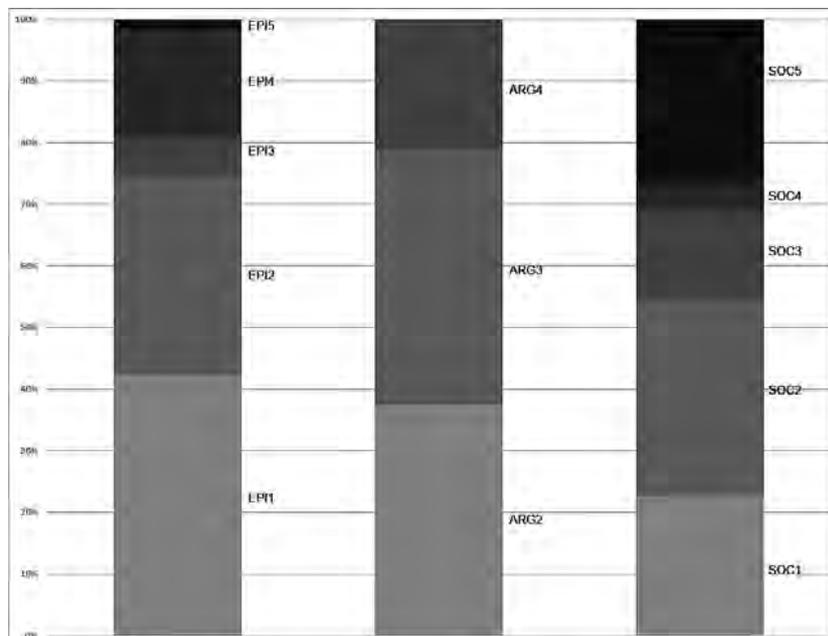


Figure 3. Frequencies of Occurrence of Epistemic, Argumentative and Social Acts

Table 3: Excerpt from a help discussion on the computational concept of “inheritance”

Actor	Message Content
Help-seeker (HS)	<p>Hello again, Guys I've read this tutorial : http://java.sun.com/docs/books/tutorial/java/land/subclasses.html As one of you suggested in one of your replies to my question on topic about inheritance and that's fine. Unfortunately in this tutorial two sentences contradict each other:</p> <ol style="list-style-type: none"> 1. A subclass inherits all the members (fields, methods, and nested classes) from its superclass. Constructors are not members, so they are not inherited by subclasses, but the constructor of the superclass can be invoked from the subclass. 2. A subclass does not inherit the private members of its parent class. <p>Moreover in netbeans while debugging it shows that private members of superclass are inherited by child class. So what's going on? Which sentence is true? And really, but on sun's official website this kind of error shouldn't have place. Looking forward to your reply</p>
Help-giver (HG) #1	<p>It's really semantics. The only difference it makes if they're inherited or not is what the definition of inheritance must be. However, according to the Java Language Specification, I would say that private members are not inherited.</p> <p><i>Quote: 6.4.3 The Members of a Class Type says: Members are either declared in the type, or inherited because they are accessible members of a superclass or superinterface which are neither private nor hidden nor overridden.</i></p> <p>Although it's entirely possible that the JLS contradicts itself elsewhere in that regard. As I said though, it makes no difference.</p>

HG #2	If it's not clear, please provide feedback, so Sun can improve the tutorial (the link from the bottom of the tutorial page): http://developers.sun.com/contact/tutorial_feedback.jsp
HS	Well it's not really semantics. Its fundamental very important concept. Well you saying that it makes no difference but I cannot agree. It is fundamental to know and understand how the mechanism of inheritance work and in order to understand that one has to know about basic concepts like inheritance and which members are inherited and I really wish that sun would make it clear and do not make any more such a basic mistakes. Is there in java (I'm sure there is) a method which will allow me to check size of an object of any class? Doing so I would answer to my question myself but logically thinking every member is inherited otherwise on what member would we operate in derived class through public methods? Regards [user id removed]
HG #2	Maybe they could have been clearer, but then how much did you pay for that tutorial (or for any Java-technology from Sun, for that matter) to Sun? Do as I suggested and provide feedback to Sun. I don't recommend to access fields of superclasses directly. If you need to access them, use either public or protected accessors.
HG #3	<p><i>Quote HS: Well it's not really semantics. Its fundamental, very important concept.</i> Same thing. The semantics of the fundamental concept.</p> <p><i>Quote HS: I really wish that sun would make it clear</i> They have. See the Java Language Specification.</p> <p><i>Quote HS: And do not make any more such a basic mistakes.</i> Such as??</p> <p><i>Quote HS: Is there in java (I'm sure there is) a method which will allow me to check size of an object of any class?</i> No.</p> <p><i>Quote HS: Doing so I would answer to my question myself but logically thinking every member is inherited otherwise on what member would we operate in derived class through public methods?</i> Every member is inherited, but they are not all accessible to derived classes, depending on the access modifiers concerned. It's in the JLS. Have a look. And stop guessing.</p>

Discussion and Conclusion

Online communities represent a significant avenue to understand many learning and education related issues such as newcomer socialization (Ducheneaut, 2005), free user-to-user help (Lakhani & von Hippel, 2003), and distribution of tasks (von Krogh, Spaeth & Lakhani, 2003; Mockus et al., 2002). Through examination of the epistemic, argumentative and social tendencies of help-givers and help-seekers in an open unscripted learning environment, we gained deeper insights into how help discussions can be characterized and how learning is taking place in unscripted online discussion. The high number of off-topic messages raises a concern whether the forums are a productive avenue for significant acquisition of domain-specific knowledge and echoes the concerns raised by Stegmann and colleagues (2007). Both help-givers and help-seekers were likely to veer off-topic towards socialization and conflict-oriented social interactions. In our analysis, we often find that the quality of help is highly dependent on the social interactions between the help-seeker and the help-givers. Taking into account that help-givers are volunteering their time, effort and offering their expertise, it is not surprising that specific social and community practices have been established to allow them to maximize their voluntary efforts. As such, help-seekers who are accustomed to these practices are more likely to attain highly productive learning experiences. In our data analysis, we observed that help-seekers, as newcomers to the community, were often not cognizant of the community practices such as using code-tags to organize their program codes and provide sufficient contextual information desired by the help-givers.

In sum, we found that while voluntary and open online discussion in this educational setting are rich in discourse and can span a long period of time, they seldom feature advanced levels of knowledge construction that result from the integration of multiple perspectives and argumentations. This runs contrary to other researchers who found high levels of arguments in discussion forums tied to formal instructional settings (Pena-Shaff & Nicholls, 2004). Advanced discussions are most likely not to occur because the multiple perspectives that the experienced help-givers provide through their arguments are most likely not able to be absorbed or understood by help-seekers. This can be attributed to the existence of a core group of help-seekers (see Table 2) who may possess a high level of domain-specific knowledge but may not capably provide learning support to maximize help-seekers' learning experiences. In the statistical examination of differences between the two groups of learners, we found that the help-seekers are more likely to engage in argumentative moves and engage in integration-oriented consensus building moves. These findings highlight an opportunity for community leaders to address the design of tasks and roles to address the expertise gap between the experts and newcomers as well as to foster more engaged learning experiences. Online communities have much to gain by encouraging participation from the larger pool of newcomer learners as it will in turn encourage higher levels of collective contributions and learner diversity. Finally, consistent with prior research on expertise we found that due to their ability to organize knowledge and connect concepts – metacognitive proclivity – experts were able to engage more centrally with the discussion, argue strongly, and thereby also construct knowledge and learn (NRC, 2000). Our findings hint at a unique problem with informal learning in that those who know more are able to learn more as they can argue more both due to prior knowledge, and in some instances, the culture of their profession or community.

Acknowledgement(s)

This material is based upon work supported, in part, by the U.S. National Science Foundation under Grants EEC-0954034 & EEC-0935143. 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 funding agency.

Endnotes

- (1) As other scholars have suggested, the role of ‘help-givers’ and ‘help-seekers’ in online communities is correlated with ‘experts’ and ‘newcomers’ respectively and we use the terms interchangeably.

References

- Andriessen, J. (2007). Arguing to learn. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 443 – 460). Cambridge, England: Cambridge University Press.
- Astleitner, H., Brunken, R., & Leutner, D. (2003). The quality of instructional materials for argumentative knowledge construction. *Journal of Instructional Psychology*, 30(1), 3–11
- Berland, L.K., & Reiser, B.J. (2008). Making sense of argumentation and explanation. *Science Education*, 93(1), 26-55.
- Bruckman, A. (2006). Learning In Online Communities. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences*. (pp. 461-472). Cambridge University Press, New York, NY.
- Bruner, J. (1999). Some reflections on education research. In E. C. Lagemann & L. S. Shulman (Eds.), *Issues in education research: Problems and possibilities* (pp. 399-409). San Francisco: Jossey-Bass Publishers.
- Chen, G., Chiu, M. M., & Wang, Z. (2010). Group micro-creativity in online discussions. *Proceedings of the 9th International Conference of the Learning Sciences*, volume 1 (pp. 357-364).
- Clark, D. B., & Sampson, V. (2008). Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. *Journal of Research in Science Teaching*, 45(3), 293–321.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20(1), 37–46.
- Crowston, K., & Howison, J. (2003) The social structure of open source software development teams. *Paper presented at the OASIS 2003 Workshop (IFIP 8.2 WG)*.
- Crowston, K., Wei, K., Howison, J., & Wiggins, A. (2012). Free/Libre open-source software development: What we know and what we do not know. *ACM Computing Surveys (CSUR)*, 44(2), 7.
- Dillenbourg, P. (1999). What do you mean by "collaborative learning"? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1-16). Amsterdam, NL: Pergamon
- Ducheneaut, N. (2005). Socialization in an Open Source Software community: a socio-technical analysis. *Computer Supported Cooperative Work*, 4(4), 323-368.
- 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), 395–429
- Hmelo-Silver, C. E. (2003). Analyzing collaborative knowledge construction: Multiple methods for integrated understanding. *Computers and Education*, 41, 397-420.
- Jamaludin, A., Chee, Y.S., & Ho, C.M.L. (2009). Fostering argumentative knowledge construction through enactive role play in Second Life. *Computers & Education*, 53(2), 317-329.
- Johri, A. (2005). Online, Offline, and In-Between: Analyzing Mediated-Action among American and Russian Students in an Online Class. In Roberts, T. (Ed). *Computer-Supported Collaborative Learning in Higher Education*. Idea Group Publishing. pp. 259-280.
- Johri, A., Teo, H., Lo, J., Dufour, M., & Schram, A. B. (2013). Millennial Engineers: Digital Media and Information Ecology of Engineering Students. *Computers in Human Behavior*, <http://dx.doi.org/10.1016/j.chb.2013.01.048>
- Johri, A., Nov, O. & Mitra, R. (2011a). Environmental Jolts: Impact of Exogenous Factors on Online Community Participation. *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 649-652.
- Johri, A., Nov, O., & Mitra, R. (2011a). Cool or monster?: company takeovers and their effect on open source community participation. *Proceedings of the 2011 iConference* (pp. 327-331).
- Kirschner, P. A., Buckingham Shum, S. J., & Carr, C. S. (2003). *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*. London: Springer Verlag.
- Lakhani, K. & Wolf, B. (2005). Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects. In: Feller, J., B. Fitzgerald, S.Hissam, K. Lakhani (eds.), *Perspectives on Free and Open Source Software*, Cambridge, MA, The MIT Press
- Lakhani, K. R., & von Hippel, E. (2003). How open source software works: "Free" user-to-user assistance. *Research Policy*, 32(6), 923-943.

- Law, N., Yuen, J., Wong, O. W., & Leng, J. (2011). Understanding Learners' Knowledge Building Trajectory Through Visualizations of Multiple Automated Analyses. In S. Puntambekar, G. Erkens & C. Hmelo-Silver (Eds.), *Analyzing interactions in CSCL: Methodologies, approaches and issues*. (pp. 47-82). New York, NY: Springer.
- Mitra, R. (2011). Collaborative learning in Open Source Software (OSS) communities: The dynamics and challenges in networked learning environments. M.S. Thesis, Virginia Tech.
- Mockus, A., Fielding, R., & Herbsleb, J.D. (2002). Two case studies of open source software development: Apache and Mozilla. *ACM Transactions on Software Engineering and Methodology*, 11(3), 309-346.
- Munneke, L., Andriessen, J., Kanselaar, G., & Kirschner, P. (2007). Supporting interactive argumentation: Influence of representational tools on discussing a wicked problem. *Computers in Human Behavior*, 23, 1072-1088.
- NRC (National Research Council) (2000). *How People Learn: Brain, Mind, Experience, and School*. National Academy Press.
- Pena-Shaff, J., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education*, 42(3), 243-265.
- Renninger, K. A., & Shumar, W. (2002). Community building with and for teachers at the math forum. In K. A. Renninger & W. Shumar (Eds.), *Building virtual communities* (pp.60-95). New York: Cambridge University Press.
- Resnick, M., Maloney, J., Monroy-Hernandez, A., Rusk, N., Eastmond, E., et al. (2009). Scratch: programming for all. *Communications of the ACM*, 52(11): 60-7.
- Rourke, L., Anderson, T., Garrison, D.R. and Archer, W., (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education*, 12.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 97-115). New York, NY: Cambridge University Press.
- Singh, V. (2008). *Knowledge Creation, Sharing and Reuse in Online Technical Support for Open Source Software*. Unpublished Doctoral Dissertation Submitted to University of Illinois, Urbana-Champaign.
- Singh, V., Kathuria, S. & Johri, A. (2012). Newcomer integration and learning in technical support communities for Open Source Software. *Proceedings of 2012 ACM Conference on CSCW*.
- Singh, V. and Twidale, M.B. (2008). The Confusion of Crowds: non-dyadic help interactions. *Proceedings of Computer Supported Cooperative Work 2008*, (San Diego, CA, 2008), ACM.
- Software. Unpublished Doctoral Dissertation, University of Illinois at Urbana-Champaign, 2008.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. MIT Press.
- Stahl, G. (2011). How to study group cognition. In S. Puntambekar, G. Erkens & C. Hmelo-Silver (Eds.), *Analyzing interactions in CSCL: Methodologies, approaches and issues*. (pp. 107-130). NY: Springer.
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 409-426). Cambridge, UK: Cambridge University Press.
- Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, 2(4), 421-447.
- Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2007). Collaborative argumentation and cognitive processing: an empirical study in a computer-supported collaborative learning environment. *Paper presented at the Proceedings of the 8th international conference on Computer supported collaborative learning*, New Brunswick, New Jersey, USA.
- Suthers, D.D., Dwyer, N., Medina, R., & Vatrappu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. *International Journal of Computer-Supported Collaborative Learning*, 5(1), 5-42.
- Toulmin, S. (1958). *The uses of argument*. Cambridge, MA: Cambridge University Press
- van de Sande, C. & Leinhardt, G. (2007). Help! Active student learning and error remediation in an online calculus e-help community. *Electronic Journal of e-Learning*. 5(3), 227-238,
- van de Sande, C.. (2010). Free, open, online, mathematics help forums: the good, the bad, and the ugly. *Proceedings of the 9th International Conference of the Learning Sciences*, pp. 643-650.
- von Krogh, G., Spaeth, S. & Lakhani, K. (2003). Community, joining, and specialization in open source software innovation: a case study. *Research Policy*, 32(7), 1217-1241.
- Weinberger, A. (2003). *Scripts for computer-supported collaborative learning: Effects of social and epistemic cooperation scripts on collaborative knowledge construction*. Unpublished doctoral dissertation. Germany: University of Munich.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, 46(1), 71-95.

Identification of patterns of tool use and sketching practices in a *learning by design* task

Kate Thompson, David Ashe, Dewa Wardak, Pippa Yeoman, Martin Parisio, CoCo Research Centre, The University of Sydney, Building A35, The University of Sydney, NSW 2006
Email: kate.thompson@sydney.edu.au, david.ashe@sydney.edu.au, dwar9402@uni.sydney.edu.au, philippa.yeoman@sydney.edu.au, martin.parisio@sydney.edu.au

Abstract: The complex interaction of tool use (both physical and digital) in face-to-face collaborative learning situations, and the role that these tools play in facilitating group work is increasingly important as tools for learning become more sophisticated and specialized. In this paper, a group of five high school students is studied as they engage in a *learning by design* task to design an educational resource about a local waterway. They carried out this design work in The Design Studio at the University of Sydney, using an iPad projected onto a whiteboard wall. Multiple streams of data were collected, visualized and analyzed, which allowed the overall patterns of tool use for all members of the group to be identified in relation to the development of their design. Two patterns of tool use are identified and analyzed according to the practice of sketching identified in other fields of design.

Introduction

As learners engage in more complex tasks, the ability to capture the processes of learning is increasingly challenging. This paper presents initial research into the combination of tools used by high school students during their design work in a dedicated design space. As part of a project funded by a regional government organization, a group of school students engaged in the design of a learning resource about a waterway of local significance, to be shared with other schools throughout the region. Students were observed and guided through a design process, intended to help them learn about issues relating to water quality in their area. Students did their design work in a physical space resourced with tools for educational designers. Many aspects of the processes of design are of importance to learning, particularly in *learning by design* tasks. As yet, these have not been discussed in the research literature. In this research, work on combinations of tool use in the process of learning in online environments, is aligned with knowledge of the importance of sketching in other areas of design, in order to identify patterns of productive activity in *learning by design* tasks.

The paper begins with a brief overview of research involving integrated tool use and the processes of collaboration (Thompson & Kelly, 2012a, 2012b; Zenios & Holmes, 2010; Thompson, Ashe, Carvalho, Goodyear, Kelly & Parisio, accepted), as well as background to *learning by design* (LBD), and the role of sketching during the design process. The physical and digital space in which the study was conducted will be described, as well as the participants and context of the LBD task. Multiple streams of data are then visualized and analyzed to describe the integrated, and specialized way in which students used and appropriated the tools available to them, for their purposes.

Literature Review

Much of the work examining the processes of learning during complex collaboration with regard to combinations of tool use, has been performed in online spaces (see for example Zenios & Holmes, 2010; Thompson & Kelly, 2012a; 2012b). These studies have examined the affordances of using a combination of online tools for discussion and idea development (such as chat or video conferencing tools), for the permanent recording of the progress of the group, and for in-depth investigation of a problem (such as wikis). In these studies, the permanence of the recording space (the wiki) was important for the progression of the collaboration; more complex discussions were able to take place once ideas were recorded. The recordings enabled participants to refer back to previous ideas. In work on face-to-face collaboration, Thompson et al. (accepted) have shown that differentiated tool use can be an indicator of successful collaboration, in a group, as students develop expertise in the role and tasks related to the tool used.

Analyzing the processes of learning in CSCL is important, in particular as the context of learning environments are able to be better described (Thompson et al., accepted). Work in this field has discussed this importance in terms of social interactions, analyzing the discourse, as well as non-verbal interaction (for a review of this, see Goodyear, Jones & Thompson, forthcoming). By extracting additional data from complex datasets, analysis of additional processes of learning can be added to our understanding of collaborative learning (Thompson et al., accepted; Kennedy-Clark & Thompson, accepted). This paper addresses the processes involved in *tool use* in a dedicated design space.

Design activity has been demonstrated to facilitate deep learning about complex systems (Hmelo, Holton, & Kolodner, 2000; Kolodner et al., 2003). The goal-orientation of design-based science learning, where

the content is fundamental to producing a working artifact, acts to contextualize the process of inquiry. “Design is used as a vehicle through which scientific knowledge is constructed and real-world problem-solving skills are cultivated. The design challenge provides impetus for identifying what needs to be learned and for sustaining engagement in inquiry over time, as well as providing need for cultivating and using a variety of skills.” (Vattam & Kolodner, 2008, p. 407). Vattam and Kolodner (2008) describe LBD as internally driven by a “need to know” and a “need to do”. During this iterative process students share experiences and ideas as they articulate what they will need to learn in order to successfully complete the design challenge. In describing this process they highlight the importance of the *design rule of thumb*, which assists learners to connect content and application.

The early conceptual phase of design is cognitively very demanding. During this phase, ideas are generated rapidly. Designers need to try out new ideas quickly and cheaply before committing to one for further development (McGown, Green, & Rodgers, 1998). During this process, the intentions of the designers are represented, or externalized, in the form of verbal explanations, written documents, sketches, drawings, diagrams, and sometimes as models. Free-hand sketches are an indispensable tool at this stage not only for externalizing design ideas in order to communicate with others, but also for examining the potential of a solution before further commitment. For designers, drawing is the act of exploration (Hokanson, 2008), of discovery (Berger, 2007), and an aid to their thought process (Buxton, 2007). Designers draw in order to spot problems, see new features and relations among elements, discover or promote new ideas, and refine current ideas (Suwa & Tversky, 2002).

The importance of sketching in design is not the sketch as a final product of design, but as a process that explores, questions, suggests, proposes, and provokes thoughts and ideas (Buxton, 2007). In other words, “sketching is thinking” (Goldschmidt, 1991, p. 130). This process enables the designer to find new aspects of the problem and to generate new ideas (Suwa, Gero, & Purcell, 2000). Designers from various fields rely on the use of visual representations as thinking tools in their design process. For example, Henderson (1999) quotes an engineer who stated “I can’t think without my drawing board”. Sketches can greatly improve communication in collaborative team situations (Eppler & Burkhard, 2006); they can be used to assist a group’s reflection and communication process. This is because sketches and drawings bring the main features of the object of design to the conversation by making them explicit and thus debatable.

Methods

A group of 16 students and 11 adults participated in the project funded by a regional organization of councils, and supported by their high-school and a local environmental rehabilitation organization. The aim of the project was to develop an innovative fieldwork and multimedia framework to engage other students in water and land management issues. South Creek, the focus of the project has been described as the most degraded creek in the region. Threats to its health include vegetation clearance, urbanization and agriculture; resulting in concerns about increased nutrient levels and an increase in the number of weed species within the riparian zone. Over a number of months students and stakeholders participated in a multidisciplinary design process to create a learning resource for use in schools across the region. The students participated in planning sessions, a site visit, and a day of hands-on site restoration at the creek, before attending a *Design Day* at the University of Sydney.

The *Design Day* was an opportunity for the students to develop their ideas about the design of the educational resource, propose possible formats, identify constraints and generate a consensus upon which a brief for the multimedia designer could be written. Participants included expert learners (the students), who ranged in age from 12 to 17 years, and experts from education, environmental science and multimedia design. The *Design Day* began with each of the expert groups outlined their desires and constraints for the design of the educational resource, and these initial parameters were summarised and referred to during the rest of the day. In the Design Studio⁽¹⁾ participants were led through the first three stages of a design process: *empathize*, *define*, and *ideate* (following the Stanford University Institute of Design – An Introduction to Design Thinking⁽²⁾). The *empathize* stage took the form of a whole group brainstorming activity during which the initial desires and constraints were discussed and the critical components extracted. Participants worked in pairs for the *define* stage; each dyad performed a needs analysis to help define the resource. During the third stage, *ideate*, the participants worked in groups according to discipline area. One group contained all the adults including educators, multimedia designers and environmental scientists. The other three groups each included between five and six students – expert learners. During the *ideate* stage, participants were asked to generate ideas. They were asked not to limit themselves to their knowledge of technology and were instructed to record all ideas. The intention was to explore a wide solution space so that, later in the *Design Day*, these ideas could be distilled into one coherent solution and a brief presented to the multimedia designer for the creation of a resource prototype.

This paper follows one of the student groups during the *ideate* stage of the design day. There were five members of the group, three females and two males, whose ages ranged from 12 to 16. Video, audio and photographs of the collaborative design work were collected. Each group was given a choice regarding the digital tool they used (computer, interactive whiteboard, or iPad). This group chose to use an iPad that was

projected onto a wall painted in whiteboard paint (a ‘white-wall’, see Figure 1). Their interactions with both the physical tools (whiteboard/pens/erasers) and digital tools (iPad and projected screen) were recorded using video. In addition, two members of the group wore audio recorders, and photographs were taken every 12 seconds of the whitewall on which students were working.

Data describing the tool use, for each member of the group, in 30 second intervals was extracted from the video. The tool, action (writing or erasing), and location (inside or outside the projected area) were coded. The transcription of their discourse was analyzed in order to identify ideas important in the development of their design. These were then discussed and agreed upon by all authors of the paper in a group meeting (Thompson, Ashe, Yeoman & Parisio, accepted). In addition, the photographs of the whitewall taken every 12 seconds, were coded using The Collaborative Process Analysis Coding Scheme (CPACS) (Kennedy-Clark & Thompson, accepted). CPACS is a multi-level coding scheme that includes macro-levels (action and content) as well as micro-levels (pronouns, tense, modality, and attitude) of collaboration; only the macro-level code, *Content*, was used in this paper. The Content section of CPACS contains six codes: phatics (social – phatics, salutations, leave taking), planning, topic, task, tool use, and off-task. The photographs were coded by one researcher, and half were coded by a second researcher. Inter-rater reliability of 55% was achieved in the first instance, and 86% after further discussion. The ideas were plotted, over time, in combination with the tools used by each person and taking into account the *content* to which they were contributing. We used this visualization to determine patterns of tool use, and analysed the photographs taken of the whitewall from the perspective of sketching practices of designers.

Analysis

The design studio was equipped with a variety of design tools and participants were able to use them as they wished. This group of students was assigned to an area of the studio that had a white-wall (the entire wall set up as a whiteboard) an assortment of coloured marker pens and the ability to project either a desktop computer or a tablet computer, onto the wall (see Figure 3). They were also given blank paper and pens, sticky notes and various paper-based templates along with reference materials, which included maps and curricula resources. The students made no use of the paper-based tools; at no point did any of the group use, or even make reference to using, the paper-based resources provided. All activity revolved around the digital projection onto the white-wall; even the use of the physical marker pens was influenced by the projected image.

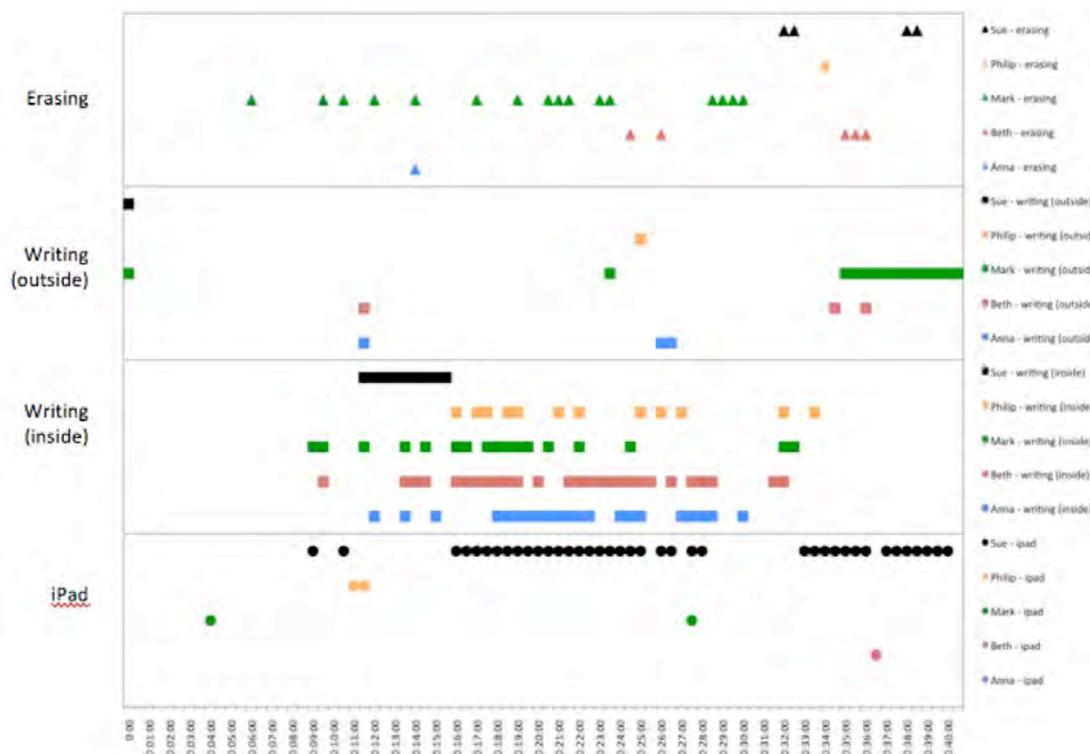


Figure 1. Tool use by participants over time

Figure 1 shows tool use during their collaborative design task. Each student is represented by a different color, and each section of the graph represents the use of a different tool. Circles, at the bottom, show the use of the iPad. Sue (black) was the main user of this tool. Other students had brief turns, Mark (green, 4, 27 minutes), Philip (yellow, 11 minutes) and Beth (red, 36 minutes). The next two sections on Figure 1 refer to writing on the

whitewall, inside and outside of the space illuminated by the iPad projection (Figure 3). All members of the group undertook this activity; there were periods of time when four of the five students were writing on the whitewall in the projected space simultaneously. Some work was performed outside the projected space, the most notable from about minute 34 by Mark (green). Finally, erasing activity is shown in Figure 1. While all members erased markings from the whitewall, Mark was primarily responsible for this activity.

Based on an analysis of the video recording, the session was divided into five phases; setup, briefing, design (I), technical problems, and design (II). Initially Sue requested that a computer (used by another group, operated via a wireless keyboard and mouse) be projected onto the whitewall. Her opinion was that the task would be simpler with a desktop system rather than an iPad. Sue took control of the computer keyboard in anticipation that the group would accept her request. When the group elected for the iPad to be projected, Sue relinquished control of the iPad to Mark. Sue soon asked Mark to hand her the iPad, which he did. During setup and briefing, the iPad was not used significantly; it was mostly used during the two design phases. In both of these phases Sue was, by far, the dominant user. During the task, there were only two minutes of the design work in which Sue relinquished the iPad; and during this time she was actively writing on the whitewall.

The students recorded their ideas by writing and drawing with both the physical marker pens on the whitewall and the digital pens on the iPad. These physical and digital tools were used in combination. Even though the entire wall was available, the students mainly confined their design work to an area formed by the projection of the iPad. As this space was relatively small, it became necessary to keep a record of what was written on the whitewall before erasing to make more space. Sue copied the text from the whitewall onto the iPad, making a permanent digital record of the writing. This text could be projected onto the whitewall, from the iPad, after the original had been erased. By taking control of the iPad, Sue effectively controlled the pace of the design task and also had some autonomy over which sections of the written text were preserved in digital form.

At various times, one or more students broke away from the accepted practice of writing within the projected area to build up an independent section on the whitewall. This could be regarded as a prototyping area, a place to sort out ideas before committing them to the projected area (and hence into the iPad), or as a method of rebelling against the group with ideas that had not been accepted.

Sue was an active member of the group. Once the task began, she took control of the iPad and dominated the creation of the digital artifact. Twice Sue relinquished the iPad; during the first, she became very active writing on the whitewall, effectively controlling the wall space, and the other was when the iPad had a technical fault. Sue stopped using the device when it was no longer being projected onto the whitewall. During this time, Sue did not use the iPad, however she did hold onto it, releasing it for technical assistance.

The use of the tools in relation to the development of ideas was visualized (Figure 2). Each photograph was also coded to illustrate any contributions the students made to the development of the *content* through writing, in addition to discourse. Figure 2 shows the ideas (the larger circles), with each individual idea represented by a horizontal line. Photographs of the whitewall were coded for the content of what was drawn, written and displayed. The CPACS *content* codes were phatics (x.1), tools (x.3), planning (x.5), topic (x.7) and task (x.9). Each person is represented by a different color, in addition to the five members shown in Figure 1, Steve, the facilitator, is also included.

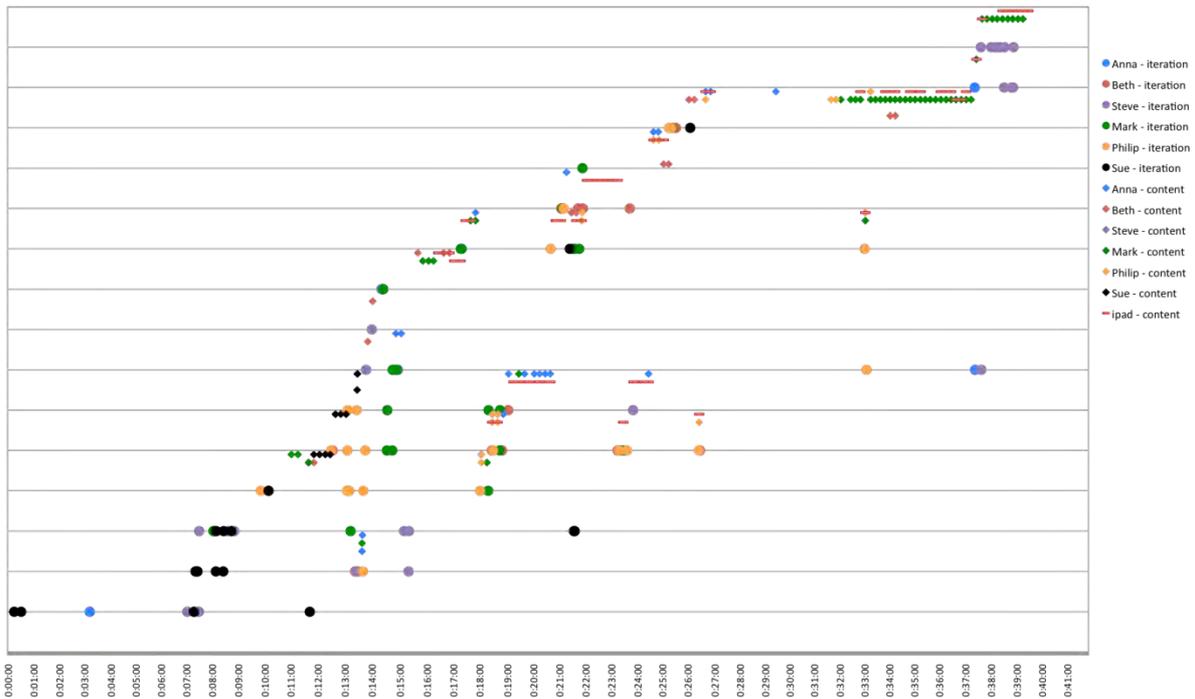


Figure 2. Additions made by participants during idea development, according to the content added over time

Figure 2 shows that the design process was not linear, that students often returned to earlier ideas and continued discussion. The whitewall was mostly used for recording information about the *task* and the *topic*, rather than for planning, phatics, or further reference to the tools. Often, the iPad was used just before a new idea was suggested; this could be because students were waiting for decisions to be recorded before they moved on. There are two occasions that were chosen for in-depth analysis of sketching practices due to the distinct patterns in students' use of the tools. The first, *collaboration*, occurs from 18 minutes to 23 minutes, when students simultaneously write on the whitewall while the iPad is also being used. It is at this time that they return to the ideas that were generated earlier in the collaboration. The second, *specialization*, occurs from 31 minutes to 37 minutes, when two members of the group dominate the use of two of the tools, for different purposes (Sue, with the iPad, focuses on the task, and Mark, on the whitewall, focuses on the topic).

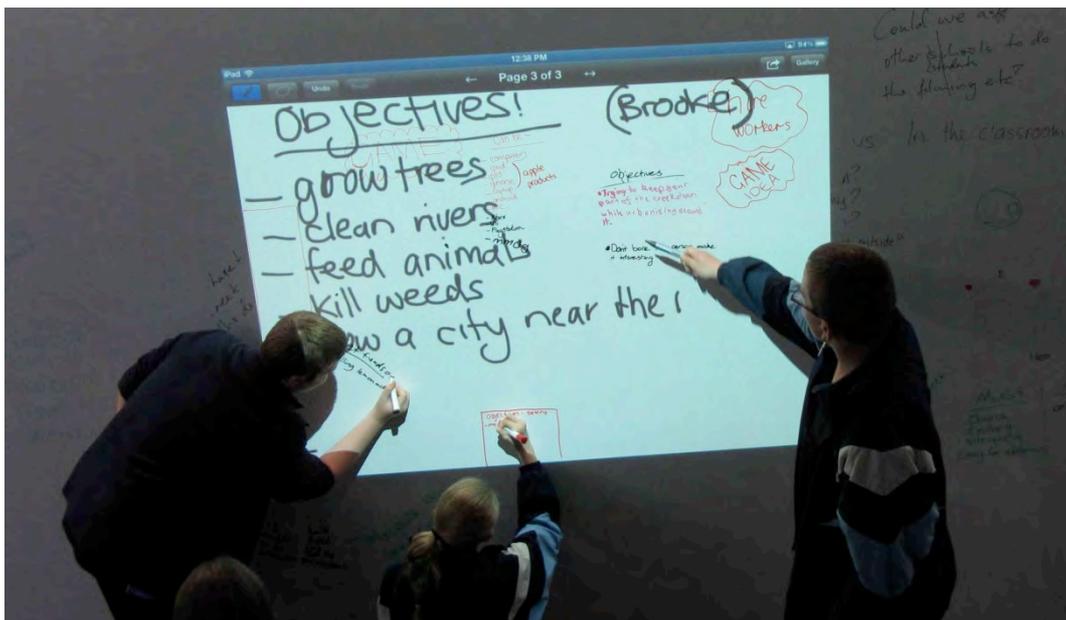


Figure 3: Collaboration

The *collaboration* analysis included an indepth analysis of 30 photographs, taken every 12 seconds (images 4248-4277). The students (shown in Figure 3) are collaborating to produce the 'objectives' for their idea of a game. It is notable that the students demonstrate behavior similar to practicing designers in many design fields. For example, the students use a variety of ways, using the available tools, to communicate their ideas to the other members of the team. Their verbal expressions are supported by drawings and text on the whitewall. This is similar to 'talking sketches' produced in the early stages of engineering design where ideas are exchanged and negotiated using sketches, text and diagrams (Ferguson, 1992). Visual representations, such as these drawings and marks on the whitewall, hold the design ideas stable so they can be argued and negotiated.

Students made several types of marks on the whiteboard during *collaboration*. A bulleted list of possible platforms for the game was produced, including computer, iPad, and Xbox. Some of the main ideas are written in capitals and enclosed by irregular shapes. Important terms and headings are underlined for emphasis. Groups of ideas, such as similar device types, are bracketed to illustrate similarities and relationships. At one point, lines were drawn across the bottom of the screen, projected on the wall, to symbolise the creek.

At this initial stage, ideas were still developing as they were being negotiated; some ideas were further developed and some were erased and replaced by new versions. For example, Phillip adds (and underlines) the words 'get funds' on the whiteboard. Beth develops the idea by adding 'selling lemonade' underneath. Phillip realizes that Beth has misunderstood the idea and continues with his line by adding 'off entrepreneurs'. The vagueness of the ideas and words on the whitewall does not seem to be an issue of concern for the students. In fact, it stimulates the conversation as the students try to make sense of the design task. This is a typical situation in most design fields where visual representations of ideas are initially vague and not fully explained but become clearer as the designers' thoughts progress (Hansen, 2000). The drawings and text on the whitewall have a fluid status. They represent ideas that are still changing. Sue is in control of the iPad and once an idea is relatively crystallized she commits the idea to digital ink, thereby giving the idea more permanence. After Sue copies the ideas onto the iPad, the other team members erase the text on the whiteboard wall.

Another similarity with typical design practices that this session demonstrates, is that drawings and marks on the whiteboard can be considered as either individually owned or as a shared entity (Eppler, 2007). For example, at one stage Anna decides to write her objectives for the game on the board by commandeering one section of the available space and drawing a line around it (see the lower, centre area of the projected image in Figure 3). She makes sure that the other students know that these are her ideas and they are not to be changed or erased. This is an example of individual ownership of a visual representation. The rest of the representations on the whiteboard are mostly shared where anyone can add to them; no one has declared their ownership.

This instance also shows how design representation can take both physical and digital formats (Eppler, 2007). This is clearly seen by observing how Sue records the team's ideas using the iPad, which is projected onto the wall, while the rest of the team draw and write on the whitewall with physical pens. The interplay between these two tools, the iPad and the whitewall, is worth noting. All of the team members are at once able to see the wall and what is being added to the iPad. They often comment on the text added to the iPad and even suggest more additions to the perceived permanent, master version residing on the iPad.

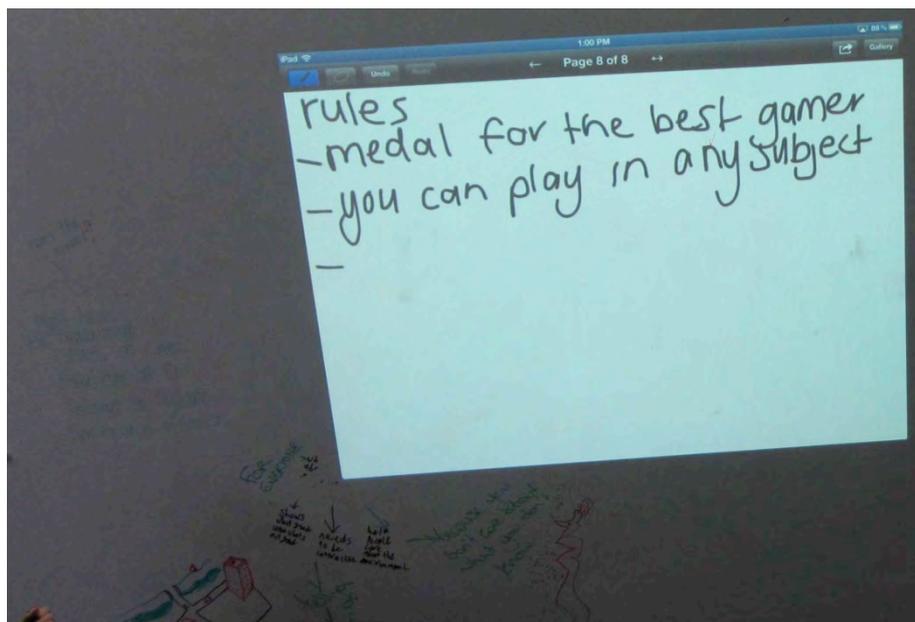


Figure 4: Specialization

Later in the session, the design needed to be finalized and a set of objectives put in place. This *specialization* session was analyzed using 25 images, taken every 12 seconds (4334-4359). Shown in Figure 4, ideas and representations on the whitewall were slowly erased as they were stored on the iPad. The idea of permanence given to the digital form is obvious here as the previously individually owned, and fiercely contested ideas on the wall are erased and replaced by the compiled list on the iPad. For example, Beth tried hard to protect her ideas written on the wall throughout the session. At one stage she proclaimed “*don't rub out my ideas*”. Nevertheless, after Sue copied her ideas onto the iPad, Beth took the eraser and wiped the text from the whitewall herself. This is typical in most design areas where the initial sketches and representations are disposed of in favor of more formal representation of the design, often expressed in textual format (McGown, Green, & Rodgers, 1998).

While this sorting and storing stage was underway, Mark started drawing on the whitewall away from the collaborative area. Mark's drawings did not appear to be intended for sharing with the group since they were drawn outside of the screen area illuminated by the iPad projection. Mark was visualizing the previous ideas that he shared on the whitewall under the heading ‘Objectives’. These ideas were about “*trying to keep your part of the creek clean while urbanising around it*”. Mark visualized this idea by drawing a creek in blue pen with roads surrounding it in black pen. He also drew high-rise buildings in red and added grass areas in green. Mark did not share or discuss his drawings with anyone else. This seems to be an example of a ‘thinking sketch’ (Ferguson, 1992). Thinking sketches are often drawn by designers in order to better understand a situation and to develop personal ideas. Mark appears to be drawing for a similar purpose; he used the whitewall, away from the projected screen, because the drawing was not meant to be a permanent record; it was not intended for sharing with the group.

Conclusions

The collection of multiple streams of data allowed a detailed, multimodal analysis of the processes of collaboration. Visualisations of tool use, and design processes over time allowed the recognition of patterns, and directed the in-depth analysis of sketching practices of students. Students used the tools available to them in complex ways and for different purposes – they used the whitewall for ideas development and the iPad for recording in a more permanent way. The iPad became central to their collaboration. When the iPad failed, the group also ceased writing on the whitewall. Perhaps without the ability to create a permanent record, they were reluctant to move on, despite the availability of the whole wall. During the 40 minute collaborative design task, students developed expertise, and adopted specialized roles within the group. This was seen most clearly with Sue and Mark. Only one person could control the iPad at any one time; multiple people could use the white-wall simultaneously. Figure 1 shows significant overlap at the wall, with multiple people contributing to the task. This may indicate that the task was indeed collaborative in nature and that the use of the iPad was more pragmatic (to keep records of the ideas) than an attempt to control the design task. All students were key in the development of ideas, demonstrated in their collaborative activity at the whitewall. The ownership of ideas was important, as was the requirement for a permanent record of them to exist.

This paper represents the initial analysis of a large data set, collected over several meetings. Future work will include analyzing the other groups who participated in the project, and tracking the development of their ideas to the final brief given to the multimedia designer. Understanding behavior associated with the intersection of the social interactions of students, the physical and digital tools, and the development of ideas as part of the design process is vital to the design of *learning by design* projects in the future.

Endnotes

- (1) The Design Studio is a multimedia educational design research facility at the University of Sydney: <http://sydney.edu.au/research/stl/facilities/EDRS/index.shtml>
- (2) Retrieved from: <https://dschool.stanford.edu>

References

- Berger, J. (2007). *Berger on drawing* (2 ed.): Aghabullogue, Co. Cork : - Occasional Press.
- Buxton, B. (2007). *Sketching User Experiences: Getting the Design Right and the Right Design*. Redmond, Washington: Microsoft Research.
- Eppler, M. J. (2007). Toward a visual turn in collaboration analysis? *Building Research & Information*, 35(5), 584-587. doi: 10.1080/09613210701355724
- Eppler, M., & Burkhard, R. (2006). *Knowledge Visualization Encyclopedia of Knowledge Management*: Idea Group Inc.
- Ferguson, E. S. (1992). *Engineering and the mind's eye*. Cambridge, Mass: MIT Press.
- Goldschmidt, G. (1991). The dialectics of sketching. *Creativity Research Journal*, 4(2), 123-143. doi: 10.1080/10400419109534381

- Goodyear, P., Jones, C. & Thompson, K. (forthcoming) Computer-supported collaborative learning: Instructional approaches, group processes and educational designs. In Spector, M., Merrill, D., Elan, J., & Bishop, M. J. (Ed) *Handbook of Research on Educational Communications and Technology*.
- Hansen, Y. M. (2000). Visualization for Thinking, Planning and Problem Solving. In R. E. Jacobson (Ed.), *Information design* (pp. 193-220): MIT Press.
- Henderson, K. (1999). *On line and on paper: visual representations, visual culture, and computer graphics in design engineering*. Cambridge, Mass: MIT Press.
- Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to Learn about Complex Systems. *The Journal of the Learning Sciences*, 9(3), 247-298.
- Hokanson, B. (2008). The Virtue of Paper: Drawing as a means to innovation in instructional design In L. Botturi & T. Stubbs (Eds.), *Handbook of Visual Languages for Instructional Design: Theories and Practices* (pp. 76-89). New York Information Science Reference.
- Kennedy-Clark, S. & Thompson, K. (accepted). Using multimodal discourse analysis to identify patterns of problem solving processes in a computer-supported collaborative environment. In Sudweeks, F. & Lim, H. (Eds) *Innovative Methods and Technologies for Electronic Discourse Analysis*.
- Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S., et al. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design into practice. *The Journal of the Learning Sciences*, 12(4), 495–547.
- McGown, A., Green, G., & Rodgers, P. A. (1998). Visible ideas: information patterns of conceptual sketch activity. *Design Studies*, 19(4), 431-453. doi: 10.1016/s0142-694x(98)00013-1
- Suwa, M., Gero, J., & Purcell, T. (2000). Unexpected discoveries and S-invention of design requirements: important vehicles for a design process. *Design Studies*, 21(6), 539-567. doi: 10.1016/s0142-694x(99)00034-4
- Suwa, M., & Tversky, B. (2002). External Representations Contribute to the Dynamic Construction of Ideas. Paper presented at the Proceedings of the Second International Conference on Diagrammatic Representation and Inference.
- Thompson, K., Ashe, D., Carvalho, L., Goodyear, P., Kelly, N., & Parisio, M. L. (accepted). Processing and visualizing data in complex learning environments. *American Behavioral Scientist*.
- Thompson, K., Ashe, D., Yeoman, P., & Parisio, M. (accepted). Phases of design: Following idea development and patterns of collaborative discussion in a *learning by design* project. *The 10th Int'l Conference on Computer Supported Collaborative Learning*, Madison, Wisconsin.
- Thompson, K., & Kelly, N. (2012a). Combining collaboration spaces: identifying patterns of tool use for decision-making in a networked learning environment. Paper presented at the *Eighth International Conference on Networked Learning*.
- Thompson, K., & Kelly, N. (2012b). Processes of decision-making with adaptive combinations of wiki and chat tools. Paper presented at *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012)*.
- Vattam, S. S., & Kolodner, J. L. (2008). On foundations of technological support for addressing challenges facing design-based science learning. *Pragmatics & Cognition*, 16(2), 406–437.
- Zenios, M., & Holmes, B. (2010). Knowledge creation in networked learning: combined tools and affordances. Paper presented at *the 7th International Conference on Networked Learning 2010*.

Acknowledgments

This work was funded by the Australian Research Council grant FL100100203. We would like to acknowledge the intellectual contribution of Peter Goodyear to the ideas developed in this paper, as well as the dedicated work of the students, teachers, and other experts who participated in this project.

Using automated and fine-grained analysis of pronoun use as indicators of progress in an online collaborative project

Kate Thompson, CoCo Research Centre, The University of Sydney, kate.thompson@sydney.edu.au
Shannon Kennedy-Clark, Learning and Teaching Centre, Australian Catholic University, shannon.kennedy-clark@acu.edu.au

Nick Kelly, CoCo Research Centre, The University of Sydney, nick.kelly@sydney.edu.au
Penny Wheeler, Learning and Teaching Centre, Australian Catholic University, penny.wheeler@acu.edu.au

Abstract: Multimodal discourse analysis has been shown to be useful in adding depth to our understanding of the processes of computer supported collaborative learning. We take one of the codes used in a multimodal coding scheme (CPACS), and apply automated data extraction techniques to a large corpus of data aimed at one code in the micro-level (pronouns). The results of this are plotted over time, and patterns of pronoun use identified for further investigation using an in-depth systemic functional linguistics (SFL) approach. Complications concerned with singular and plural second person are discussed, and patterns of pronoun use that indicate movement of the group from one phase of design work to the next are identified. Further refinement of the techniques of automated extraction are required to capture additional patterns noted in the SFL analysis.

Introduction

Multimodal discourse analysis adds depth to our understanding of the processes of computer supported collaborative learning. The use of a coding scheme such as CPACS (Kennedy-Clark & Thompson, accepted) combines macro- and micro-level approaches to discourse analysis in a systematic way. In this paper our focus is at the micro-level, particularly the use of pronouns in online group interactions. Other work has focused on the macro components of CPACS, such as decision-making or content (see Kennedy-Clark & Thompson, accepted). We apply automated extraction of pronoun usage to a well-researched corpus of data from a month-long online group collaboration on a design task. Patterns of pronoun use were identified, and an in-depth analysis from a systemic functional linguistics perspective was applied. We related these micro-level patterns to the macro-level patterns previously identified in the data (Reimann, 2009; Thompson & Kelly, 2012).

Background information

Without entering into the debate of the actual definition of a pronoun, personal pronouns can be viewed as a person marker (Kummerow, 2012): that is, as a part of speech, pronouns identify speaker(s), hearer(s)/addressee(s) and ‘spoken of’. One of the functions of pronouns is to establish a continuity of reference; that is, they show how locally constructed reference forms are used to construct relationships between episodes and sequences of events over time: as Kitzinger, Shaw & Toerien (2012) state, “People don’t just find themselves having a focus – rather, they create it” (p. 133). An analysis of the use of person markers is of interest in CSSL as it enables researchers to construct a visual representation of how participants arrange their actions and relationships within an online group through the choices they make in presenting themselves and others in the discourse. Traditional English grammars distinguish pronouns in terms of person, number, gender (in the third person) and case (Table 1):

Table 1: Traditional grammatical categorization of English pronouns

SINGULAR				PLURAL		
Person	Subject	Object	Possessive	Subject	Object	Possessive
First	I	me	mine	we	us	ours
Second	you	you	yours	you	you	yours
Third	he / she / it	him / her / it	his / hers / its	they	them	theirs

Systemic functional linguistics (SFL) focuses on description and emphasizes choices within a specific situation or cultural context (the ‘systemic’ in its title summarizing the concept that each instance of writing or talk or other text is a selection from a set of alternative features which are systematically related to each other by the particular language or language subset). The foundation of SFL is four connected observations on language, that language is (Eggs, 2005): (i) *functional*, that it serves a use; (ii) *semantic*, that its function is to make meaning; (iii) *contextual*, that meaning is influenced by social and cultural context; and (iv) *semiotic*, that using language involves a choice of which signifiers to use. These principles encourage an “activist bent” (Jacobs & Ward, 2000) in its practitioners and educators, who can, for example, help students examine unfamiliar

academic registers as the first step to mastery, or reveal how language enables justice to be restored or to fail to be done, although SFL analysis can also be used, as here in an examination of CSCL, for humbler purposes.

In SFL, personal pronouns can be analyzed as key features in the ‘interpersonal’ metafunction of the text, “enacting our personal and social relationships with the other people around us” (Halliday & Matthiessen, 2004, p. 29), and as linked to other interpersonal components such as the Mood element (which in English is made up of the Subject and Finite operator of a clause, and realizes the declarative, interrogative, or imperative mood of a clause); modality (the uncertainty or possibility in a clause, for example ‘can’ or ‘must’), and the role of clauses in offering and exchanging services and information. The other two linguistic metafunctions in SFL theory are the ideational, which looks at the categories and processes being identified and experienced by the speakers, and the textual, those parts of a text that provide cohesion and organization for the discourse as a whole. Personal pronouns can be examined for the agency that they realize in the clause, and also for their cohesive role in providing a continuous thread of reference to the participants in the conversation, written text or, as here, computer-mediated chat.

Word counting has been used in text analysis in the social sciences for many decades. Two notable examples are DICTION (Hart, 2001) and LIWC (Chung & Pennebaker, 2012; Pennebaker & Francis, 1996), both of which use pre- and user-defined dictionaries to count and subsequently analyze corpora of text for features such as attentional focus, emotionality, social relationships, thinking styles, and individual differences (Tausczik & Pennebaker, 2010). Natural language processing has significantly advanced over the past few decades (Jurafsky & Martin, 2008; Manning & Schütze, 1999), and advanced techniques have recently begun to be applied to the learning sciences, such as the mapping of words in vector spaces to discover use of concepts through clustering (Sherin, 2012) and the use of parts-of-speech (PoS) tagging for segmentation and classification. In this work we make use of the Python programming language (Bird, Klein, & Loper, 2009) to apply some of these techniques to our analysis of pronouns used during CSCL. We use these tools as aids in reconstructing the contextual use of person markers in relating to other members of the group, before more closely examining their functions in this context via a grammatical analysis based on systemic functional linguistics.

Methods

The corpus of data used in this analysis is the online chat record of a group of four postgraduate students (one male, three females) collaborating over one month on a design task. The task was to identify an existing system dynamics model and turn it into an educational resource both by adding to the model and providing materials to accompany the model. Students had access to a chat and a wiki to support their work. Previous research has described the data collection (Reimann, Weinel & Thompson, 2007) and the patterns of decision-making in the chats (Reimann, Frerejean & Thompson, 2009) as well as in the wiki (Thompson & Kelly, 2012). This paper returns to focus just on the chat data. The collaborative exercise took place over one month in 2006; students met as a whole group on eight occasions, and 2182 utterances were used in the automated process.

The Collaborative Process Analysis Coding Scheme (CPACS) was used to code the synchronous chat data. (For a full description of the coding scheme, see Kennedy-Clark & Thompson, accepted). This coding system considered several SFL phenomena, such as those described by Halliday (1994) and elaborated upon by Martin and Rose (2007). CPACS functions on two levels. At the macro level, the action code concerns goal identification and solution, and the content code covers the content of the utterance. At the micro level, the attitudinal code describes the type of attitude each utterance was taking towards the problem-solving design; the tense is the marker of temporality; the modality is the degree of certainty; and pronouns relate to personal references. In this study we use the pronouns code: 1 (first person singular); 2 (second person singular); 3 (third person singular); 4 (first person plural); 5 (second person plural); and 6 (third person plural).

We hypothesized that by combining the macro elements of discourse (e.g. Poole & Holmes, 1995) with the micro or grammatical levels that are often the target of discourse analysis (e.g. Nivre et al. 1999) we could obtain a better understanding of the conversations than if we analyzed the macro levels only, using a coding system such as the Decision Function Coding System (Poole & Holmes, 1995) as we had done in our earlier analysis (see, for example Kennedy-Clark & Thompson, 2011; Thompson, Kennedy-Clark, Markauskaite, & Southavilay, 2011). We found that by drilling down into the discourse and examining the conversations at a finer grain size we could get a detailed understanding of the function of language in establishing roles and responsibilities within a group.

Naturally occurring classroom language and patterns of discourse produced by the students were mined for linguistic and social insights about how the online groups managed their shared task through the language that they used (Gee, 2005; Schallert, et al., 2009). SFL analysis was carried out by considering correspondences in the data between the variables of DFCS, personal pronoun use, and author, and investigating the language in use at this point. To carry out these analyses, one researcher worked on the utterances as filtered by the quantitative analyses, and provided additional columns of annotation on specific linguistic aspects, namely: modality; mood of the clause, that is, imperative, interrogative, or declarative; annotations to the automated

pronoun identification data, where, for example, the author has omitted the Subject that should be understood and produced an elliptical utterance, particularly by omitting a first person singular pronoun; cohesion, by marking the chain of reference for each neutral third person pronoun used; and process type, from the three principal categorizations in SFL of material; mental; and relational processes.

These general categorizations are ones that are immediately evident to the reader, and these types of annotations are rapidly made: by working directly on the filtered data source, the researcher can compare the occurrence of different linguistic aspects within the target variables, counting occurrences and looking for 'marked' clauses, that is, linguistic constructions which show a non-typical structure. This level of human linguistic analysis was possible because the preceding automated analysis and visual interpretation of the quantitative aspects of the linguistic data reduced the number of segments and sequences that require individual assessment, and because the categorizations are at the most general level of SFL analysis. Aspects of appraisal and analysis of circumstantial elements (such as realizations of location, manner, or cause), both of which can provide deep insights into interactions and experiences as represented in more highly structured oral and written texts, did not form part of this SFL analysis.

Analysis

Automated extraction of data and identification of patterns

Two phases of automation were used in the analysis of this corpus of chat text⁽¹⁾. The first involved counting pronouns, whilst the second used PoS tagging to assist in understanding the word counts. Pronouns function as both person markers and indications of focus. A count of the pronouns uttered by each author gave an overview of how the author positioned their own experience of the collaboration, whom they addressed during the group exchanges, and when they addressed entities outside the exchange. The count was conducted using a Python script to iterate through chat text and count each utterance. The pronouns counted were each assigned to a CPACS pronouns code: 1 (my, me, mine); 2 (you); 3 (her, him, it, he, she); 4 (us, we, ours) and 6 (they, them). The result of this analysis was a count of each CPACS code for each author. A limitation of this automated approach to pronoun counting is that it loses the context of the utterance. This means that it always assumes that the use of a pronoun by a participant is in their own voice, when in some cases it may be a quote of something that another participant or third party has said. An analysis based upon use of the words 'said' and 'says' shows that this use of another voice occurs 23 times in the corpus.

Further analysis was possible by breaking these numbers down in two ways: (i) seeing how they changed over time; and (ii) seeing how they changed within context. Students met in eight separate chat sessions and separating the count of pronouns in each session allowed for the analysis of changes that occurred over time. In addition, each utterance within this corpus of text had previously been grouped into five tasks: choosing, adding, implementing, overall and coordinating (Reimann, Frerejean & Thompson, 2009). The counting of pronouns was indexed to these categories, allowing for an analysis of the changing use of pronouns within these different contexts.



Figure 1: visualization of pronoun use for each participant, for each session

Figure 1, in combination with results from coding and counting not presented here, was used to identify instances of variation. Figure 1 shows the percentage of pronoun use (the CPACS code is along the y-axis,

session number along the x-axis) for each member of the group. Each section (for example, between 1 (me, my, our) and 2(you)) represents 100% of the pronouns extracted. An issue with second person is that the same word is used for both singular and plural, and so all second person was recorded as '2': this is the first complication to be discussed below. Five patterns were observed during examination of the extracted pronoun data. (1) Pronouns and tasks: the combined percentages were different when examined across tasks (not presented in Figure 1). Higher percentages of use were observed in coordination and implementation than choosing or additions. Within coordination, Charles and Jane used this more than other members, while Beth had the highest proportion in implementation. (2) Use of first person changes over time: if we examine the patterns across time, within each pronoun code, it can be seen that the use of first person, while remaining quite consistent, differs amongst the group members. (3) The task of choosing had a different distribution of pronouns than other tasks: coding and counting results indicated that with respect to task, choosing had a different distribution amongst team members than other tasks. (4) Session 5 was different to other sessions with respect to pronoun use: Figure 1 shows a different distribution amongst team members in Session 5. (5) First person plural is used differently over time: Sessions 7 and 8 show a slightly lower percentage across all group members than other sessions. The use of third person showed a higher proportion apparent when participants were discussing implementation.

'You' – plural or singular?

PoS tagging (Mu, et al., 2012; Schmid, 1994) was utilized in the deeper analysis of the word 'you' within the corpus. Because English no longer has a singular and plural form of the second personal pronoun, a simple count of the occurrences of 'you' does not distinguish between the author's focus upon one person or upon some or all of the people in the exchange. A Python script was used to apply PoS tagging to the corpus⁽²⁾, to produce an output consisting of, each time the word 'you' is uttered; the two words before and after this utterance of the word 'you', and the PoS indicated. The assumption in doing this was that the PoS tagging would reveal patterns that could allow for automation of the context of the word 'you', and it was partially successful in this, highlighting two linguistic features particularly relevant to the successful conduct of computer-mediated discussion: a marked use of the vocative, and the conversational workarounds that English speakers use to indicate plural reference for the 'you' pronoun. In the latter case, the PoS tagging clearly revealed the uses of 'you both' (1 instance), 'you all' (4 instances), 'you guys' (25 instances: note this reference is non-gender-specific) as well as 'you girls' (6 instances) and 'you ladies' (4 instances), both of these authored by the sole male speaker within the group of 4 collaborators, and all of these examples of the author disambiguating the 'you' pronoun to indicate all addressees in the context of the discussion in order to check for their consent, for example, or direct their action. (In subsequent human analysis, reviewing the PoS tagging output, an additional refinement of this distinction between 'you' singular and plural was identified in the use of 'u', a text-chat specific variant of the pronoun, which was used by one author as a singular 'you' reference only.) The PoS tagging also highlighted where authors were employing the vocative as a means to target their utterance to only one of their collaborators: a few examples of this linguistic choice, one frequently found in this corpus, are 'jane, do you have that link handy?' (segment 14); 'what d'ya say beth and jane?' (98); 'beth do you what [want] to start with your questions?' (2958); 'lydia, jane, what about you guys?' (5065); 'beth you mentioned you had some resources' (7743); and 'charles what do you think about beths question?' (6212).

This marked usage of the vocative in CSCL is paralleled by a similarly marked rejection of the use of personal pronouns in favour of names as the subject or object of a clause. For example, in standard conversation or writing, the context often makes unnecessary the specific naming of your partner in an action: it is the de-contextualised nature of CMD that makes it preferable to choose 'lydia and me' instead of standard 'we' in an utterance such as 'In that way Lydia and me should find additional resources' (2989) or to use a name rather than a third person pronoun to emphasize to which person's utterance your response refers: 'my that's right is with charles answer' (6510) or to reject 'you', when Jane and Lydia are both present, in favour of 'we need to see if jane or lydia want to add anything / amalgamate anything' (7279).

Guiding behavior

An important way in which personal pronouns provide a shortcut to an analysis of what is going on in the discussion is by indicating the kind of exchanges occurring in the text, and the salient personal pronoun is 'you'. Figure 1 shows that both Charles and Jane are strong users of 'you', particularly in segments that are coded as overall and coordination phases in the DFCS. SFL sees an analysis of what is being 'exchanged' by these 'you' clauses as a fundamental indication of the roles and statuses adopted by the authors of these clauses.

Clauses can be structured as an exchange of information or as an exchange of goods-and-services: most simply, the choices can be represented as giving information (making a statement) or giving goods or services (making an offer); or as demanding/requesting information (questioning) or demanding goods or services (issuing a command, also discussed in the previous section). An example of how clauses function in these exchanges can be seen in Table 4.

Table 4: Utterances from Session 5, Jane's use of 'you' as part of overall decisions (planning, trouble-shooting)

Author	Clause	Role in exchange	'You' refers to:
Jane	did you tell lydia about what Kate told you guys ?	question – asking for information	'you' (here, Sayer and Subject), 'you guys' (here, Audience) = Beth and Charles
Jane	have you guys seen the storytelling feature how it works?	question	you guys (Subject) = Beth, Charles and Lydia
Jane	beth you were interest in seeing how storytelling works see the immigration and Easter Island models	statement, command ('see the ... models')	you (Subject) = Beth
Jane	but I don't see how lydia and I help you guys	==	you guys = Beth and Charles
Jane	I could work with you charles	offer	you = Charles

It is clear, in tracing the referent(s) of the 'you' pronoun, that additional mapping beyond pronoun counting and PoS tagging would be required to precisely outline who was working with whom, but what is evident from looking at the clauses as exchange is the role that the author can variously assign herself/himself.

Another example where the distribution indicates a strong use of 'you' is Charles (see Table 5), in coordination mode, requesting goods-and-services from others in the group. It appears that his earliest demands on his group members (here, Beth) are couched as a request rather than a requirement, but he does not use this strategy in later sessions.

Table 5: examples of Charles' use of 'you' in coordination

Session	Clause	Role in exchange
1	sure beth. are you going to link it in the wiki?	Structured as a request for information, this is actually functioning as a command (asking for Beth's action).
2	jane, do you want to put in your response to beth's questions first	Request for information in structure, by implication a command for action, an instruction
2	if you like set up a forum.	Suggestion: a command ('set up a forum') modulated by a concessive clause
2	jane, can you also archive the notes	Command structured as a yes/no question
3	for sure beth! you get some sleep girl!	Command
4	ok. first let me tell you that you have made a huge contribution to the team	Both a command (for 'you') and an offer (from 'me'): 'let's'
6	ok. jane. were you ok about tomorrow night at 10:30?	Request to attend, structured as a yes/no question
7	can you give me a link?	Request for action

The examples shown in Table 5 are in fact not the typical clause structures for realizing a command, as they are marked for person. Usual or unmarked imperatives cannot be identified in an automatic count of pronouns, as they do not contain the Subject component of the clause: take, for example, the instances of unmarked imperatives from Charles in the first session (see Table 6).

Table 6: Imperative clauses (unmarked): author, Charles, coordination

Session	Clause	Role in exchange
1	check page now	Command
1	bullets in now = refresh	Command (= 'Now, refresh!')
1	check now!	Command

Making suggestions

We can extend the discussion of the clause as exchange by examining a marked difference in first person pronoun use within the 'choosing' component (DFCS coding) for team members than in the other DF tasks. Using SFL tools to probe this more closely, we can identify the linguistic choices accompanying this first person pronoun use (chiefly, the use of 'I' or 'we' in focus or thematic position) as contributing to the successful use of 'suggestion'.

There are two choices that the group must make to progress with the collaborative task: the selection of model, and the selection of audience or class level. The selection of model comes first, and, it being early in their joint history (completely contained in Session 1), group members are initially unwilling to baldly assert a choice: instead, they make suggestions. The linguistic choices required for successful suggestions are seen by SFL as part of the interpersonal resources of language, and, matching pronoun use to DFCS choosing, and following this with human analysis, we found that strategies used in suggestion and negotiation include a modulation in the Finite verb used when 'I' was the Subject (for example, using modals with a lower representation of certainty ("Charles: i agree. having a look at the some the models we might just need to pick the one that best suits us"): mitigation of the imperative mood (that is, softening the command, as in "Beth: not yet, but come prepared with one each and then vote, i suppose" (segment 9); or, "Beth: how about we give each other a day or two" (segment 3); and omission of first person singular pronouns to conceal who is the Actor, in SFL terms, in these processes (as Charles does repeatedly as he reports his additions to the wiki, perhaps concealing his dominant role here: "Charles: just updated that element"; "Charles: just put in the link to the October 5th archive", and so on).

Creating active artifacts

A higher pattern of third person usages within segments coded as implementation warrants a closer look at these occurrences. The English third person singular pronoun distinguishes its referents by genders, and by simply using the filters in the pronoun count view, it is apparent that the human referents for these pronouns are straightforward: 'he' = the supervising male teacher for the activity, 'she' = the female teacher or, in one sequence, one group member who was temporarily absent. More frequent is the neuter pronoun 'it', and, in those utterances related to the task of implementation, the referent is the group's collaboratively created digital artifact and/or the task of creating this artifact. For example, when Beth directs the group's attention to the visuals and sound effects that she has added to the model in the wiki, they respond:

Jane: **it's** very good, girls (10238)
 Charles: will just open **it** now (10239)
 Beth: **it's** a little rough...i need to adjust **it** now i've spoken to Kate (10240)

Beth: and then **it** goes into storytelling mode... and by pressing the space bar, the 'story' appears step by step according to the sequel in trac (10252)
 Lydia: the female learners will luv **it** yah :) (10253)

By this stage in the operation of the group, the design task is getting close to completion, and this is indicated by the increased reference to the model they have worked on as one of the 'participants', that is, in SFL transitivity terms, one of the entities involved in the events or processes being referred to, and its emergence as a possible subject and Actor (as in 10252).

Contributing independently and within the group

The pronoun count indicates a decrease in use in first person plural in the later sessions, and a follow-up by a human reader indicates that in part this is because the participants have organized themselves into pairs, and report on the pair's activity as, for example, 'lydia and me'. The addition of a broad classification of process types (the major categories of SFL transitivity analysis, that is material, mental, verbal and relational processes) associated with the first person plural pronoun reveals that the actions the group is depicted as taking are not often material, and certainly not physically effective processes. For example:

Charles: we did some refinements [Material]
 Charles: we came up with a "test some values and come up with something better as an activity basically [Material]
 Charles: can we get away with two? [Material]
 Charles: I think we can probably whip another scenario up pretty quickly. [Material]
 Lydia: shd we add a short story to explain the stock and flow concept? [Material]

These clauses with Material processes are most importantly realizing the actions of the group members in *creating* or extending the enhanced model: 'came up with'; 'whip up'; and 'add'.

In clauses with relational processes, the pronoun 'we' is more likely to be referring to the group of four participants as a whole (Table 7). The relational process subsystem of 'possession' is realized strongly here, as well as the idea of representation, an identification process again related to the idea of creation. Metaphors of process here are relational processes realizing mental states: 'got any ideas' = 'thinking'; 'are we on the sane [same] page' = 'agree'

Table 7: Examples of relational and verbal processes from the chat

Author	Clause	Process type
Jane	I think we all have that impresion	relational
Charles	we need the stocks to be different for each scenario.	relational
Lydia	if we represent the scenarios with graphs,	relational
Jane	in trac we have limited space to upload files	relational
Jane	charles are we going to have 1 or two scenarios	relational
Charles	got any ideas for it?	relational
Lydia	are we on the sane page?	relational
Charles	we are nearly there!	relational
Jane	should we talk about our progress ?	verbal
Beth	lydia, we need to talk about the text	verbal
Charles	so we still agree that [we get them done on Wednesday]	verbal
Lydia	is there anything else we need to discuss?	verbal
Charles	we chat more tomorrow	verbal

By contrast, the first person singular pronoun is extensively employed in these last two sessions, particularly in segments coded as ‘implementing’ in the DFCS, and this may indicate that the group members are operating independently, having negotiated particular sets of tasks during the earlier sessions of choosing, planning and coordinating. Looking at the processes which the group members self-describe themselves as performing in broad SFL categories, it is evident that ‘I’ is associated with more effective material actions. Note the over-representation of Beth in using ‘I’: she has a significant share of the utterances coded as implementation and particularly utterances using ‘you’ as the object of a request for help. The other types of processes done by ‘I’ in these sessions are, in the segments coded for implementation, evenly split between mental, verbal, and relational, with just a few instances from each category.

Conclusions

This initial combination of methods of analyzing this data has indicated areas for follow-up research in other corpuses of data. There is some evidence provided here that changes in patterns of pronoun use could indicate changes in phases of design work (from choosing to additions, and also the beginning of implementation). In several instances, automation did not identify the patterns identified by human raters. There is significant scope for further application of automation to systemic functional linguistics in education and refinement of this work, as was carried out with ‘you’, and reported in this paper. One of the key areas for development is in the calibration of the diverse methods available with human input. Once we know what patterns to look for, we will be better able to identify automated techniques that could be used to help students manage their own progress, as well as instructors to monitor the progress of groups through their collaborative work. Future work too needs to tie these findings more explicitly to the macro-level of discourse analysis.

Endnotes

¹ All scripts used in the analysis were written in the Python programming language and are available by request

¹ The Python package Natural Language Toolkit was used for the PoS tagging, specifically the algorithm trained on the Wall Street Journal corpus of over 270 000 words (Bird, et al., 2009)

References

- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*: O'Reilly Media, Incorporated.
- Chung, C. K., & Pennebaker, J. W. (2012). *Linguistic Inquiry and Word Count (LIWC): Pronounced “Luke,” and Other Useful Facts Applied Natural Language Processing: Identification, Investigation and Resolution* (pp. 206-229): IGI Global.
- Eggs, S. (2005). *Introduction to systemic functional linguistics*: Continuum.
- Gee, J. P. (2005). *An Introduction to Discourse Analysis: Theory and method* (2nd ed.). London: Routledge.
- Halliday, M. A. K. (1994). *Introduction to Functional Grammar*. London: Arnold.
- Halliday, M. A. K. & Matthiessen, C. (2004). *An Introduction to Functional Grammar* (3rd ed.). London: Arnold.

- Hart, R. P. (2001). Redeveloping DICTION: theoretical considerations. *Progress in communication sciences*, 43-60.
- Jacobs, G. M., & Ward, C. (2000). Analysing Student-Student Interaction from Cooperative Learning and Systemic Functional Perspectives. *Electronic Journal of Science Education*, 4(4). Retrieved <http://ejse.southwestern.edu/article/view/7639/5406>
- Jurafsky, D., & Martin, J. (2008). *Speech and Language Processing* (2nd Edition) (Prentice Hall Series in Artificial Intelligence): Prentice Hall.
- Kennedy-Clark, S., & Thompson, K. (2011). What Do Students Learn When Collaboratively Using A Computer Game in the Study of Historical Disease Epidemics, and Why? . *Games and Culture*, 6(6), 1-25.
- Kennedy-Clark, S. & Thompson, K. (accepted) Using multimodal discourse analysis to identify patterns of problem solving processes in a computer-supported collaborative environment. In Sudweeks, F. & Lim, H. (Eds) *Innovative Methods and Technologies for Electronic Discourse Analysis*.
- Kitzinger, C., Shaw, R., & Toerien, M. (2012). Referring to Persons Without Using a Full-Form Reference: Locally Initial Indexicals in Action. [Article]. *Research on Language & Social Interaction*, 45(2), 116-136. doi: 10.1080/08351813.2012.673376
- Kummerow, D. (2012). The Person That Isn't: On Defining the Third-Person, Negatively and Positively So. *Australian Journal of Linguistics*, 32(2), 259-290. doi: 10.1080/07268602.2012.669099
- Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*: MIT press.
- Martin, J. R., & Rose, D. (2007). *Working with Discourse: Meaning Beyond the Clause* (2nd ed.). London: Continuum International.
- Mu, J., Stegmann, K., Mayfield, E., Rosé, C., & Fischer, F. (2012). The ACODEA framework: Developing segmentation and classification schemes for fully automatic analysis of online discussions. *International Journal of Computer-Supported Collaborative Learning*, 7(2), 285-305. doi: 10.1007/s11412-012-9147-y
- Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition & Emotion*, 10(6), 601-626.
- Poole, M. S., & Holmes, M. E. (1995). Decision Development in Computer/Assisted Group Decision Making. *Human Communication Research*, 22(1), 90-127.
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4, 239-257.
- Reimann, P., Frerejan, J. & Thompson, K. (2009). Using process mining to identify models of group decision making in chat data. Paper presented at the CSCL Conference, Rhodes, Greece.
- Schallert, D. L., Chiang, Y.-h. V., Park, Y., Jordan, M. E., Lee, H., Janne Cheng, A.-C., Song, K. (2009). Being polite while fulfilling different discourse functions in online classroom discussions. *Computers & Education*, 53(3), 713-725. doi: 10.1016/j.compedu.2009.04.009
- Sherin, B. (2012). Using computational methods to discover student science conceptions in interview data. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24-54.
- Thompson, K., & Kelly, N. (2012). *Processes of decision-making with adaptive combinations of wiki and chat tools*. Paper presented at the The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012).
- Thompson, K., Kennedy-Clark, S., Markauskaite, L., & Southavilay, V. C. (2011). Capturing and analysing the processes and patterns of learning in collaborative learning environments. Paper presented at the 9th International Conference on Computer Supported Collaborative Learning. CSCL2011: Connecting computer supported collaborative learning to policy and practice, Hong Kong.

Acknowledgments

Part of this work was funded by the Australian Research Council grant FL100100203.

Phases of design: Following idea development and patterns of collaborative discussion in a *learning by design* project

Kate Thompson, David Ashe, Pippa Yeoman, Martin Parisio, CoCo Research Centre, The University of Sydney, Building A35, The University of Sydney, NSW 2006

Email: kate.thompson@sydney.edu.au, david.ashe@sydney.edu.au, philippa.yeoman@sydney.edu.au, martin.parisio@sydney.edu.au

Abstract: Learning by design (LBD) has a long association with learning about complex environmental systems. This investigation traces the development of ideas within a group of five students engaged in a collaborative design process. Tasked with the design of an online educational resource, about a waterway of local significance, this group was one of three for which multiple streams of data (audio and video) were collected. Ideas central to the progression of their design were identified and represented visually over time, showing the impact of each group member and the facilitator, and discourse was coded according to the content code of the CPACS scheme. Four phases of design were identified and Markov-transition diagrams of the content were interrogated. This paper makes a contribution to our knowledge of the phases of design evident during LBD tasks, which could have implications for the design and management of such projects in the future.

Introduction

Engaging school students in tasks that support collaboration, technology enhanced learning, and the understanding of complex systems is challenging. As part of a project funded by a local government organization, a group of school students engaged in the design of a learning resource to be shared with other schools in their region. The group included 16 students across multiple year groups and the focus of the project was a waterway of local significance. The students were observed and guided as they worked through a design process structured to help them learn about issues critical to water management in their area. This paper begins with a brief overview of the existing research in the field of *learning by design* (LBD) and the processes of learning about complex systems. The analysis examines one group of students during a forty-minute session. Their task was to develop ideas about their design in order to present the design concept to the group as a whole. Attention is given to the way in which the design progressed during this session with a focus on the initial development of ideas, the use and reuse of ideas, the way the task, topic, tools, and social interactions were communicated during verbal exchanges; and the role of the facilitator in the process. The implications of this in-depth analysis are discussed in terms of what is known about the processes of learning during a design task.

Literature Review

Learning by design (LBD) is the blending of what is known about case-based reasoning, with what is known about problem-based learning, resulting in a project-based inquiry approach to science learning (Kolodner et al., 2003). A case is described as a “contextualized piece of knowledge representing an experience that teaches a lesson fundamental to achieving the goals of the reasoner” (Bergmann, Kolodner, & Plaza, 2006, p. 209). LBD leverages what is known about promoting deep and effective learning by situating it in activity that is both purposeful and engaging. Involving learners in the process of design facilitates their movement between evaluation and creation. Designing presents an opportunity for deep learning to occur because student-designers are required to use their knowledge of natural systems to build an artificial working replica of a functional system (Hmelo, Holton, & Kolodner, 2000). In order to complete the task the designer needs to: (1) select an approach, (2) understand nuances in function and behavior, (3) sequence and interrelate multiple functions, and (4) evaluate compliance with functional requirements. The process of building a working replica, rather than an appearance-model, presents the learner with a project that will inevitably require multiple iterations. With each iteration the learner is presented with an opportunity for reflection on current levels of fidelity requiring them to revisit their knowledge of the natural system. Linn (1996) describes how navigating between the model as studied and the model as currently under construction helps build an ever more sophisticated understanding of the system being studied.

Vattam and Kolodner (2008) describe LBD as internally driven by a “need to know” and a “need to do”. During this iterative process students share experiences and ideas as they articulate what they will need to learn in order to successfully complete the design challenge. Numerous LBD studies have illustrated positive learning outcomes for students (Bamberger, Cahill, Hagerty, Short & Krajcik, 2010; Fortus, Dershimer, Krajcik, Marx & Mamlok-Naaman, 2004; Hmelo, Holton, & Kolodner, 2000; Sadler, Coyle, & Schwartz, 2000), and the role of providing support for these projects with paper-based reflective journals and computer based scaffolding (Domeshek & Kolodner, 1994; Guzdial, 1998; Kolodner et al., 2003; Puntambekar & Goldstein, 2007;

Puntambekar & Hubscher, 2005; Puntambekar & Kolodner, 2005; Vattam & Kolodner, 2008). In investigating the role of the teacher in LBD tasks, Puntambekar & Stylianou, (2007) highlight the need for students to make connections between the design activity and the learning, both of which were evident in classrooms where teachers helped students to connect prior learning to the topic studied, and where they assisted students in the generation of goal-related questions. Despite positive gains across a number of areas, Vattam and Kolodner (2008) identify two significant challenges to the implementation of design-based science learning (DBSL): (1) the need to bridge the design-science gap, and (2) finding a way to manage time and material constraints. They investigate software solutions that integrate explanation-construction scaffolding with modeling and simulation, and conclude that their strategy enhanced collaborative understanding and social construction of knowledge in DBSL environments. In a more recent study Bamberger, et al. (2010) reveal that students who engaged in LBD tasks were better able to understand scientific content and, in particular, the workings of scientific systems.

Generally, systems are characterized as having components or definable elements, interactions or interrelations between the elements; and in open systems such as ecological systems, fluxes across the system boundaries (Reimann & Thompson, 2009). To understand the local creek and its surrounding environment, students in this study needed to consider all of the components of the system as well as the fact that the system may change over time (Limburg, O'Neill, Costanza, & Farber, 2002; Ossimitz, 1997). Another characteristic of complex systems is emergence, where aggregate level structures affect the behavior of the elements of which they are composed (Wilensky & Reisman, 2006). Many learners have trouble understanding complex systems even when they are illustrated using tools such as models. It was hypothesized that a *learning by design* project, in which the design task was to create an educational resource, may result in a greater understanding of, the connections between elements of the system, as well as changes over time, and emergent features of the system.

Visualizing and analyzing the processes of learning is a relatively new area in the learning sciences. Reimann's (2009) seminal work outlines the importance of time and order in considering the processes of learning. Generally, work in this field has concentrated on decision-making (Reimann, Frerejean & Thompson, 2009; Kapur, 2011), and has used a variety of methods of analysis, such as heuristics mining (Reimann et al., 2009), first-order Markov models (Thompson & Kelly, 2012), and hidden Markov models (Southavilay, Yacef, & Calvo, 2010). In this study, we examine the processes of *design* in a learning context, with a focus on the development of ideas and the content of the discourse (using Kennedy-Clark & Thompson's CPACS (accepted).

Methods

A group of 16 students and 11 adults participated in the project funded by a regional organization of councils, and supported by their high-school and a local environmental rehabilitation organization. The aim of the project was to develop an innovative fieldwork and multimedia framework to engage other students in water and land management issues. South Creek, the focus of the project has been described as the most degraded creek in the region. Threats to its health include vegetation clearance, urbanization and agriculture; resulting in concerns about increased nutrient levels and an increase in the number of weed species within the riparian zone. Over a number of months students and stakeholders participated in a multidisciplinary design process to create a learning resource for use in schools across the region. The students participated in planning sessions, a site visit, and a day of hands-on site restoration at the creek, before attending a *Design Day* at the University of Sydney.

The *Design Day* was an opportunity for the students to develop their ideas about the design of the educational resource, propose possible formats, identify constraints and generate a consensus upon which a brief for the multimedia designer could be written. Participants included expert learners (the students), who ranged in age from 12 to 17 years, and experts from education, environmental science and multimedia design. The *Design Day* began with each of the expert groups outlined their desires and constraints for the design of the educational resource, and these initial parameters were summarised and referred to during the rest of the day. In the Design Studio (the Design Studio is a multimedia educational design research facility at the University of Sydney: <http://sydney.edu.au/research/stl/facilities/EDRS/index.shtml>) participants were led through the first three stages of a design process: *empathize*, *define*, and *ideate* (following the Stanford University Institute of Design – An Introduction to Design Thinking (<https://dschool.stanford.edu>)). The *empathize* stage took the form of a whole group brainstorming activity during which the initial desires and constraints were discussed and the critical components extracted. Participants worked in pairs for the *define* stage; each dyad performed a needs analysis to help define the resource. During the third stage, *ideate*, the participants worked in groups according to discipline area. One group contained all the adults including educators, multimedia designers and environmental scientists. The other three groups each included between five and six students – expert learners. During the *ideate* stage, participants were asked to generate ideas. They were asked not to limit themselves to their knowledge of technology and were instructed to record all ideas. The intention was to explore a wide solution space so that, later in the *Design Day*, these ideas could be distilled into one coherent solution and a brief presented to the multimedia designer for the creation of a resource prototype.

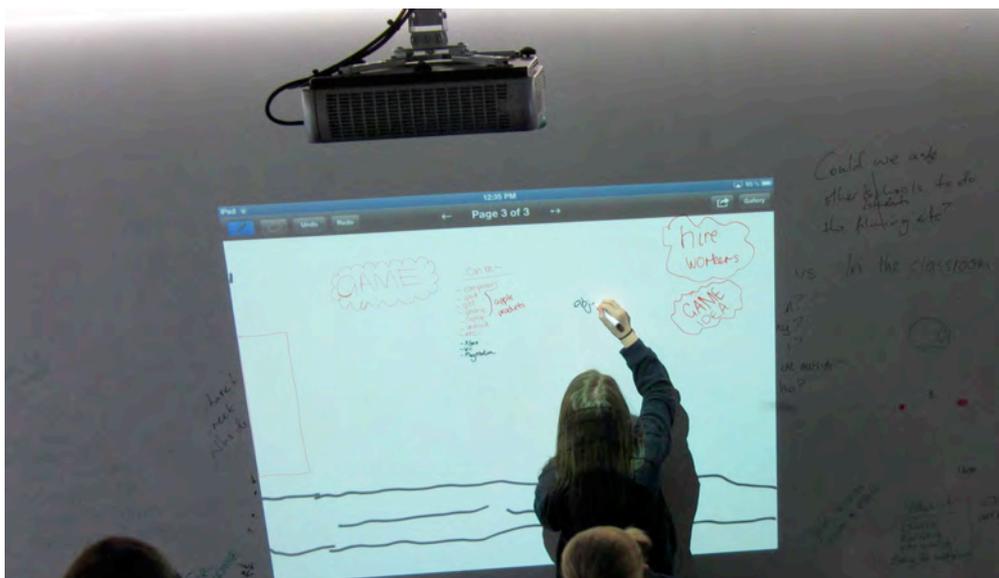


Figure 1: Students working on the design task

This paper follows one of the student groups during the *ideate* stage of the design day. There were five members of the group, three females and two males, whose ages ranged from 12 to 16. Video, audio and photographs of the collaborative design work were collected. Each group was given a choice regarding the digital tool they used (computer, interactive whiteboard, or iPad). This group chose to use an iPad that was projected onto a wall painted in whiteboard paint (a ‘white-wall’, see Figure 1). The transcription of their discourse was analyzed in order to identify ideas important in the development of their design. These were then discussed and agreed upon by all authors of the paper in a group meeting. The Collaborative Process Analysis Coding Scheme (CPACS) was used to code the transcript. CPACS is a multi-level coding scheme that includes macro-levels (action and content) as well as micro-levels (pronouns, tense, modality, and attitude) of discourse; only the macro-level code, *Content*, was used in this paper (Kennedy-Clark & Thompson, accepted). The Content section of CPACS contains six codes: phatics (social – phatics, salutations, leave taking), planning, topic, task, tool use, and off-task. Initial agreement between raters was 52%, after discussion 96% agreement was achieved. Other work using the *Content* code has shown that a periodic oscillation between phatics, tool use, planning, topic, task, with all elements included, is indicative of successful collaborative work. This tends to correlate with observable patterns in other macro-level processes, such as decision-making (Kennedy-Clark & Thompson, accepted). The generation of Markov transition probabilities has been shown to be a useful tool to visualize the patterns of decision-making (Thompson & Kelly, 2012) and content (Kennedy-Clark & Thompson, accepted) in discourse, and will be used in the analysis presented here. Markov transition diagrams illustrate the probability of each state transition (for example, from topic to task), and are appropriate for processes in which there is an expected order of states.

Analysis

The ideas considered to be important to the development of the group’s design were identified in a transcript of the group’s discourse. Selection was based on key descriptors of the final design, rather than methods of implementing the design, or examples of existing games the students used as inspiration. They are presented in Table 1, below.

Table 1: Idea development

Idea	Description	Proposed by
1	Original illustration	Sue (Mark)
2	A computer game	Sue
3	Access to technology	Steve
4	“taking care of”	Philip
5	Managing	Philip
6	The Creek	Philip
7	“challenge others”	Steve/Mark
8	Levels of difficulty	Steve/Anna
9	“a player in the game”	Anna/Steve

10	Gaming platforms	Mark
11	Different game “views”	Beth
12	Role playing	Mark
13	Violence	Philip
14	Score, points and awards	Anna
15	School subject and class	Steve

There were 15 ideas that were directly relevant to the development of the group’s design. The number also represents the order in which the ideas were suggested. Not all ideas moved the group forward in their design in the same way. For example, three of the ideas (“*taking care of*”, *Managing*, and *The Creek*) occurred in quick succession; they are closely related and were suggested by the same person (Philip). Violence, however, was less related to other ideas and could almost be considered tangential. It is possible that this was an idea stimulated by a discussion taking place in one of the other groups. There are many ways in which the ideas could be rated or classified; however, for the purposes of this analysis all 15 have been assumed to be of equal importance to the final design. Table 1 also shows the name of the group member who first put forward an idea. Steve, the facilitator, played an important role during the design process and on more than one occasion he reflected an idea suggested by one student, helping the others to recognize the contributions of all to the design process and the final product.

Reporting the counts of *content* codes for each person, or for each idea, adds little to the understanding of the design process. Instead, the *content* codes were plotted over time, for each participant (all five students and the facilitator), taking the generation of ideas into account. This is presented in Figure 2.

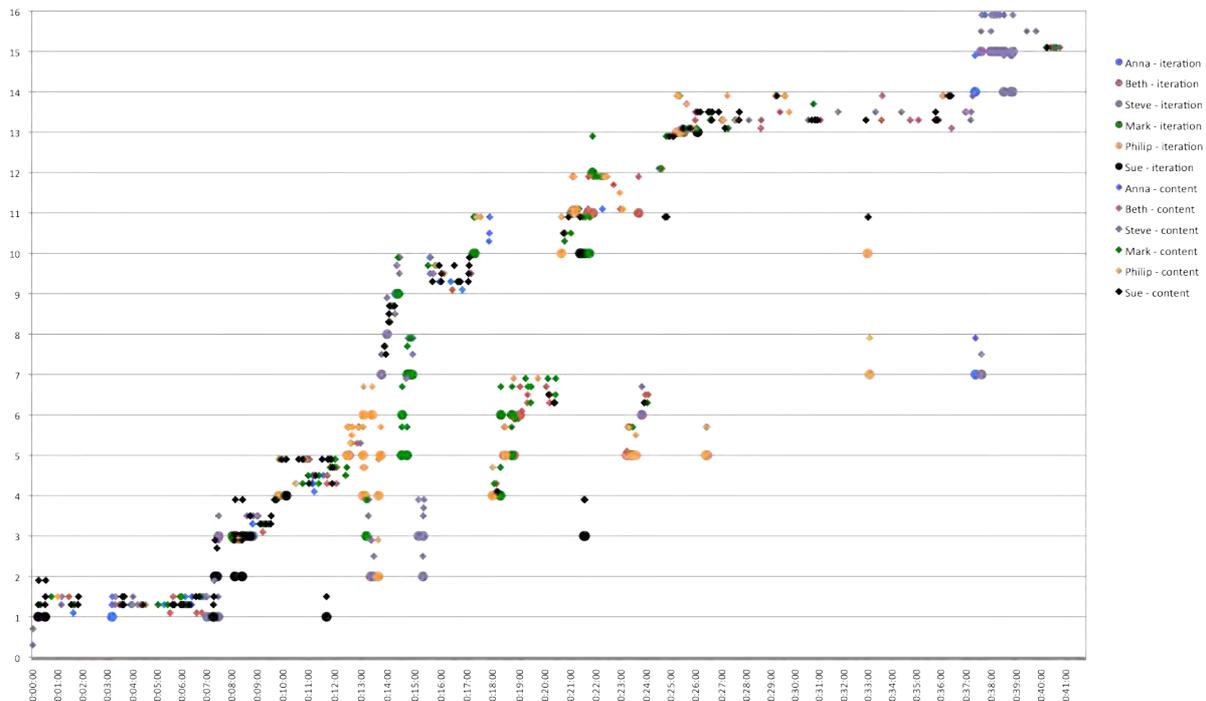


Figure 2: Iterations of the design, and associated conversation

Figure 2 shows the development of ideas during the process of design, for each member of the group, and the *content* of discussion. Time (from 0 to 41 minutes) is shown on the x-axis, and each new idea (see Table 1) corresponds to the number on the y-axis. Where an idea is built on, or developed, the preceding codes are only repeated if the reference clearly takes the conversation back to an earlier idea. We did this for two reasons: (1) the notion of design assumes a cumulative and revised whole; (2) it was those moments where iterations emerged that we examined more closely. Each participant has been given a different colour. Circles are the utterances that include reference to a specific idea, and diamonds are the utterances that have been coded with *content*. Each *content* code was given a numerical value: phatics – x.1, tool use – x.3, planning – x.5, topic – x.7, and task x.9. For example, at about 12 minutes, participants had generated ideas to number 4 (“*taking care of*”), and the subsequent discussion revolved around the task, and the tool (*content* codes): Sue was active in this discussion (black). At about this time, she also mentioned an earlier idea, idea 1 (*the original drawing*), however subsequent discussion by the rest of the group was related to the idea of “*taking care of*” the environment,

specifically the creek. In order to explain the patterns observed in Figure 2, the video of the collaboration was analyzed in detail.

The collaborative design work analyzed here opens with a longer period of discussion, following on from the initial idea. This consists of activity and conversation that seeks to organise and orientate the group in terms of physical space, available tools and the task at hand. It includes discussion about location and centers on an illustration drawn by Sue and Mark before the official ‘start’ of the group work. Sue starts by making use of the visual representation of their efforts thus far, she uses words economically but, from the beginning, her presence can be seen in how she records ideas and scaffolds the group’s thinking. There are times when Sue takes ownership of ideas; however, her influence can be seen predominantly in how she maintains ownership of the tools and space in which she records ideas and how she returns the group to the objectives of the task. Anna who seems reluctant to start without adult help precipitates the second phase of the design work. It is apparent in the recordings that she is engaged at this stage of the collaboration; however, the rest of the group seems not to hear her suggestions or her questions. During this time, the phrase: ‘*what are we supposed to do?*’ is used by many members of the group. It is Anna who approaches the facilitator, Steve, and he responds with a series of questions that directs the group to value and develop the ideas they have already had. Phrases such as: *You are stealing all these ideas* (which referred to the results of the earlier *define* stage); *What are you actually designing?*; *There’s some ideas* (pointing again to the white-wall); *It’s got to be fun so what is it?*; *It was your must have thing that you just drew up so what does it do?* (this question was asked of Sue); and *What are you actually designing?* Sue continues the discussion with reference to what she has drawn and following further goal-oriented questioning from Steve, says “It’s a computer game”. This moves the group to the second phase of their design.

The activity that follows includes familiarizing themselves with the iPad and the orientation of its projected screen, allocating space within the projected iPad screen on the wall, selecting whiteboard markers and trying them out. It also includes an extended discussion about games and gaming platforms. Knowledge of, and access to, different digital devices and games stimulates discussion about the design task which leads to documentation of objectives on the white-wall. It is not until Steve (the facilitator) returns that Philip justifies their discussions by relating it back to the project. After this, a chain of closely related ideas is put forward that includes *taking care of the creek* and *managing* a game. This ‘stacking of ideas’ (seen in Figure 2 between minute 14 and minute 16) is led by Philip who suggests several ideas but seldom dominates the subsequent conversation. The rapid turn-taking in the discourse, starting with Steve at idea number 2, is carried on by the students, particularly between Philip and Mark to idea number 5. At this point, Anna, who can be seen observing Beth, Mark and Sue sketching characters on the wall, gains Steve’s attention for feedback on their progress. Steve continues using the techniques he employed earlier, goal-oriented questioning and seeding ideas. He says: *Ok, so are you yourself in this game or do you choose a character?* Mark answers with: *you as a character in the game* and adds the notion of *personal challenge*.

The next observable phase in the collaborative design work can be seen in Figure 2, as idea 12 (*role playing*) is suggested for the first time, utterances related to the task are further apart, and there is no obvious pattern to their distribution. As the discussion returns to *gaming platforms*, *role playing* and *violence*, Philip says *Make it a violent game*. Mark objects, saying that he doesn’t like violence or violent games, and attempts to get support from other members of the group. During this time, the social interactions are important and threaten the cohesion of the group. Anna directs negative comments at Mark, and Beth attempts to distract and refocus the group by embellishing their sketched characters on the wall. This marks the end of the most productive period of idea generation, all of which occurred in a relatively short amount of time.

The final phase, during the discussion around idea 13, is marked by the failure of one of the tools. During this time, the projection of the iPad onto the white-wall ceases to work, although the iPad itself still works and the entire wall is available for writing on. At this point, members of the group seem distracted by the other groups also working in The Design Studio, and physically withdraw from the space in which they had been working. Even when the iPad projection is restored, some members do not return to the shared space. Mark’s absence is clearly visible in Figure 2, from the point where the notion of violence is introduced, which results in a heated verbal exchange after which he withdraws altogether from participation in idea generation. During this time he attempts to take ownership of the iPad in order to record ideas, however, Sue retains control. After trying to help restore the iPad projection to the wall he sits down against the wall - and draws a city near a river. Philip and Sue work on documenting objectives with some help from Beth and again it is not until Steve returns that a few more ideas are generated and a conclusion is reached.

The collaborative design process was divided into four phases, identified through in-depth analysis of the discourse in combination with the visualization of the content codes and idea generation over time. Phase 1 involved idea 1; Phase 2, ideas 2-10; Phase 3, ideas 11 and 12; and Phase 4, idea 13. In order to determine if there were discernable patterns to the discussion of content in these phases, Markov transition probabilities were calculated for each of these phases, and the diagrams can be seen in Figure 3.

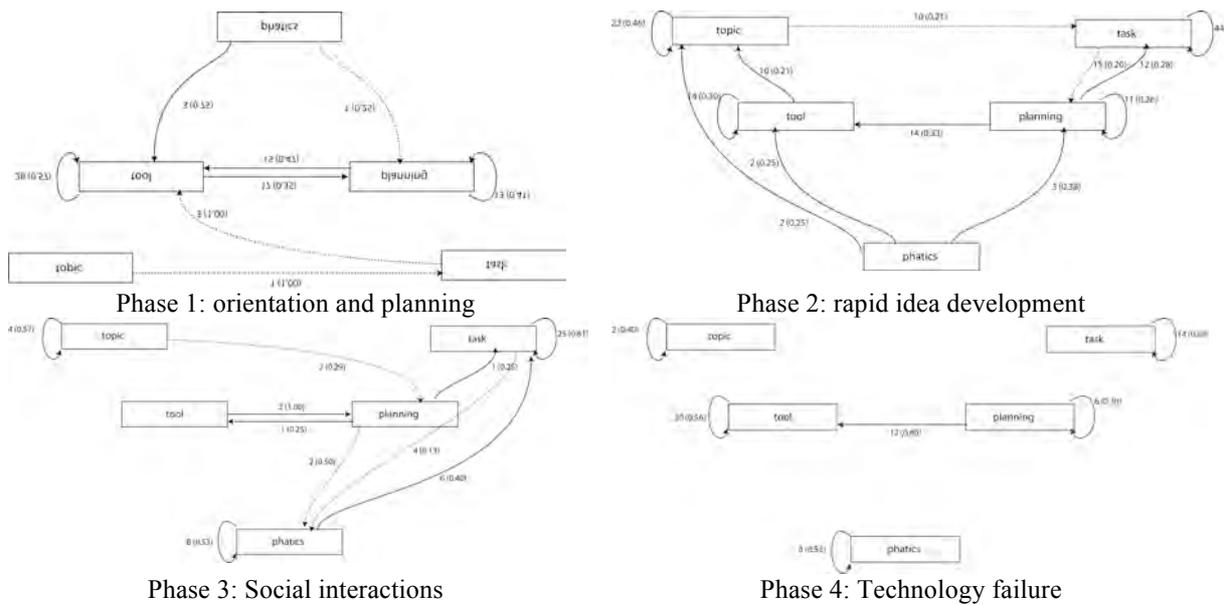


Figure 3: Markov transition diagrams for the four phases of design work

Figure 3 shows four distinct patterns of discussion with regards to content, during the design task (see Thompson & Kelly, 2012 for more on interpreting Markov transition diagrams). In each, the five elements of the content codes are displayed: phatics (the social interactions), tool (reference to the physical and digital tools provided), planning, topic (anything that would be seen in the game, for example the creek, hiring workers etc), and task (utterances related to designing the game). Links between these elements are shown only if the probability is greater than 0.25, or if the number of transitions is greater than ten.

In the Phase 1 transition diagram, there is a clear emphasis on planning and the tools that would be used. Socially based utterances (phatics) were followed by those connected to either tool use or planning. If members discussed tool use, they either continued discussing this (0.57) or moved to planning (0.35). Very little of this discussion was related to either the topic or the task to be performed. This aligns with the in-depth analysis, which identified this phase as an orientation and planning phase.

Phase 2 of the design work was described as rapid idea development. The transition diagram above shows a regular cycle through the elements of content: from planning, to tool use, to topic and then to task, with a return to planning. In all cases, when students begin to discuss each of these elements, they tend to focus on that element. The social interactions are not related to any one of these and this supports the in-depth analysis. This design work was characterized by productive discussion, with few distractions; the group members appear to be focused on their design work.

The in-depth analysis identified phase 3 as the end of idea development, and a focus on social interactions. The First-Order Markov transition diagram supports this finding. The link from phatics to itself (53%) shows that group members remained discussing this content. In addition, links from other elements to phatics (planning and task) shows that the members were distracted from the previous cycle of ideas development. The link between planning and tool use still exists, but for very small frequencies. Initial observations of the design work had led to an assumption that the failure of the iPad was the trigger for the group dispersal; however, the social interactions may have preceded this.

The key event in phase 4 was the failure of the projection of the iPad. During this time constructive work ceases as the group waits. In Figure 3, the transition diagram shows the only link between elements: from planning to tool use, presumably indicative of problem solving related to the technical issues. This figure indicates a lack of connection between elements of the discussion: the links were distributed between so many different elements that none had a high enough probability or frequency to be displayed.

Discussion and Conclusions

The aim of this paper was to describe the processes of design observed in a group of five students during a *learning by design* task. Multimodal analysis has provided insights into relationships between the patterns of the content of discussion, the generation of ideas, and the phases of design work. One measure of the success of this group was the organic nature of its designerly behavior. While we only concentrated on the idea generation stage of design in this analysis, Figure 2 showed that this group returned to ideas raised earlier in the design work. An important feature of the group's design process was the way in which the participants revisited earlier

ideas. This was done naturally, within the processes of conversation, and new ideas were built on the new, common understanding of the design.

Steve, the facilitator, was an important element in this system; his goal-oriented questions (Puntambekar & Stylianou, 2007) provided students with the opportunity to navigate between the design studied and the resource under construction (Linn, 1996). Opportunities for reflection were incorporated into the productive phase of idea generation without disturbing the generation of new ideas. Steve's role in connecting their activity to the task at hand, and his persistent use of goal-directed questioning, is clearly instrumental in aiding this group to develop their ideas over time. Their freedom to access and talk about gaming systems, with which they were already familiar, and the ease with which they could record, draw and build upon prior iterations (all with the express aim of designing a tool to connect other students with the project) was well supported in this environment.

Students' use of tools was a key part of their idea development and record keeping. Further analysis (Thompson, Ashe, Wardak, Yeoman & Parisio, accepted), investigated how the tools available in the dedicated design space effectively supported the collaborative design work of the members of the group. The tools were a central element of discussion and productive design work (Figure 3). They were an aid in collaboration allowing for the articulation of the ideas of individuals on the white-wall and in creating a space in which all ideas were combined in the more permanent record keeping on the iPad. Interestingly, it was the social interactions that seemed to interrupt the regular pattern of idea generation as was seen in the Markov diagrams.

The learning outcomes for students participating in this project relate to knowledge about water quality issues as well as an understanding of the process of design for learning. In the analysis presented, the focus has been on the social interactions and interactions with both physical and digital tools. However, the ideas concerned with managing the creek occurred early in the design work, and when students returned to these ideas, they added to them, rather than revising them. Implicit knowledge about the inputs and outputs from the ecosystem, impacts, and links between elements of the ecosystem was present in the development of their ideas. Knowledge about the ecosystem was present in their discourse, and in their drawings (Thompson et al., accepted), throughout the collaborative task. The students agreed that the ecosystem in question needed to be managed in terms of human impact on this system, and that there should be shared use; mostly for urbanization (most familiar to these students) rather than agriculture or industry. They put *themselves* in the pictures that they drew of the creek, and a strong personal connection between the creek and the group members was observed. There was no demonstration of an understanding of the global/local relationship in the ecosystem, nor specific reference to possible impacts of decreased water quality.

By extracting multiple streams of data, visualizing and analyzing the processes of learning, phases of idea generation, as well as patterns of discourse, were identified. Future research will apply these to the analysis of the other groups that participated. Comparisons between the groups will then be made using lag sequential analysis techniques. It is expected that the orientation phase would be common across all groups. One question to investigate is what moves the group from the orientation phase into design work. In this study, the facilitator played a significant role. Given that he visited all groups, his role can be further interrogated. A larger question for this research is what role the interaction between the task, the social interactions, and the coordination of tool use play in influencing the processes of learning.

The aim of this paper was to describe the processes of design evident in the interactions between group members and the tools provided in a *learning by design* task. We have shown that students were able to articulate an understanding of some (often challenging) ideas about systems. In addition the development of ideas as they related to the final design were identified and visualized. Through these visualizations, phases in the design work were identified that corresponded to recognizable patterns in the discourse. Social processes and interactions with tools were essential to the progression through the design process. The management of time and materials was identified as a challenge of LBD projects (Vattam and Kolodner, 2008). Part of the challenge is knowing when to intervene in a group's collaborative design work. Identifying naturally occurring phases, and the indicators of movement into and out of these phases is essential to managing this process.

References

- Bamberger, Y., Cahill, C., Hagerty, J., Short, H., & Krajcik, J. (2010). Learning science by doing design: How can it work at the middle school level? *Journal of Education, Informatics, and Cybernetics*, 2(2), 41-46.
- Bergmann, R., Kolodner, J., & Plaza, E. (2006). Representation in case-based reasoning. *The Knowledge Engineering Review*, 20(03), 209–213.
- Domeshek, E., & Kolodner, J. L. (1994). The design of a tool kit for case-based design aids. *Artificial intelligence and design '94* (pp. 109–126). Dordrecht: Kluwer Academic.
- Fortus, D., Dershimer, R. C., Krajcik, J., Marx, R. W., & Mamlok-Naaman, R. (2004). Design-based science and student learning. *Journal of Research in Science Teaching*, 41(10), 1081-1110.

- Guzdial, M. (1998). Technological support for project-based learning. *Association for Supervision and Curriculum Development*, 39–42.
- Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to learn about complex systems. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 9(3), 247–298.
- Kapur, M. (2011). Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of Computer Supported Collaborative Learning*, 6(1), 39–56.
- Kennedy-Clark, S. & Thompson, K. (accepted). Using multimodal discourse analysis to identify patterns of problem solving processes in a computer-supported collaborative environment. In Sudweeks, F. & Lim, H. (Eds) *Innovative Methods and Technologies for Electronic Discourse Analysis*.
- Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S., et al. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design into practice. *The Journal of the Learning Sciences*, 12(4), 495–547.
- Limburg, K. E., O'Neill, R. V., Costanza, R., & Farber, S. (2002) Complex systems and valuation. *Ecological economics*, 41, 409–420.
- Linn, M. C., & Muilenburg, L. (1996). Creating lifelong science learners: What models form a firm foundation. *Educational Researcher*, 25, 18–24.
- Ossimitz, G. (1997) *The development of systems thinking skills using system dynamics modeling tools*. Retrieved 6 April 2004, from http://www.uni-klu.ac.at/users/gossimitz/sdyn/gdm_eng.htm
- Puntambekar, S., & Goldstein, J. (2007). Effect of visual representation of the conceptual structure of the domain on science learning and navigation in a hypertext environment. *Journal of Educational Multimedia and Hypermedia*, 16(4), 429–459.
- Puntambekar, S., & Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, 40(1), 1–12.
- Puntambekar, S., & Kolodner, J. L. (2005). Toward implementing distributed scaffolding: Helping students learn science from design. *Journal of Research in Science Teaching*, 42(2), 185–217.
- Puntambekar, S., & Stylianou, A. (2007). Comparing classroom enactments of an inquiry curriculum: Lessons learned from two teachers. *Journal of the Learning Sciences*, 16(1), 81–130.
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4, 239–257.
- Reimann, P., Frerejan, J. & Thompson, K. (2009). Using process mining to identify models of group decision making in chat data. Paper presented at the CSCL Conference, Rhodes, Greece.
- Reimann, P. & Thompson, K. (2009). Ecosystem modeling for environmental education: From stocks and flows to behaviour and interactions. In Blumschein, P., Hung, W., Jonassen, D. & Strobel, J. (Eds.) *Model-based approaches to learning: Using systems models and simulations to improve understanding and problem solving in complex domains*, pp. 111–148. Sense Publishers, Rotterdam.
- Sadler, P. M., Coyle, H. P., & Schwartz, M. (2000). Engineering Competitions in the Middle School Classroom: Key Elements in Developing Effective Design Challenges. *Journal of the Learning Sciences*, 9(3), 299–327. doi: 10.1207/s15327809jls0903_3
- Southavilay, V., Yacef, K., & Calvo, R. A. (2010). *Analysis of Collaborative Writing Processes Using Hidden Markov Models and Semantic Heuristics*. Paper presented at the Proceedings of the Third International Workshop on Semantic Aspects in Data Mining 2010, in conjunction with the 10th IEEE International Conference on Data Mining (ICDM 2010).
- Thompson, K., Ashe, D., Wardak, D., Yeoman, P., & Parisio, M. (accepted). Identification of patterns of tool use and sketching practices in a learning by design project. *The 10th Int'l Conference on Computer Supported Collaborative Learning*, Madison, Wisconsin.
- Thompson, K., & Kelly, N. (2012). *Processes of decision-making with adaptive combinations of wiki and chat tools*. Paper presented at the The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012).
- Vattam, S. S., & Kolodner, J. L. (2008). On foundations of technological support for addressing challenges facing design-based science learning. *Pragmatics & Cognition*, 16(2), 406–437.
- 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

Acknowledgments

This work was funded by the Australian Research Council grant FL100100203. We would like to acknowledge the intellectual contribution of Peter Goodyear to the ideas developed in this paper, as well as the dedicated work of the students, teachers, and other experts who participated in this project.

Individualistic appropriation as a primary mechanism of collaborative conceptual change: a case study

Michael Tscholl, University of Central Florida, 12461 Research Parkway, Orlando, FL 32826, USA, mtscholl@cs.ucl.ac.uk

John Dowell, University College London, Malet Place, WC1E 6BT London, UK. jdowell@cs.ucl.ac.uk

Abstract: Collaborative learning with cases and problems is characterized by the contribution of disparate knowledge and varying interpretations. The way in which this public knowledge is exploited individually to construct a conceptualization of the problem is examined here. The paper presents a microanalysis of a collaborative case-centered learning dialogue between three learners where a novel conceptualization is constructed re-using selected surface and structural elements contained in a prior conceptualization. How the novel conceptualization is constructed is shown by tracing surface and structural knowledge in the sequence of contributions and by identifying the point-of-view adopted by the learners. We argue that a mechanism of individualistic appropriation accounts for this construction, and this mechanism may be central in collaborative learning. This entails a revision of the notion of co-construction in collaborative learning.

Introduction

In collaborative problem- or case-based learning, a divergence of views on how to conceptualize and frame the problem or case is a common occurrence. Groups' discussions are characterized by the effort to identify the right way to understand the problem, a process that may then produce the conceptual change we usually regard as learning. Through the sequence of discursive exchanges, different viewpoints on a problem will change and so also will the way in which problem aspects are conceptualized.

A fundamental research issue concerns how a change in conceptualization is achieved within a group and whether the change results from an individual developing it, or whether the group develops it jointly. An individual may put forward a view of the case or problem that may become the basis for the groups' conceptualization, or a different viewpoint may be proposed becoming the new basis. In other words, a conceptualization may *evolve* through transforming a prior conceptualization, or it may be created *de novo*, in part with selected elements of surface and structural knowledge contained in prior conceptualizations. By implication, effective collaborative problem-based learning may not depend on one person internalizing another's view; their own learning may develop through incorporating elements of the views put forward by others, possibly then resulting in the advancing of alternative conceptualizations, which may then be recognized as correct (given a specific task context) and adopted by others.

This paper is a case study of a discussion between three medical students working in a virtual learning system (featuring a chat tool and a means to navigate the learning material) to interpret a case of alleged medical negligence concerning a patient's consent. The students were tasked to interpret the case within given concepts of medical law, interpreting the facts of the case and relating them through consideration of those concepts. Individual viewpoints are central to our analysis and by identifying the viewpoints of learners' contributions, we describe how knowledge deployed within a conceptualization is taken up and re-used to support a different conceptualization. In this way, the development of conceptualizations as well as the change between them can be captured.

The analysis adopted the perspective of problem structuring and its corollaries of story abstraction and concept-fact bindings to interpret the interactive nature of the students' activities. This allows us to trace the introduction of facts and relational knowledge (in support of one or another viewpoint) and then, more importantly, to trace how facts and relational knowledge are taken up subsequently within a different viewpoint resulting in a different structure of the case. By adopting this perspective, we can show what a learner does with the knowledge being shared with them and how the reuse of knowledge relates to the development of the abstract representation of the case.

Through this analysis, we show that the development of an abstract representation of the case conforming to the set learning goals – namely the structuring of the case in relation to domain knowledge – results from an individual appropriating selected knowledge elements and the restructuring of knowledge conveyed in prior contributions.

Iterative development of representations

Joint construction of knowledge is deemed to lie at the heart of collaborative learning and is a complex and multi-dimensional phenomenon (Hausmann, Chi & Roy, 2004). An influential model treats joint construction as the process of conceptualizing the problem within a shared problem space (Rochelle, 1992). Within this problem space a more or less virtual representation of the problem is produced and iteratively improved through successive contributions made by different people. The group converges on a set of shared meanings and shared view of the problem and problem solution by gradually aligning their views. Conceptual change, according to this model, results from the tendency of convergence: in the effort to be communicative and collaborative, peers' cognitive responses are gradually aligned resulting in a shared representation of the problem.

The prime attraction of the shared representation model is the explanation it affords of how conceptualization of the problem develops within the group with each group member contributing conceptually. The group produces a shared representation of the problem, albeit incomplete and flawed, and continues to refine it collectively. The conceptualization evolves as different meanings for concepts and different ways of structuring the problem converge on a common meaning and structure.

Individual representations

However, the development of a shared problem representation is not intrinsic to collaborative knowledge construction, and learners' individual construction of representations is at least equally plausible. For example, Miyake & Shirouzu (2002) argue that whilst collaborative learning from problem solving may seemingly correspond with a 'one voice' interpretation, close inspection of a group dialogue reveals that each person is instead using the others as "stepping stones" to enable them to elaborate a distinctive, individualistic representation of the problem. Peers within a group provide a vital monitoring role for each other, allowing them to develop their own individual representations of the problem (Shirozou, Miyake & Masukawa, 2002). This individualistic approach allows that people can develop conceptualizations through integrating different considerations about the problem features and the relevant underlying concepts; however the conceptualizations remain essentially personal and re-interpretation occurs within the person.

The individualist model of collaborative problem-based learning supports a characteristically different explanation of how conceptual change occurs wherein each individual is seen to generate and modify their own conceptualization of the problem but influenced by others' conceptualizations. One way in which this influence may work is that a person internalizes and adapts the conceptualization offered by one of their peers in the group. Alternatively, the conceptualization is created wholly *de novo* by the individual but likely incorporating elements of the conceptualizations offered by the peers.

Interaction

The individualist model sees interaction within the group as having a fundamentally different role of supporting the learners in developing their own perspectives, engendering, for example, more thorough checking and modifying of individual understandings (e.g. Miyake, 2007). In this view, individuals' construction of their own knowledge is a consequence of being challenged and critiqued (e.g. Glachan & Light, 1982) rather than a result of individuals building on each other's knowledge. In interactive situations learners are seen to readily produce explanations from which they themselves benefit most; this 'self-explanation effect', (Chi et al., 1994) is naturally consistent with the individualist model. More generally, for the individualist model the effect of learners being in a collaborative situation is primarily pragmatic: the situation engenders specific roles (the monitor, the task-doer; the explainer, the listener) that in turn result in re-elaboration of an individual's knowledge, and learners may change their understanding independently on the alternative understandings proposed by someone else (e.g. Schwarz, Biezuner & Neuman, 2000).

Central to research on interaction is the question of the mutual dependency of learners' activities and the level of the content, i.e. knowledge deployed in the groups' cognitive environment. In CSCL research, various concepts have been proposed to capture 'interactivity' in order to articulate claims about the role of interaction in learning (e.g. Fischer & Mandl, 2005). The concept of uptake (Suthers, 2006) is specifically intended to capture not only the pragmatic aspects of an interaction, but also the way in which the content of learners' contributions are related, and hence possibly dependent. Uptake describes how learners move within a shared problem space, each one adding, modifying, elaborating or combining knowledge and understanding of the problem at hand and influencing each other by the continuous refinement of that problem space. However, radical changes in conceptualization during a group learning session may be difficult to capture without also analyzing the point of view or perspective learners currently work within. These perspectives may not only be held and maintained by an individual, but may also be at the heart of changes in conceptualization that often represent the most advanced forms of collaborative learning. By identifying the point of view or perspectives of individual learners we can pinpoint the mechanisms through which such changes are engendered. Such perspectives are visible in the way in which a case or problem is structured, that is, in the way in which facts are

interpreted and related to each other, forming an often hierarchical multi-level abstract structure whose ‘peak’ is the ‘point’ or ‘story’ of the case. The notions of conceptualization and structure, and the processes of structuring and concept-fact binding are thus central in our analysis, and will be introduced now.

Structures, structuring and conceptual change

In general, problems or cases (in this paper we use the terms interchangeably) consist of surface features and structural features. The surface features are the concrete aspects of the problem, both given and implied, and can include agents, objects, actions and events. For example, in the case of medical negligence used in the case study that follows, a surface feature is the low level of patient risk associated with the medical procedure. A problem also has a structure that represents its point or story. The structure is an abstraction that binds the surface aspects of the particular problem to the domain concepts. For example, in the medical case the structure is the causal relationship between the injury suffered by the patient in the operation and the failure to inform the patient of the risk associated with the medical procedure (i.e., if the patient had known about the risks, would they have consented to the operation). The question of the surgeon’s liability turns on this point.

To conceptualize a problem is to develop a representation of the structure of a problem. It requires a person to have knowledge of the domain from which the appropriate concepts can be drawn to produce a correct structure. In the medical negligence case, to produce a correct structure the learners need to already know that the law requires one to show that harm to a patient was *caused* by the actions of a doctor, not simply that the doctor’s actions were deficient.

Method

The first analysis identified the argumentation structure of the dialogue. Utterances were categorized into claims, counter-claims, confirmations and elaborations. This gave us an overview of the overall dynamic of the discussion as well as insight into the individual perspectives adopted by the students.

A second and more important analysis we employed is verbal analysis (Chi, 1997) that focused on the content of the chat messages. This analysis assumes that specific terms represent a learner’s view of the problem, and that the cognitive structural representation of a problem can be uncovered by analyzing the use of relational terms (such as ‘because’, ‘but’, etc.). To capture the relations between contributions, a further analysis, based on the earlier verbal analysis, identified topical relationships between parts of or the complete contribution. So, for example, we checked whether a structural relation evident in a contribution was mentioned in a subsequent one; or whether contributions mentioned the same fact of the case. Based on this analysis, speculations about the relations of earlier utterances on subsequent ones were made, therefore giving us a picture of the interactivity between the learners.

The study

The learning setting

The setting for this study was a purpose built distributed and collaborative virtual learning environment. The students were remote from each other and all their communication was mediated by the system through a chat tool. The chat window can display 30 lines of text at one time, allowing the most recent exchanges to be reviewed. The user interface also contains a menu that allows users to select the content of a display window containing the case text and the background to the core notions of medical negligence. The system recorded all text-communication and the student’s use of the hyperlinks. The study lasted about 20 minutes.

A group of three undergraduate medical students participated in the study. The students were given five minutes to read a short summary of the lecture on medical negligence they had attended. The students were then presented with the description of the case and were asked to explain the judge’s verdict.

The case

The students discussed the ‘Chester case’, a case of negligent liability. The text given to the students read:

“Miss Chester suffered from chronic back pain and consented to neurological surgery. After the surgery, she had extensive motor and sensory impairment. The kind of surgery done, it was accepted, carried some unavoidable risks and a one to two per cent chance of serious complications. It was also accepted that a patient such as Miss Chester should be informed of those risks.

The extent of this warning was disputed. The doctor claimed she had been warned adequately. Miss Chester countered that her fears about being ‘crippled’ were brushed aside and she had been told that the operation was merely a routine procedure. She went on to claim that if she had known about these risks she would not have agreed so easily to surgery without giving it more thought and/or seeking a second opinion.

Additional information: There was evidence of her aversion to surgery – judged unlikely to be so reduced in three days, if she had been adequately informed of the risks.“

Analysis results

Introduction

We will begin the analysis of the dialogue by describing the views taken by each of the three students – Erin, Claire and Susan (not their real names) – because the dynamic between the different views is central to our claim for the reuse of knowledge.

The essential dynamic is simply described as follows: Erin proposes a judgment of the case based on its conceptualization in terms of informed consent. Susan, while initially accepting and contributing to that conceptualization, changes, at the end of the first half of the discussion, radically her view of the case, and correctly applies the concepts and conceptual relations pertinent for negligent liability. Claire contributes more significantly than Susan to Erin's conceptualization, but analyses it critically, and by doing so, furnishes Susan with structural knowledge upon which Susan then builds her new conceptualization.

Erin and Claire

Erin opens the discussion by introducing the concept of informed consent to conceptualize the case (figure 1, line 77). As described earlier, this conceptualization relies on ethical principles (primarily the principle of non-paternalism, i.e. the prohibition to limit a person's autonomy or liberty even if it is thought to be for their own good; the requirement of informed consent derives from this principle) and is complementary to the 'legal view' that the students *should* adopt. She concomitantly judges the surgeon's behavior by appeal to this view.

77	<i>Erin: I think the doctor has failed to achieve appropriate consent</i>	conceptualization (judgment)
78	<i>Claire: the risk was small 1-2%</i>	countering 77
79	<i>Susan: he got consent, just not an informed one</i>	elaborating 77
80	<i>Claire: he must have had the consent</i>	countering 77
81	<i>Claire: informed but not well enough informed</i>	countering 79

Figure 1. Students' initial argumentations within the concept of 'informed consent'

Erin's view is therefore not incorrect, but rather misplaced. Claire attempts to counter the severity of that view arguing that that 'the risk was small' (figure 1, line 78), and that what has been told the patient suffices that the consent obtained is an informed one. It is important to point out that this argument is essentially incorrect because the requirement of informed consent stipulates that the patient needs to be informed independently of any other consideration.

82	<i>Susan: even risk of death from GA is small, but patients have to be told all the same</i>	countering 78, confirming 77
83	<i>Erin: exactly he didn't fulfill the criteria of legal consent. i.e. the patient should understand the risks involved be able to weigh them in a balance and recall them.</i>	confirming 77
84	<i>Claire: but if she had been told the risks and heard that there was only a 1-2% chance that it would go wrong, she might have dismissed it thinking that it probably wouldn't happen to her</i>	countering 82
85	<i>Susan: do you think that the patient should be told of every possible risk?</i>	
86	<i>Erin: but there is evidence of her aversion to surgery</i>	countering 84

Figure 2. Claire's countering of the severity of the lack of information. It introduces the relation between the knowledge of the risk and the patient's decision, within the perspective of informed consent.

Erin continues to maintain that informed consent was not obtained (figure 2, line 83); while Susan confirms the absolute requirement to inform a patient (principle of non-paternalism; line 82, figure 2). Claire counters these absolute stances by reiterating that the risk was so small, that even a fully informed consent wouldn't have changed the patient's decision (figure 2, line 84). We view the introduction of this relationship as an essential stepping-stone towards the switch in perspective and the development of the correct conceptualization (see section on Susan, below). Through this argument, Claire introduces, for the first time in the dialogue, a relation between the knowledge about the risk and the patient's decision, albeit this relation is 'deployed' within the ethical perspective and in order to counter Erin's insistence on the principle of non-

paternalism. Claire’s thinking in relation to the ‘smallness’ of the risk is evident in line 92 (figure 3), giving us proof of Claire’s working within the ethical perspective.

The introduction of this relation (a structural knowledge) is crucial for the discussion, and indeed this relation is a core relation of the structure of the case within a conceptualization centered on negligent liability. Indeed, the plaintiff (Mrs. Chester) will have to convince the court or jury that if she had been fully aware of the risk, she would have not consented to the surgery. Through this, a relation of causality or co-causality between the surgeon’s action and the harm is established.

Susan

It is Susan who reuses the relation between the knowledge of the risk and the patient’s behavior, to propose the novel (and correct) conceptualization of the case based on the concept of negligent liability (figure 3, line 90)

While not expressing the novel conceptualization as a question (the question to be answered in court), Susan’s utterance displays the correct relations between the facts of the case in terms of negligent liability. Susan’s reference to both aspects (knowledge of the risk and patient’s decision) in utterance 90 is evidence that she had generated her new conceptualization on the basis of aspects that had become the main focus of the discussion. Susan’s accomplishment is to isolate the aspects used earlier to support a different view, and insert them into a new structure.

90	<i>Susan: so had she known of the risks, she would have continued to suffer from chronic back pain</i>	(new) conceptualization
91	<i>Erin: that would be up to her</i>	
92	<i>Claire: patients should be told all the risks, but some risks are so small that there’s no point emphasizing them too much otherwise you’ll just end up worrying the patient (probably unnecessarily)</i>	countering 89

Figure 3. Susan’s different (and correct) conceptualization (line 90): the correct relation between the ‘knowledge about the risk’ and the patients’ decision.

Claire

Claire switches between the views established by her peers. She initially embraces Erin’s view, but remains critical of it. Claire’s contributions are mostly counter-claims to Erin’s conceptualization, but occur within the view of informed consent. They are shown here in context.

Claire’s first contribution (line 78, figure 1) counters the judgment entailed by Erin’s use of the concept of informed consent. Claire’s use of the ‘smallness of the risk’ is unclear in 78, but is reiterated and becomes clear in 84 (figure 2). Claire’s position on the case, reiterated and evident also in 92 (figure 3), is the following: she accepts that the surgeon has violated the patient’s right to information, but proposes to include, in an eventual judgment, mitigating circumstances. Specifically, she argues that since the risk is so small, there was no need to acquire informed consent.

It is crucial for our interpretation to understand that Claire, despite her proposal to take into consideration the special and mitigating circumstances, views the case in terms of the concept of informed consent, not the concept of negligent liability. Claire, as Erin, views the case as representing an issue of patients’ rights and professional obligations, not an issue of responsibility and compensation for damage caused.

Later however, Claire internalizes the new conceptualization put forward by Susan in utterance 90. Claire’s contribution in 98 (figure 4) is clear evidence for this internalization: indeed this is the question that will be answered in court. This is in contrast to Erin who continues to maintain her point-of-view (figure 4, line 97) and still contributes to the conceptualization offered by Susan (and later Erin), thus remaining well interactive.

93	<i>Susan: she had a fear of being crippled, had an aversion to surgery, the risk of impairment from the surgery was small</i>	
94	<i>Erin: and then she wouldn’t be complaining now as she would have fully consented to the surgery and understood the risks</i>	
95	<i>Susan: it seems to me that the doctor was trying his best to get her to have the operation</i>	
96	<i>Susan: true</i>	
97	<i>Erin: but at the end of the day it is not the doctors right to decide whether she should have surgery it is her decision</i>	countering 95
98	<i>Claire: do you really think that she would have refused the operation even if she had known the risks?</i>	

Figure 4. The last exchanges. Erin continues to view the case within the ethical perspective; while Erin changes to the perspective of negligent liability producing (in 98) the question that will be answered in court

Reinterpretation of facts

Further evidence of the different perspectives and that a new (and correct) structure has been created by Susan, derives from the role and relevance given to an apparently negligible fact: that the surgeon was aware of the patient's aversion to surgery.

Erin had emphasized that fact to claim, 'it was even more important to fully inform the patient' (figure 2, line 86). The fact, therefore, acquires some relevance in her conceptualization. However, within the new 'negligent liability' conceptualization, this fact is differently explained. The surgeon's withholding of information is seen now as evidence for his caring attitude towards his patient (Susan, figure 4, line 95) rather than evidence for his professional deficiency. It can clearly be both, albeit in a civil court where only the question on the causal relation between an action and harm is discussed, it is not relevant. It is important to point out that it is Susan giving a new role and relevance to this fact: it is her who completes the binding of the facts within the new structure.

Discussion

Two perspectives on the Chester case are evident in the dialogue: one based on ethical principles and one based on the legal concept of negligent liability. Each is a valid way of understanding the case, but the task the students were set demands to apply the legal perspective. Specifically, they need to consider whether the surgeon's withholding of a more extensive warning had a causal or 'causally contributing' effect to the damage sustained by the patient.

The two perspectives do overlap: both reference explicitly the failure to fully inform the patient about the risks of the surgery. It is correct, as Erin states in #77 and continues to maintain throughout the discussion, that the clinician failed to obtain informed consent from Ms Chester and therefore violated her right to be informed. However, the ethical perspective is not sufficient to identify the causal relation that is at the core of cases of negligent liability.

Our interest in the discussion lies in the conceptualizations that are constructed within the two perspectives, which enable us to show how knowledge constructed and made relevant within one perspective is taken up within a different perspective. We have identified these conceptualizations by analyzing the interpretations the students give to aspects, and, more specifically, the role and relevance they assign to the aspects. Erin proposes the first conceptualization of the case, embedded within the ethical perspective. Within it, the absence of adequate warning is uniquely sufficient to propose a 'judgment' on the surgeon's behavior because it violates the patient's right to information and, more generally, it violates the ethical principles of non-paternalism and autonomy. Most other aspects of the case have no relevance in this conceptualization. Later, Erin gives some relevance to the aspect 'aversion to surgery' by pointing out that, given Ms Chester's aversion, it is all the more important to inform her. Erin's interpretation of this fact is a further indication for her framing of the case within an ethical perspective; in a sense, she is clearly focused on providing a judgment of the surgeon's behavior in relation to ethical principles ("he failed (...)"). His knowledge about the patient's aversion is, for Erin, a further indication for a general ethical failure.

Claire accepts Erin's perspective on the case, but analyzes it critically by checking whether a specific concept – the concept of informed consent – of the ethical perspective is applicable to the case. She points out that since the risk was so small, informed consent was not needed. She mentions the 'smallness' of the risk, the aspect that according to her trumps Ms Chester's right to be informed, as a fact to counter Erin's quite severe indictment of the surgeon. The discussion revolves, as a consequence, on the significance of the 'size' of the risk, which is made relevant and becomes the focus for all students' contributions. In this way, they remain interactive and collaborative (Trognon, 1993; Barron, 2000), and indeed Claire, in #84, *works on* from this fact to propose a relationship between the size of the risk and the patient's decision. It is important to point out that the relation between the 'size of the risk' and the patient's decision is deployed into the group's cognitive environment to counter Erin's 'evidence' (the absence of adequate warning); it is an attempt to lessen the impact of Erin's 'evidence' on the judgment. The relation is hence deployed within the ethical perspective, which by itself is, at this stage of the discussion, not questioned.

But it is this relation between the 'size' of the risk and the patient's decision that provides the basis for Susan's re-conceptualization of the case. As we have shown, Susan constructs the different conceptualization within the perspective of negligent liability. She isolates the relation between the extent of the warning and the patient's decision from Claire's earlier relation between the smallness of the risk and the extent of the warning and the patient's decision, and so correctly identifies the core question of the case: whether being fully informed about the risk would have made the patient decide to not to undergo the surgery. She then also interprets the other aspects on the basis of this structure, as indeed the new conceptualization assigns different roles and relevance to the aspects. The 'smallness' of the risk is irrelevant within this structure, except in relation to the patients' decision (in the ethical perspective it was discussed as being central); another fact, the surgeon's knowledge about the patient's aversion to surgery, while not being crucial, is given an interpretation that is opposite to the one Erin gave it within the ethical view: rather than strengthening the view of the surgeon as

having failed, Susan points out that he may have withheld some information on the basis of his professional opinion that surgery was the best option available to the patient; he is, in this sense, a caring doctor.

Susan's novel conceptualization is only gradually taken up and developed. However, while Claire at the end accepts the new conceptualization, and indeed poses the question that will have to be answered in court (figure 4, line 98), Erin continues to persist on the ethical view.

By tracing the introduction and subsequent take-up of surface and structural knowledge (the causal relation) we show that the construction of a new (and, given the task demand, correct) conceptualization builds upon knowledge elements introduced previously (cf. Trognon, 1993). We argue that Susan's new conceptualization selectively exploits structural and surface elements of the earlier conceptualization, and that Susan remains interactive when creating the new conceptualization.

To what degree then is Susan's construction of the conceptualization a co-construction? Theberge-Rafal (1996) describes co-construction as the phenomenon where utterances by different speakers represent a complete idea or where a contribution extends a previous contribution. Co-construction is a mark for interactivity, and, in Barron's (2000) analysis, of successful coordination. Co-construction may however also indicate jointly produced novel ideas (e.g. McGregor & Chi (2002)); whether they are produced as a result of collaboration or are joint articulation of pre-existing ideas, is clearly difficult to establish. However, if co-construction were defined as the joint construction of a complete idea, then the dialogue between Susan, Claire and Erin would not count as an example of it. It is more the case that the learners take up one another's contributions, modifying the content and developing new ideas from them (Suthers, 2006).

Roschelle's (1992) notion of joint construction is also relevant for the interpretation of the students' discussion. He defines joint construction and, by extension, collaborative learning as the mechanism by which a new conceptualization is generated on the basis of a peers' partial conceptualization. The phenomenon described in this paper calls for a refinement of this definition. Roschelle's definition hinges on the notion of partial conceptualization. On the surface, Claire's conceptualization cannot be called a partial conceptualization of Susan's because it is embedded within a different perspective. However, we might only decide that Claire's conceptualization is indeed partial if we know what interpretation Susan gave it. Susan interpreted Claire's contribution in terms of negligent liability, and completed this partial structure. Moreover Claire's contribution becomes a partial conceptualization because Susan interpreted it in this way. Which is the mechanism within which collaborative learning occurs: a student gives meaning to information written in a text or knowledge introduced into the group's cognitive environment, which may then be taken up and given a different meaning by her peers.

It is important to discuss also the possible impact of the virtual environment of the discussion. It is likely that the availability of prior deliberations in the chat window (it displays about 8 messages from all learners) has made it possible to revisit those prior deliberations, and facilitated their reuse in the construction of the novel conceptualization. Face-to-face conversations are typically more strongly constrained by adjacency with the strict ordering of utterances organized – in addition to content relevance – by non-verbal cues and rules of turn-taking. Though 'far' references are not uncommon, they represent a significant effort on the part of the speaker to close a current thread, while needing to justify the opening of a new one. In a computer-supported learning environment many conversational rules do not apply and typed messages are more persistently present than spoken ones, allowing learners re-visits of prior contributions and interpretations less constrained by the currently agreed perspective.

Conclusions

As recognized by many in CSCL, it is essential that the basic notion of interactivity – namely that an action is at least partially influenced by a prior action – remains the focus of collaborative learning research. More specifically, it is important to adopt frameworks that enable us to capture the determination of one peer's action in relation to another peer's action. Individualistic accounts allow, in this respect, quite under-determined actions: *what* knowledge is constructed when being, for example, criticized, is under-determined and indeed depends very strongly on an individual's own knowledge. The study of collaborative learning should therefore include a strong focus on what knowledge is constructed and how that new knowledge is specifically related to knowledge in the group's environment.

We have adopted a framework for studying collaboration that is focused on this analysis, as well as allowing an influence of individualistic processing. Our analysis illustrates how collaboration is sustained through a joint focus on shared information and knowledge (Barron, 2000, 2003) that provides opportunities for interaction, but may also, as a consequence, restrict what new knowledge is constructed. The group's achievement becomes, within this view, quite remarkable. Despite an early focus on a fact that is, at the end, not relevant within the sought conceptualization and the emphasis on the (ethical) perspective that all 3 students work within, a novel conceptualization is constructed interactively. Our analytical focus on knowledge elements and especially the introduction and uptake of structural knowledge, leads to the conclusion that this new

conceptualization results from an individualistic appropriation of elements (cf. Schwartz, 1995), rather than progressive shared construction (Rochelle, 1992).

In problem-based learning, where microanalyses are *de rigueur*, case studies have a special significance. However, a single case study cannot be the basis for a general claim – case studies are useful for illustrating particular processes or mechanisms that may then be the subject of further inquiry, both theoretical and experimental. This paper is not intended to provide evidence that individualistic appropriation is a key mechanism for collaborative learning; rather, it is intended to demonstrate the viable application of this concept to a collaborative learning situation; and also to demonstrate the difficulties of applying a notion of interaction to a real-world learning session where that notion does not take into account the positioning resulting from different viewpoints of the learners. But more importantly, the case study illustrates how individual viewpoints can be uncovered through a verbal analysis and by carefully constructing the alternative structures each viewpoint entails (fact interpretations and relations between facts). As such, our analysis may be seen as contributing to a much needed and increasingly sought framework for the analysis of conceptual change through collaboration.

References

- Barron, B. (2000). Achieving coordination in collaborative problem-solving groups, *Journal of the Learning Sciences*, 9(4), 403-436.
- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12 (3), 307 – 359.
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, 6(3), 271–315.
- Chi, M. T. H., de Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding, *Cognitive Science*, 18, 439-477.
- Doise, W., Mugny, G. & Perret-Clermont, A.-N. (1975). Social Interaction and the Development of Logical Operations, *European Journal of Social Psychology*, 9, 245-247.
- Fischer, F., & Mandl, H. (2005). Knowledge Convergence in Computer-Supported Collaborative Learning: The Role of External Representation Tools *Journal of the Learning Sciences*, 14(3), 405-441.
- Gick, M. and Holyoak K. (1980). "Analogical Problem Solving." *Cognitive Psychology* n. 12, pg. 306-55.
- Glachan, M. & Light, P. (1992). Peer Interaction and Learning: can two Wrongs make a Right? In G. Butterworth & P. Light (Eds.), *Social Cognition: Studies in the Development of Understanding*, (pp. 238-262). Chicago: University of Chicago Press.
- Hausmann, R. G. M., Chi M. T. H., and Roy M. (2004). Learning From Collaborative Problem Solving: An Analysis of Three Hypothesized Mechanisms. *Proceedings of the 2004 Cognitive Science Conference*. Mahweh: NJ: LEA.
- Jeong, H. & Chi, M. (2007). Knowledge Convergence and Collaborative Learning, *Instructional Science*, 35:287–315
- McGregor, M. U. and Chi, M. T. H. (2002). Collaborative Interactions: the Process of Joint Production and Individual Reuse of Novel Ideas, *Proceedings of the 2004 Cognitive Science Conference*. Mahweh: NJ: Lawrence Erlbaum Associates, 2004.
- Miyake, N. (2007). Conceptual Change through Collaboration, in Vosniadou, S. (Ed.) *International Handbook of Conceptual Change*, pp. 453-478, Routledge.
- Miyake, N. and Shirouzu H. (2002). Learning by Collaborating Revisited: Individualistic Vs. Convergent Understanding. *Proceedings of the 2002 Cognitive Science Conference*. Hillsdale: NJ: Lawrence Erlbaum Associates.
- Roschelle, J. (1992). Learning by Collaborating: Convergent Conceptual Change, *The Journal of the Learning Sciences*, 2 (3), pp. 235-76.
- Schwartz, D. L. (1995). "The Emergence of Abstract Representations in Dyad Problem Solving" *The Journal of the Learning Sciences*, vol. 4, no. 3, pg. 321-54.
- Shirozou, H. (2010). Collaboration as constructive interaction and the jigsaw method as its enhancer, *Workshop Proceedings of the 9th International Conference on the Design of Cooperative Systems*, *International Reports on Socio-Informatics*, 7(1). International Institute for Socio-Informatics.
- Shirouzu, H., N. Miyake, and H. Masukawa (2002) "Cognitively Active Externalisations for Situated Reflection." *Cognitive Science*, vol. 26, n. 4, pg. 469-501.
- Suthers, D. D. (2006). Technology affordances for intersubjective meaning-making: A research agenda for CSSL. *International Journal of Computers Supported Collaborative Learning*, 1(3), 315-337.
- Trognon, A. (1993). How does the process of interaction work when two interlocutors try to resolve a logical problem? *Cognition and Instruction*, 11(3&4), 325-345.

Acknowledgement

We acknowledge the support of the Economic and Social Research Council, award no. L328 25 3013.

Experiences of a Newbie Helper in a Free Open Online Mathematics Help Forum Community

Carla van de Sande, Arizona State University, Tempe, AZ, carla.vandesande@asu.edu

Abstract: Free, open, online help forums are open to the public and allow students to anonymously seek homework help from volunteers who have the time, willingness, and experience to respond. These forums offer affordable, accessible, and efficient help as a social, public endeavor. Some forums exhibit a strong sense of virtual community, especially amongst well-established helpers who are core participants. To investigate how newcomers may enter into such a community, five helpers were recruited to participate for eight consecutive weeks in an existing popular forum for mathematics homework help covering arithmetic through advanced mathematics. We modified a virtual community framework to describe the activity of the newcomer helper who made the most progress in moving from peripheral to fuller participation in terms of membership, influence, and immersion. This process involved cultivating a sense of belonging and building supportive relationships, contributing meaningfully and creating a voice, and demonstrating dedication to shared goals.

Introduction

Free, open, online, homework help forums are found on public websites and allow students to anonymously post problem-specific questions from assignments that are then visible to others. These forums are open in the sense that, unlike other asynchronous communication tools (such as course forums or discussion boards), access for students is not restricted to any particular course or institution. Also, instead of hosting discussions based on the content from a particular course, the forums cover broad school subject areas (such as mathematics, science, and business) at a range of course levels (from elementary to graduate). These forums are a help-seeking resource that is currently available to any student who has Internet access.

Students from around the world access these forums when they are in need of help completing assignments or understanding course material outside of the classroom, and this is particularly true for school subject areas such as mathematics that are homework intensive and require students to construct solutions to exercises. It is also probably no coincidence that the mathematics forums receive so much traffic given the large number of students who suffer from “math anxiety” and who approach their studies and math courses with apprehension, if not downright dread. (See Hembree, 1990 for a meta-analysis of the extensive body of research on math anxiety). Thus, many sites (e.g., www.mathhelpforum.com and www.freemathhelp.com) that offer help in arithmetic through higher mathematics have tens of thousands of members, thousands of whom are regularly active. The net result is a rich set of archived threads that stem from student mathematics homework questions and contain authentic help-seeking and help-giving discourse. These are some of the reasons why a mathematics homework help forum was chosen as the focus of this study, although the research presented here is pertinent in any school subject area covered by help forums.

The forums belong to a genre of technology-assisted education called ‘networked learning’ (Goodyear et al., 2005) that emphasizes the social dimension afforded by information and communications technology. The extent of the ‘network’ depends, in part, on the participation structure that a given help forum adopts (Martinovic, 2005; van de Sande & Leinhardt, 2007). Some forums effectively bound the network by restricting the set of volunteer helpers to those who are qualified according to some criteria (such as content knowledge or pedagogical approach). Other forums favor a broader, self-selecting network (which we refer to as Spontaneous Online Help, or SOH sites) by allowing any member to take up and respond to a student query and contribute to an ongoing threaded discussion, or thread. What this means is that SOH “tutors” can be anyone with any level of expertise and favoring any pedagogical approach from anywhere. Furthermore, and paramount to the notion of community, these SOH helpers can interact with one another in the threads as they collectively help students with their assignments.

The helpers (who are generally the core participants) in SOH mathematics forums may exhibit a strong sense of virtual community if they identify with fellow members, assume responsibility for participation, negotiate features and practices, and appear comfortable exchanging ideas (van de Sande & Leinhardt, 2007). In such sites, it is common practice for helpers to refer to one another by name, joke and kid around with one another, introduce alternative perspectives or ways of looking at the problem, engage in “math talk,” and perform peer review by critiquing or correcting mistakes and errors in others’ contributions (van de Sande, 2008). Broadly speaking, the helpers in forums with these dynamics might be considered as members of a community of practice, as conceptualized by Wenger (1998) and Lave and Wenger (2002).

The objective of this study is to explore the enculturation of newcomer helpers into an SOH mathematics help forum that already has a tightly knit group of well-established helpers. In other words, we are exploring how newcomers learn to become accepted into the community, or “members of the gang.” Adopting the lingo of contemporary youth particularly with regard to participation in a given Internet activity, we refer to these newcomer helpers as “newbies.” Using a modified framework that conceptualizes sense of virtual community, the major question addressed is: What are the characteristics of a newcomer helper who is on an “inbound trajectory” from peripheral to full participation (Wenger, 1998, 100-101) as a helper in such a forum?

Theoretical Framework

A community is chiefly characterized by relational interactions or social ties that draw people together (Heller, 1989). The connections between members of a community need not be based on physical proximity (e.g., neighborhoods or classrooms), but can be forged between people anywhere who share interests, hobbies, or ideas. The Internet, in particular, supports such relational communities by mitigating the distance between people who wish to connect and interact with one another.

The notion of a virtual community is a diverse, dynamic, and emergent construct, building largely on our understanding of physical community. According to Fernback and Thompson (1995), a virtual community consists of a set of social relationships created in cyberspace through repeated contact within some boundary. Balasubramanian and Mahajan (2001) specify a set of virtual community requisites: (1) an aggregation of people, (2) rational members, (3) interaction in cyberspace without physical collocation, (4) a process of social exchange, and (5) a shared objective, property/identity, or interest. Preece (1999) emphasizes the relational aspects of a virtual community, such as the presence of mutual support and shared emotional connections. Despite their different emphases, all of these notions of virtual community involve relationships of various kinds (e.g., help services, interest groups, etc.) and degrees (ranging from casual to formal) that exist in Cyberspace.

Like physical communities, virtual communities can be more or less tightly knit. Community researchers refer to the affective bonds that characterize closely woven communities as “sense of community,” and recognize that the presence of sense of community increases member satisfaction, commitment, and involvement. Based on McMillan & Chavaz’s (1986) theoretically based and empirically supported descriptive sense of community framework, Koh & Kim (2003-4) propose three dimensions that characterize a sense of virtual community: (1) membership – people experience feelings of belonging to the community, (2) influence – people feel that they have a voice in the community, and (3) immersion – people are dedicated to spending time within the community. Meaningful antecedents include leaders’ enthusiasm (commitment to maintaining and sustaining community) and enjoyability (pleasurable engagement in community activities). In this paper, we apply a modified version of this framework to the activity of a newbie helper on a particular online mathematics help forum to characterize what it means to move from the periphery to fuller participation in a virtual community.

Methods

We chose a forum with a strong sense of virtual community (van de Sande & Leinhardt, 2007) that already has several well-established helpers as core participants and tolerates a variety of pedagogical tactics (ranging from providing solutions to scaffolding moves). As a first step, we elected to use qualitative methods because the exploratory research involves an emergent phenomenon (e.g., forum helper enculturation), and is geared toward gaining in-depth information that is difficult to quantify (e.g., group membership) (Strauss & Corbin, 1990).

Five recruited helpers were asked to participate in an online mathematics SOH forum for at least 10 hours weekly (scheduling flexible) for eight consecutive weeks and were compensated \$10/hour. These (admittedly somewhat arbitrary) conditions were intended to support a level of commitment that would be sufficient for participants to become acquainted with the logistics of forum participation and potentially experience a sense of virtual community.

Site Description

Started in 2002 by an enterprising high school sophomore, www.freemathhelp.com (FMH) has attracted over 17,000 members to date, who have contributed upwards of 140,000 posts to more than 34,000 threads. The math help message board is split into individual forums covering mathematics from the elementary level (arithmetic) up through university (post-calculus), and also includes forums for discussing administrative issues (such as any problems encountered or suggestions for improvement) and “math odds and ends” (puzzles, clever math jokes, trivia, etc.). FMH is an SOH site, so that any registered member can act as a helper. The FMH helpers who volunteer information about themselves as part of their member profile are self-reportedly (retired) educators, professionals, and (advanced) students, who have a passion for wanting to share their expertise and help others. In contrast to some other mathematics help forums, there is no single community advocated pedagogical approach prescribing what it means to “help” and how this should be accomplished, although the dominant

perspective is that it is better practice to scaffold students by asking them questions or giving hints rather than to provide answers or worked solutions. (See van de Sande 2010, 2011 for discussion of a mathematics homework help forum that favors one particular theory of learning.) Instead, on FMH, participating helpers can choose whether and how to respond to any given student query or ongoing thread – and neither these decisions nor the quality of their contributions affect a helper's status on the forum. Instead of having a reputation system (in which status on the forum depends on others' ratings), members (whether they ask questions, answer them, or both) achieve status on the forum according to the number of distinct threads to which they have contributed: new (0-49), junior (50-249), full (250-999), senior (1000-2499), elite (more than 2500). The established helpers on this site convey a strong sense of virtual community (van de Sande & Leinhardt, 2007).

Participants

The five participants in the larger study all had considerable mathematical expertise and experience helping (undergraduate) students solve exercises face-to-face, but were newbies in online help forum participation. Participants were recruited through advertisement, and selected on the basis of enthusiasm, commitment to the study, and life experience. In particular, an effort was made to recruit young adults who were currently students themselves, as well as adults with career experience. Two of the participants were upperclassmen undergraduate engineering majors (young adults), two were graduate engineering master's students close to completion of their degrees (young adults), and one was a former instructor (adult) with an engineering degree who had recently (6 months previously) immigrated to the United States from Mexico. This last participant was a newbie to the forum in two additional ways: language (English was his second language) and culture (his experiences as a student and instructor took place in Mexico). At the time of the study, the two undergraduate students were also employed part-time by the engineering help center on the campus of a large university in the southwest.

Data

Four sources of data were collected: archived forum threads or logs, interviews, journal entries, and private messages. The primary data source consisted of the logs of participants' threads during their eight contiguous weeks as helpers. In addition, each participant was interviewed twice, once prior to participation, and, again following the eight weeks of participation. These interviews were recorded and transcribed, and addressed content knowledge, experience as helpers/tutors, pedagogical views, comparisons of face-to-face vs. online help, and perceptions of the activity, including suggestions for improvement. To further coordinate the content of forum threads with participants' reflections and intentions, participants were asked to keep journals during the study recording their thoughts, scratch work, and any significant events or encounters. Finally, some of the participants engaged in private messaging with students, which is a forum option. Although not all of these messages were kept (due to lack of foresight on the researchers' part), those that were provide a source of behind the scenes data that is normally not accessible to forum observers.

Results

In this paper, we focus our attention on JuicyBurger who, of all the study participants, appeared to make the most progress in moving from peripheral to fuller participation. Although he did not contribute to the largest number of threads nor have the largest number of posts, his activity over the eight weeks was consistent with sharing in a sense of virtual community according to the three dimensions in the Koh & Kim (2003-4) framework: membership, influence, and immersion. Figure 1 contains our hypothesized schematic depiction of how these dimensions were related to characteristics of Juicy Burger's activity on the forum. We suggest that conforming to the actions of others and affirming the contributions of others are markers of membership, that challenging others demonstrates perception of influence in the community and, conversely, that being either challenged or affirmed by others demonstrates that the community is seeking to influence the member, and, finally, that when a member indicates that s/he is actively participating in the joint effort of other members and is also engaging in extracurricular activities that this signals immersion in the virtual community.

Membership

Members in a virtual community feel connected with other members and enjoy a feeling of belonging, somewhat akin to being a part of a family or a group of close friends. This is true of the established FMH helpers, and evident in the appearance of their contributions, as well as in the ways in which they joke around, tease one another, and compliment one another on clever or novel ways of approaching a problem (van de Sande & Leinhardt, 2007). Although the predominant (and perhaps sole) form of interaction amongst most FMH helpers is on-line, there is nevertheless a strong sense of camaraderie and respect for other subject enthusiasts who are seriously engaged in this grassroots effort to band together and help students.

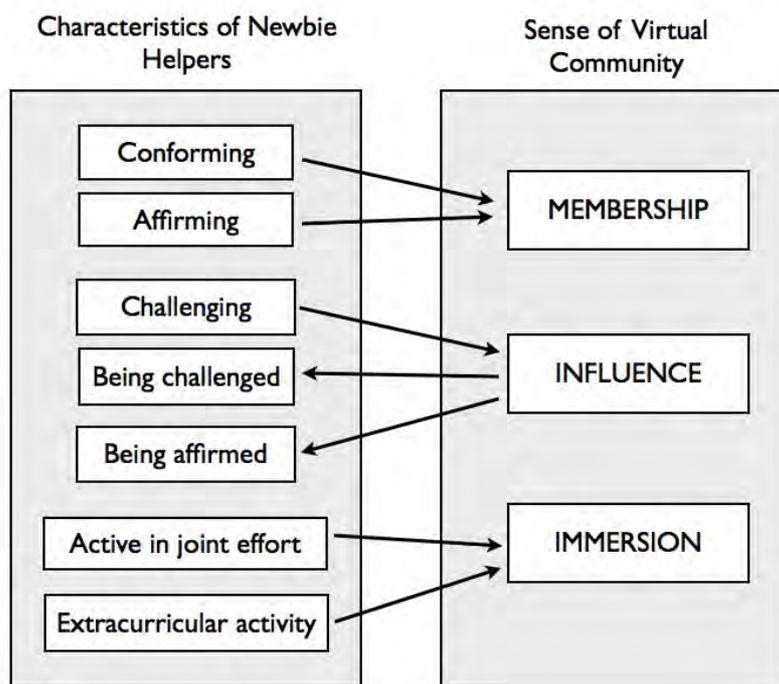


Figure 1. Schematic of relationship between characteristics of newbie helper activity and sense of virtual community.

In his forum participation, JuicyBurger consistently communicated a desire to conform or blend in with other helpers, both in the way he presented his contributions and in the way he interacted with others. First, during his first week on the forum, he adopted the preferred method of posting mathematical expressions. Using horizontal notation can make it cumbersome to communicate mathematical ideas, and many of the established helpers are proficient at using LaTeX (a typesetting system designed for the production of technical and scientific documentation) to format their contributions. (Students, on the other hand, appear to favor horizontal notation, perhaps due to their more peripheral participation in the community.) During his first week on the forum, JuicyBurger asked for advice on how to embed math in his posts, took the advice of the senior member who responded, and subsequently used LaTeX to create well-formatted contributions, some of which included quite complicated expressions. In the post-interview, JuicyBurger explained how he was able to quickly “learn” LaTeX by finding and copying (cutting and pasting) examples from experienced helpers’ threads and then revising the code for his purpose.

In addition to making his posts look like those of established helpers, JuicyBurger was friendly and congenial, bantering with other helpers, as well as recognizing and affirming others’ contributions. Table 1 contains excerpts from JuicyBurger’s logs in which he demonstrated solidarity with other helpers by voicing agreement, paying compliments, and supporting their efforts. Notice that in all of these examples, JuicyBurger personalized the interaction by explicitly naming the helper whose contribution he was affirming. By conforming to the habits of established helpers and affirming others’ contributions, JuicyBurger signaled to the established helpers that he wished to share membership in their virtual community.

Table 1: Excerpts in which JuicyBurger affirmed and supported contributions made by established helpers.

Context	Affirmation
JuicyBurger and mmm4444bot (Elite Member) were collectively working with student who was trying to graph a parabola in an online system. (posts 1-17)	‘I am with mmm4444bot on this one...’ (post 18 in thread)
BigGlenntheHeavy (Full Member) pointed out to student who is taking limit of function that the	‘BigGlenntheHeavy is correct... You only need to divide the coefficients of the highest order

lower order terms ‘become superfluous.’ (post 2 in thread)	term...’(post 3 in thread)
JuicyBurger informed student who was seeking ‘an easier method’ for finding arc length, ‘I’m sorry to inform you that this is the only way to do it!’ and suggested that ‘the only easier way would be to use your calculator to integrate for you.’ (post 2 in thread) Galactus (Elite Member) then sketched a method in which the integral is framed in terms of the independent variable instead, with the result that ‘the integral is easier.’ (post 4 in thread)	‘Ah hah! That does make it much simpler. Thanks galactus.’ (post 5 in thread)
mmm4444bot provided student with rules for logarithmic sums and differences (post 2 in thread)	‘Now use the rule that mmm4444bot posted for adding logs and you should find your answer.’ (post3 in thread)

Influence

Members of a virtual community feel influential if their contributions are taken up, reviewed by, and responded to by others. In terms of taking up ideas, established FMH helpers often work collectively with a student by continuing a dialogue established by another helper, adding to one another’s ideas, and pointing out alternative solutions (van de Sande, 2008). Because of the nature of the forum (as a help service), reviews often involve challenging, and sometimes correcting, others’ posts or self-correcting, which has been referred to as a “Wikipedia-like” quality (van de Sande & Leinhardt, 2007) because it represents a collaborative constructive activity with a quest for accuracy. Accordingly, established FMH helpers may have occasional spats that involve challenging the accuracy of one another’s mathematics, disagreeing with mathematical approaches (e.g., whether a certain method is necessary or elegant), and criticizing pedagogical tactics (such as providing the student with a worked solution). Positive mutual influence on others in the community is demonstrated when criticism and contributions are responded to in an affirmative manner, worked through, and resolved.

Having Contributions Affirmed

JuicyBurger’s contributions did not go unrecognized by the established helpers. Table 2 contains excerpts from JuicyBurger’s logs in which other helpers affirmed his contributions by crediting him by name for correct contributions, taking up and adding onto his helper moves, and supporting his mathematics. The play on words that mmm4444bot (Elite Member) used when referring to one of JuicyBurger’s contributions (row two) is further evidence of how other members of the community facilitated his shift to fuller participation.

Table 2. Excerpts in which JuicyBurger’s contributions were affirmed and supported by established helpers.

Context	Affirmed Contribution
JuicyBurger was working with a student to identify key features in order to graph a parabola. The student questioned his help, and mmm4444bot picked up the thread. (posts 1-4 in thread)	mmm4444bot (Elite Member): ‘JuicyBurger is correct. There is symmetry about the vertical line $x = -1$.’ (post 5 in thread)
JuicyBurger provided a hint to student that, in order to find the center and radius of a given circle, the formula should be converted to the ‘standard equation’ $(x-h)^2 + (y-k)^2 = r^2$ (post 3 in thread)	mmm4444bot (Elite Member): ‘I’ll contribute a little free sauce, to the juicy burger below. [quoted hint given by JuicyBurger] by completing the square in x and y. ’ (post 4 in thread) [the bold text was added onto JuicyBurger’s contribution]
JuicyBurger was helping student but questioned the accuracy of his own response: ‘Now you have...[derived expression] What is this equal to? Hmmm this doesn’t seem right.’ (post 2 in thread)	galactus (Elite Member): ‘You’re [JuicyBurger] on the right track. This is basic relativity. ...’ (post 3 in thread)

Having Contributions Challenged

Of course, not all of JuicyBurger’s contributions deserved, and therefore received, affirmation; his contributions were also reviewed and challenged when they contained errors. However, as reported in his interview, JuicyBurger generally took his time framing his contributions (for example, by working through solutions on scratch paper before publishing hints and suggestions), and, perhaps for this reason, did not receive many challenges. One of the earlier challenges to JuicyBurger’s contributions came in his first week on the forum

from mmm4444bot (Elite Member) when he incorrectly took the square root of an expression: ‘Juicy Burger forgot that the square root of $(x+6)^2$ is not $x+6$. It’s the absolute value of $x+6$.’ JuicyBurger responded in a manner and with a tone consistent with forum practice; he (sheepishly) acknowledged his mistake (‘Yes I forgot the +/- [confused emoticon]’) and edited his earlier post to correct the error.

Challenging Others’ Contributions

In addition to having his own ideas taken up and challenged, JuicyBurger made efforts to create a voice on the forum by challenging others’ contributions, even those of high status members. As is generally FMH practice, he was usually polite (by forum standards which appear to have a higher tolerance for sarcasm and impertinence than face-to-face communications) when he questioned others’ contributions. For instance, when galactus (Elite Member) incorrectly parameterized two planes, JuicyBurger simply provided the correct solution for the student (who had also questioned galactus’ response), noting the nature and location of the error: ‘This should give you: $z = t - 1$ (not $z = -t + 1$) [which is how galactus had erred].’ Galactus responded with an apology and an edit of his previous response: ‘Sorry, that was a typo. I fixed my foolish error. Thanks. Goes to show you, one little negative sign [sic] misplaced and it’s all kaput. [smiling emoticon]’ This exchange occurred in the first week following JuicyBurger’s forum debut, and shows how members can influence one another through constructive critique as they work collectively to provide correct mathematical help for the students.

JuicyBurger also confronted an established member whom he felt interfered with a dialogue he was trying to establish with a student. The student had requested help finding the area between two curves, and JuicyBurger suggested that the student start by finding the points at which the curves intersected. Fifteen minutes later, BigGlenntheHeavy (Full Member) essentially provided the solution to the problem by posting the resulting integrals (which contained the intersection points) together with the final numerical answer, $937/6$. JuicyBurger responded within minutes by chastising BigGlenntheHeavy, as shown in Figure 2.

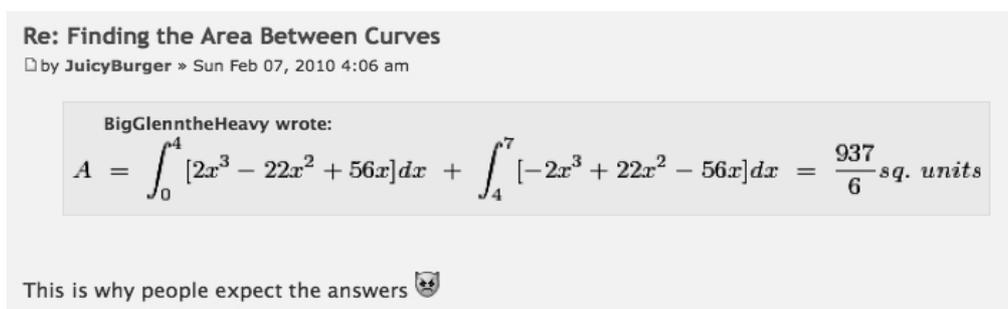


Figure 2. JuicyBurger challenges established FMH helper.

In reply, BigGlenntheHeavy defended his actions, perhaps under the impression that JuicyBurger was referring to his giving away the final numerical answer: ‘JuicyBurger: It is setting up the problem that requires skill, as the answer then is academic.’ Later in the exchange, mmm4444bot (Elite Member) championed JuicyBurger: ‘I think that it is obvious that Juicy Burger was not thinking “expect the answers”. If you claim that setting up the integrals requires the greater skill, big guy, then why aren’t you helping to guide the original poster through *that* process versus spoon-feeding’ [italics in the original]. This mentoring move shows how JuicyBurger’s voice was validated from an influential member of the community. In the interview, JuicyBurger spoke of this incident, reporting that it contributed to his confidence amongst the ‘old guys who control it [the forum].’

Members who share a sense of virtual community exercise influence within the community. In a help forum, this plays out in challenging others’ contributions to order to have an influence on the content or material that the students are receiving. Conversely, if a member is influential in the community, then other members pay attention to her/his contributions so that they do not go unnoticed or ignored. Other members may indicate that a member is influential by heeding her/his contributions and responding with either challenges or affirmation.

Immersion

Members are immersed in a virtual community if they are dedicated (or even perhaps addicted) to the activity and spend large amounts of time engaged in interaction. Although we cannot determine the amount of time per se that established FMH helpers spend crafting and publishing responses, their activity level as measured by number of contributions demonstrates behavior along these lines that is consistent with immersion in the community. It is worth noting that JuicyBurger, with his 96 posts in 48 threads over the course of the study, ranked just under these established FMH helpers during that time period.

However, there are other ways, aside from time spent participating and number of contributions, which signal immersion in a virtual community. These may include demonstrating concern by taking seriously the joint efforts of community members, and signaling interest by participating in extra-curricular activities. In other words, the interest with which a member participates may also be used as part of the index for immersion.

Joint effort

Helpers in an online forum elect how many times they participate in a given thread. Having contributed to a thread, a helper may choose to return to that thread and contribute further, or may choose not to return. In FMH, well-established helpers generally return to threads multiple times, especially if they are involved in an ongoing dialogue with the student or others. This activity demonstrates that helpers have a sense that helping does not consist of a single contribution but rather a process in which they are actively involved.

JuicyBurger appeared to share this view of helping. Instead of participating just one single time in a given thread (as was more the case with other participants in our study), JuicyBurger returned to 46% of the threads in which he was helping a student. In the interview he reported that he set the forum notification system to inform him when there was further activity in one of the threads in which he was participating so that he might consider whether he could (or should) contribute more. This behavior demonstrates a sense of concern, and perhaps responsibility, for the outcome of his contributions to the community that mimicked the activity patterns of well-established helpers.

Along these same lines, JuicyBurger also allowed students to private message him (an option for forum members), and responded when asked specifically to help them. For instance, one student sent the following private message: 'sorry, my first part of the question was cut-off, I had asked you to help me, cause the other night you did help me, and someone else is trying to help, but I'm so lost. [smile emoticon]' We see here that JuicyBurger identified with the community effort through responding outside of the forum archives.

Extracurricular activity

Aside from the quantity, consistency, and thoughtfulness of the help they provide, the immersion of FMH helpers in this community is evident from their lively participation in extra-curricular activities, such as contributing ideas to the administrative forum and posing and solving math puzzles in the 'math odds and ends' forum. Not only do the established members care about sustaining and improving the quality of FMH, they also appear to greatly enjoy matching wits with other subject enthusiasts and tackling challenging problems.

JuicyBurger participated in three threads in the 'math odds and ends' forum during the study. Although he did not initiate any puzzles and riddles, he challenged others' contributions, contributed his ideas for solutions, and spent time engaging in deep thought. The most notable example of this was in a 'math odds and ends' thread involving finding the volume of a drilled sphere. JuicyBurger first challenged the framing of the puzzle ('Actually, you do need to know more information. You need to at least know the radius of the sphere.'), but then retracted his challenge just minutes later ('Actually, I think I may be mistaken (=D) I'm still working it out...'). Although JuicyBurger continued working on this puzzle for the next few days, he did not post further in the thread; another forum member, however, did contribute a solution. A week after JuicyBurger had last posted, BigGlentheHeavy (Full Member) published a calculus-based solution to the puzzle, directing his remarks specifically to JuicyBurger: 'Juicy Burger, in regards to mmm4444bot.'s [sic] poser, if we observe the great circle of a sphere...[solution to puzzle] Note: The volume is independent of R [the radius of the sphere].' The genuineness of the thought that JuicyBurger put into this puzzle was evident in his reply: 'BigGlenn: Yes I realized this after I posted the statement about it being a large sphere with a large cylinder cut out, or a small sphere with a small radius cylinder. I thought about it a bit more (tempted to google it the entire time) and decided that the question had to make sense, otherwise it would not be riddle. So if the question made sense, it must not depend on the radius, which was when I came up with 36π . I haven't posted merely because others have already answered it in the in between time whilst I was thinking about it...' We see here that JuicyBurger forewent the easier path (Internet search) in favor of thinking through the problem himself and immersing himself in the type of activity that other forum helpers enjoy. In the interview, JuicyBurger described this problem as 'cool,' noting that he had learned some mathematics from this particular interaction.

We are not claiming here that JuicyBurger became fully immersed in forum participation; indeed, he ended his participation after the eight-week period. This outcome is not surprising, though, since he was a full-time student. Also, he and the other participants were recruited and paid, which undoubtedly affected their attitude towards the activity. Many of the established FMH helpers have remained members for several years. On the other hand, we are suggesting that JuicyBurger's behavior during his eight week stint was consistent with becoming immersed in this community, and that this process manifested itself in a variety of ways (aside from simply time spent participating on the forum). Immersion can be marked by active engagement in the joint community effort and by demonstrating enthusiasm for interests shared by other members.

Discussion

Free, open, online, help forums represent the movement to democratize education through technological innovation (Larreamendy-Joerns & Leinhardt, 2006), as well as the increasing relevance of social networks in learning and instruction (Goodyear et al., 2005). The forums provide students around the world with affordable (free), efficient, and accessible help when and where they need it. Students who might not otherwise have access to resources outside of the classroom have the opportunity to ask questions of more experienced others. At the same time, the forums are a unique location where helpers, who traditionally operate individually and privately, can practice alongside others within a community. When the existing set of core established helpers form a closely-knit social group, newcomers, who are necessary for the growth and sustainability of the endeavor, must be able to ease their way into the community in order to become core participants. This paper addressed how a newbie helper navigated a trajectory from peripheral to fuller participation in a mathematics help forum with a strong sense of virtual community, learning how to work alongside established members as they helped students learn. Sheer volume of contributions was not sufficient. Another helper in our study posted more to the forum than JuicyBurger and yet did not make comparable inroads into the community. Instead, this process involved cultivating a sense of belonging and building supportive relationships with others (membership), contributing meaningfully and creating a voice (influence), and demonstrating dedication to the shared purpose of the community (immersion). This paper is a first step at establishing a framework for what it means for newbies to join in a sense of virtual homework help forum community. Our long-term goal is to be able to scaffold this experience so that people can more easily join, and thereby help sustain, these communities.

References

- Balasubramanian, S. & Mahajan, V. (2001). The economic leverage of the virtual community. *International Journal of Electronic Commerce*, 5(3), 103-138.
- Fernback, J. & Thompson, B. (1995). Virtual communities: Abort, retry, failure? Computer Mediated Communication and the American Collectivity. Available at www.rheingold.com/texts/techpolitix/VCcivil.html
- Goodyear, P., Jones, C., Asensio, M., Hodgson, V., & Steeples, C. (2005). Networked Learning in Higher Education: Students' Expectations and Experiences. *Higher Education*, 50 (3), 473-508.
- Heller, K. (1989). Return to community. *American Journal of Community Psychology*, 17(1), 1-15.
- Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21, 31-46.
- Koh, J. & Kim, Y-G. (2003-4). Sense of virtual community: A conceptual framework and empirical validation. *International Journal of Electronic Commerce*, 8(2), 75-93.
- Larreamendy-Joerns, J. and G. Leinhardt (2006). Going the distance with online education. *Review of Educational Research*, 76(4), 567-605.
- Lave, J. & Wenger, E. (2002). Legitimate peripheral participation in communities of practice. In M. R. Lea & K. Nicoll (eds.), *Distributed learning: social and cultural approaches to practice*, pp. 56-63. London: Routledge-Falmer.
- Martinovic, D. (2005). *Characteristics of asynchronous online mathematics help environments: Do they provide conditions for learning?* Proceedings of the 12th International Conference on Artificial Intelligence in Education, Workshop 6, pp. 32-38, Amsterdam, the Netherlands.
- McMillan, D. W. & Chavaz, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology*, 14(1), 6-23.
- Preece, J. (1999). Empathic communities: Balancing emotional and factual communication. *Interacting with Computers*, 12, 63-77.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. London: Sage.
- van de Sande, C. (2008). *Open, online, calculus help forums: Learning about and from a public conversation*. (Unpublished doctoral dissertation). University of Pittsburgh, PA.
- van de Sande, C. (2010, June). *Free, open, online, mathematics help forums: The good, the bad, and the ugly*. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences*, pp. 643-650, Chicago, IL: International Society of the Learning Sciences.
- van de Sande, C. (2011) *How are students' problems being solved? The quality of worked solutions on a popular, open, online, mathematics help forum*. In Spada, H., Stahl, G., Miyake, N., Law, N. (Eds.) *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings*. International Society of the Learning Sciences, pp. 280-287.
- van de Sande, C. and G. Leinhardt (2007). Online tutoring in the Calculus: Beyond the limit of the limit. *Éducation et Didactique*, 1(2), 115-154.
- Wenger, E. (1998). *Communities of practice: learning, meaning and identity*. New York: Cambridge University Press.

Multidimensional Teacher Behavior in CSCL (1)

Anouschka van Leeuwen, Jeroen Janssen, Gijsbert Erkens, Mieke Brekelmans
 Utrecht University, Heidelberglaan 1, 3584CS Utrecht, the Netherlands
 A.vanLeeuwen@uu.nl, J.J.H.M.Janssen@uu.nl, G.Erkens@uu.nl, M.Brekelmans@uu.nl

Abstract: Situated in the field of computer-supported collaborative learning, the aim of this study is to present a multi-dimensional approach for the examination of teacher behavior. Two dimensions were used for coding: focus (what the intervention is aimed at) and means (how the teacher intervenes). Teacher behavior was studied for a period of several weeks. Our results indicate that both focus and means changed as time progressed, and that type of interventions varied between groups. Characterizing teacher behavior as a particular type, as done in other research, is therefore refuted. Conclusions that hold for the whole of the studied period are that number of teacher interventions was related to amount of student activity, and that the teacher focused more on task content than group collaboration. In the discussion suggestions are given for teacher supporting tools.

Introduction

Computer-supported collaborative learning (CSCL) combines collaborative learning with the use of information and communication technologies. There is a broad range of types of supporting tools specifically aimed at helping students carry out the learning task (Soller, Martinez, Jermann, & Muehlenbrock, 2005). In settings in which these tools are used, there often is no mention of the presence of a teacher, i.e., the students seemingly work on their assignment independently.

However, recently there has been a growing interest in the role of the teacher during CSCL. This has led to the development of supporting tools not only for students, but also for the teacher guiding the students (see for example the Argonaut project; Asterhan & Schwarz, 2010). These tools could deliver information that enables teachers to better carry out their tasks, for example about group progress or collaborative processes (McLaren, Scheuer, & Mikšátko, 2010).

To develop and test the effectiveness of supporting tools for teachers, there is a need to know what kind of behavior teachers employ, and what kind of behavior is effective. Many researchers have aimed at first studying which behavior teachers during CSCL display without any supporting tools present (Azevedo, Cromley, Winters, Moos, & Greene, 2005; Greiffenhagen, 2012; De Smet, Van Keer, & Valcke, 2008). What appears from these studies is that teachers face a very complex task (Volman, 2005). The present study aims to contribute to this knowledge base by using a multi-dimensional conceptualization of teacher activities, and by studying the variation in teacher activities between lessons and between groups.

Complexity of a Teacher's Task in CSCL

There are several reasons why the teacher's task of guiding students' learning in a CSCL environment is so complex. First of all, comparable to face-to-face collaborative situations, teachers deal with several types of synchronicity. Not only do they guide several groups' activities at the same time, they also have to focus on both task content and collaboration between students. Furthermore, differences between groups require the teacher to adapt his guidance to the groups' needs (Chiu, 2004). There are also complicating factors that are unique to CSCL. Teachers are offered information not only about the learning result but also about the learning *process*: teachers are able to follow discussions between students, and in some digital environments teachers are given information about students' progress. It may be easier for a teacher to monitor students' learning processes when given this extra information. On the other hand, having access to such amounts of information could also cause an information overload for the teacher (Van Diggelen, Janssen, & Overdijk, 2008).

Analyzing Teacher Interventions among Multiple Dimensions

As was mentioned in the introduction, many researchers have aimed at studying natural teacher behavior in CSCL. The content of teacher interventions is often studied using a one-dimensional conceptualization. When the content of teacher interventions is analyzed, it is almost always the case that each intervention is assigned to one particular category. That is, one dimension is used to categorize teacher interventions, and the categories within the dimension are mutually exclusive. In contrast, it could be argued that each intervention can be studied from multiple angles. In other words, each intervention can be categorized among multiple dimensions.

The content of a teacher's intervention both has a focus and a means. For example, when a teacher tells a group of students to "Start by reading the assignment", the focus is on the regulation of cognitive activities, while the means is an instruction. A similar difference between the 'what' (focus) and 'how' (means) of teacher interventions was also noticed in the literature about teacher regulation of face-to-face collaboration in a review study about scaffolding by Van de Pol, Beishuizen, and Volman (2010). Why is such a distinction useful? Van

de Pol et al. (2010) suggest that “this distinction [...] enables us to look more precisely at interactions and results in more nuanced descriptions of teacher–student interactions” (p. 276).

Very few studies on CSCL employ such a two dimensional analysis for the content of teacher interventions. Typically, the methodologies used are similar to either focus or means. Some studies approach the focus of teacher interventions, i.e. what the teacher’s intervention is aimed at. Lund (2004) for example includes the categories ‘Social’, ‘Managerial’, and ‘Pedagogical’. It is generally considered important for teachers to focus on students’ cognitive (task-related) activities as well as social activities within a group (Kreijns et al., 2003). Other studies approach the means of teacher interventions, i.e. how the teacher intervenes. Consider the study by Greiffenhagen (2012), in which the categories ‘Making announcements’, ‘Reminding’, and ‘Suggesting’ are used. Van de Pol et al. (2010) listed the categories of means that are most commonly identified: feeding back, hints, instructing, explaining, modeling, and questioning.

Change between Lessons and Variation between Groups

The research on teacher interventions described so far has a one-dimensional character in the sense that one aspect of the teacher’s task is analyzed (either focus or means). Another characteristic of these studies is that results are often presented on an aggregated level. That is, conclusions are drawn about teacher interventions at a general level. Some researchers aim to characterize particular teacher ‘types’ or ‘moderation profiles’ (Asterhan, 2011). For example, Mazzolini and Maddison (2003) distinguish between a dominating ‘sage on the stage’ and a less pervasive ‘guide on the side’, based on summaries of teacher interventions taken as a whole.

It should be kept in mind that in CSCL environments students often work on a complex assignment that takes multiple lessons to complete. Such tasks contain multiple phases during which students perform different kinds of activities (Erkens, Jaspers, Prangma, & Kanselaar, 2005). It is therefore likely that not only the students’ behavior, but also the type of teacher’s interventions change as the number of sessions progresses (Onrubia & Engel, 2012). This importance of time scales and temporality has recently received great attention in CSCL literature, leading to a movement that calls for a temporal analysis of interaction data. It is emphasized that “temporality matters” (Kapur, 2011; Reimann, 2009), both on a small scale in patterns of turn-taking within interaction and on a larger scale as change or variation between sessions or lessons. As Kapur (2011) notes, “By aggregating counts over time, information about temporal variation is lost” (p. 41). Too little is known yet about the change in teacher interventions between lessons and the variation of teacher interventions between groups.

Aim of this Study

Two methods have been pointed out that could contribute to a better understanding of the complex task teachers face when they guide students’ learning in a computer-supported collaborative learning environment. The first method is to analyze the content of teacher interventions on multiple dimensions: both focus and means. The second method is to analyze teacher interventions taking into account change between lessons and variation between groups. This study aims to bring together these two methods. A case study is presented in which the interventions by one teacher, spread over multiple weeks of activity, are analyzed. The aim is to contribute to the conceptualization of teacher interventions during CSCL, and to contribute empirically to the existing literature on the characteristics and dynamics of teacher interventions in CSCL. The following research questions have been formulated:

RQ1 Which interventions, in terms of focus and means, does the teacher use?

RQ2 How do teacher interventions, in terms of focus and means, change between lessons and vary between groups while students work on the assignment?

Method

Participants

One secondary education history male teacher (age 43, with 15 years teaching experience), and 21 students (age $M = 15$, $SD = 0.6$) were involved in the study, who were all enrolled in the third year of the pre-university education track. Students were assigned by the teacher into groups of three students (7 groups).

CSCL Environment

Students and teacher made use of the CSCL environment called Virtual Collaborative Research Institute (VCRI, see for example Janssen et al., 2007). It was used in a synchronous, co-located setting. All students had their own computer. The assignment involved exploring the topic by reading historical sources in the Sources-tool. Students could discuss the information through the synchronous Chat-tool. Students used the Debate-tool to construct a shared diagram of their arguments. Students used the Cowriter, a shared text processor, to write their texts.

An alternative interface of the VCRI-program was available for the teacher, which allowed him to monitor the online discussions of the students in the Chat-tool in real-time and send messages in order to answer students’ questions. Teachers can examine the texts students are writing in the Cowriter or the diagrams they are

making in the Debate-tool. The teacher intervened by sending messages through the Chat-tool. The program offered the teacher some basic statistical information about students' activities in VCRI's tools (e.g., the number of keystrokes per student). Figure 1 shows the teacher interface.

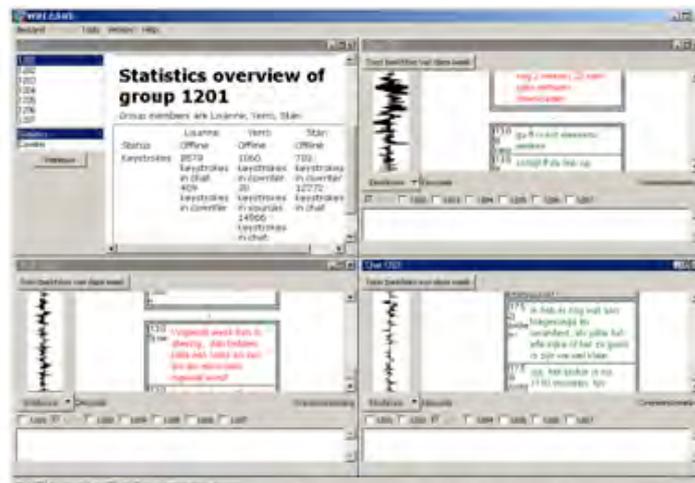


Figure 1. An example of the configuration of the VCRI teacher interface. In this case, the teacher has opened student activity statistics (upper left) and the *Chat*-tool for three of the groups.

Assignment

Students collaborated on a group task about the Cold War. It was an open-ended group task that focused on reading, comprehending, and synthesizing historical sources. The assignment was split into four parts, each resulting in an argumentative text written by the groups of students. The first two parts asked students to reflect on the consequences of World War II and the events during the Cold War. The third part required students to summarize their arguments in a graphical representation in the Debate-tool. This resulted in part four, a final text on why and how the Cold War ended. The class worked with VCRI for 8 lessons of 50 minutes.

Analysis of Teacher Interventions

Each of the teacher's interventions (messages) were coded along two dimensions: focus and means. The codes for focus (Table 1) are based on the distinction between cognitive and social activities, which are the two categories generally considered important for teachers to support students' learning process (Salmon, 2000). These two categories are further split into the object and the meta (regulative) level (Molenaar et al., 2011). Codes for means were derived from the review study by Van de Pol et al (2010). Their categories feeding back, hints, instructing, explaining, and questioning were used, and the category questioning was further expanded by adding diagnosing and prompting. Diagnosing denotes questions that are aimed at checking students' progress. Prompting denotes questions that are meant to help or activate students.

Two independent coders both coded a random sample of about 100 chat messages. Cohen's κ for focus was .77, for means it was .76. After establishing the interrater reliability of the coding procedure, all messages sent by the teacher were analyzed.

Table 1: Coding scheme for the focus of interventions.

Focus	Definiton	Example
Cognitive activities (Cog)	Utterances about task content	"SU means Sovjet Union"
Regulation of Cognitive activities (RegCog)	Utterances about planning of the task / time management Utterances about task strategies	"Start reading the sources"
Social activities (Soc)	Utterances that contribute to the mood within a group or the class or express emotions	"Come on, let's get to work"
Regulation of Social activities (RegSoc)	Utterances about the collaboration process / about strategies for collaboration	"Divide the tasks among your group"
Other	Remaining utterances	

Results

Overall Results

During the 8 lessons the teacher and students worked in VCRI, the teacher sent 391 interventions. Of these 391 interventions, 192 (49.1%) were reactions to student questions. The teacher used the system 6 times between lessons to read student contributions in the Cowriter tool. In every lesson the teacher checked the statistics right at the start of the lesson, and on average another 4 times during the remainder of the lesson.

Table 2 summarizes the focus and means of the teacher's interventions for all lessons taken together. The means used most often are Explaining (25.2%), Feedback (22.7%), and Diagnosing (20.1%). The focus of interventions was most on cognitive activities (41.8%), closely followed by focus on the regulation of cognitive activities (39.6%). Relatively few interventions were focused on (regulation of) social activities (together making up 15.1% of interventions).

Table 2: Focus (columns) and means (rows) of all 397 teacher interventions taken together, in percentages.

	Cog	RegCog	Soc	RegSoc	Other	Total
Questioning	0.3	0.0	0.8	0.0	0.0	1.0
Diagnosing	3.3	15.1	0.5	0.5	0.8	20.1
Prompting	11.3	0.0	0.0	1.0	0.0	12.3
Hinting	3.3	4.5	0.0	1.3	0.0	9.1
Explaining	16.1	6.1	0.0	1.0	2.0	25.2
Instructing	0.3	4.0	0.0	0.8	0.0	5.0
Feedback	7.1	9.6	4.0	1.8	0.3	22.7
Other	0.3	0.3	3.5	0.0	0.5	4.5
Total	41.8	39.6	8.8	6.3	3.5	100.0

Some groups were more active than others. It appeared that the higher student activity, the more teacher interventions to the group. Analysis of the individual lessons showed there was indeed a significant relationship between the number of student messages and the number of teacher interventions, $r = .38$, p (one-tailed) $< .01$.

Change between Lessons and Variation between Groups

Figure 2 and 3 display the focus and means of teacher interventions for each of the 8 lessons that students worked on the assignment.

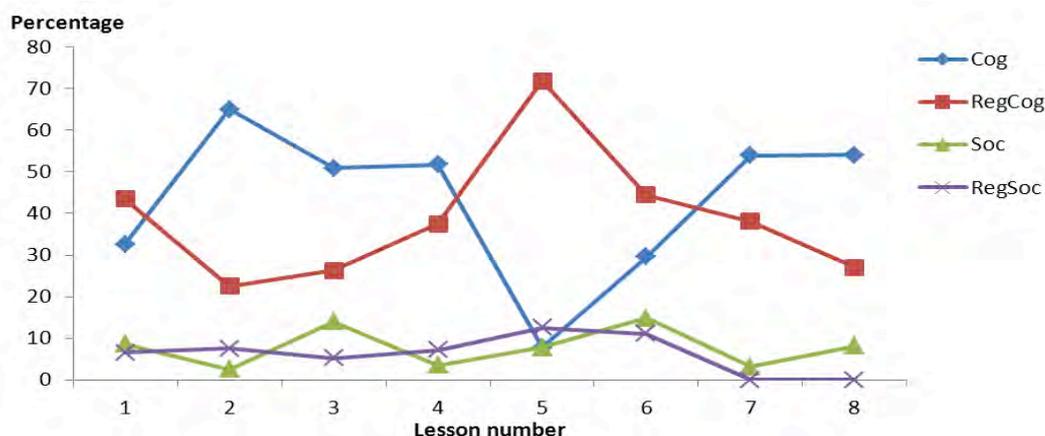


Figure 2. Focus of teacher interventions in percentages, displayed per lesson.

In line with the overall results, the focus is mostly on (regulation of) cognitive activities. However, the relative frequencies fluctuate for each lesson. For example, there is a strong predominance of focus on regulation of cognitive activities during the middle part. A relatively small portion of teacher interventions focused on (regulation of) social activities. These categories score low throughout all lessons. The variation between groups for these categories is relatively high. For example, the coefficient of variation for focus on regulation of social activities in lesson 3 is 1.3 where the corresponding score in Figure 2 is 5%. This means that the standard deviation was relatively high and that the teacher focused on regulation of social activities only in selected groups. The same can be said for the focus on regulation of social activities in later lessons (5 and 6).

As for the means (Figure 3), the most occurring categories found in the overall results (Explaining, Feedback, Diagnosing) are not predominant when viewing each lesson separately. For example, the relative frequency of Explaining is especially low during the middle part, and the relative frequency of Feedback is quite low at the beginning. The other way around it is apparent that categories that score relatively low overall, like Instructing (5% of total), occur relatively often during some lessons. For example, Instructing reaching 20% during the middle part (lesson 6). It thus seems there is a change in the type of interventions as time progresses.

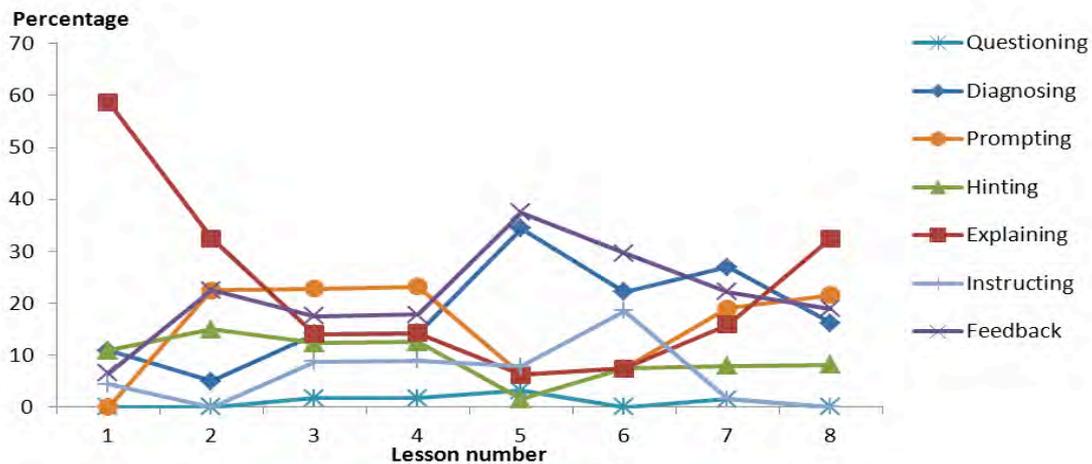


Figure 3. Means of teacher interventions in percentages, displayed per lesson.

Below, some examples of teacher-student interaction are given to illustrate the change between lessons and the variation between groups.

Example Lesson 1: Explaining

At the start of working with VCRI, there is a peak of Explaining (Figure 3). The focus of the teacher’s interventions is equally distributed among cognitive activities and the regulation of cognitive activities (Figure 2). In the first lessons lesson, the teacher explains how to start working and he also explains some basic concepts that students need throughout the assignment. There are also some explanations that deal with the program itself.

Example Lesson 3: Prompting

Next, a period follows in which there is more focus on cognitive activities (Figure 2). The teacher uses more open forms of interventions, such as Prompting (Figure 3). In Table 3, the students are discussing the results of the Second World War. Instead of giving away the answer, the teacher responds by posting counter questions.

Table 3: Example of teacher interventions during lesson 3.

User	Message	Focus	Means
Student 3	I can’t find any reports on damage to the US after WOII		
Teacher	If you can’t find it, maybe there isn’t any? Where did they fight?	Cognitive activities	Prompting
Student 3	the pacific?		
Teacher	What does that mean?	Cognitive activities	Prompting

Example Lesson 5: Diagnosing

In lesson 5 there is a peak in focus on Regulation of cognitive activities (Figure 2) as well as a peak in Diagnosing and Feedback (Figure 3). Looking at the chat protocols, there are two things that the teacher is paying attention to. First of all, he is checking how the groups are doing in terms of progress on the assignment and instructing them to move on to the next part. Secondly, in one of the small groups, it appears there is a collaborative problem (Table 4). One student has been absent for a few lessons and has contributed less to finishing the assignment. The teacher does not think this is a valid reason for lagging on the assignment and tells the students to improve their collaboration.

Table 4: Example of teacher interventions during lesson 5.

User	Message	Focus	Means
Teacher	Have you finished part 2?	Regulation of cognitive activities	Diagnosing
Student 2	no, because sally has been absent the last two lessons		
Teacher	You are perfectly able to finish the assignment without her, I think	Regulation of social activities	Feedback
Student 3	She has written something I don't understand		
Teacher	You're working as a team, so everyone is responsible for the assignment	Regulation of social activities	Feedback

Example Lesson 7: Cognitive Activities

After the dip in focus on cognitive activities, the final lessons show an increase in focus on this area again (Figure 2). The means used in lesson 7 are varied (Figure 3). The chat protocols show that during lesson 7, there is again more time spent on the content of the task. Students start the final part of the assignment, which also accounts for the increase in Explaining.

Discussion

It was argued that an adequate description of teacher interventions during CSCL warrants a multi-dimensional and temporal analysis. In this study, focus and means of teacher interventions were analyzed and compared between lessons and between groups.

Focus

The results show that overall, the focus was mostly on (regulation of) cognitive activities. Within the individual lessons, this dominance was also visible. However, the relative frequencies fluctuated for each lesson. The number of teacher interventions focusing on (regulation of) social activities was much lower, reaching a maximum percentage of about 15%. There was little change between lessons, but large variation between groups. The results from other studies do not uniformly fit with our findings. Asterhan (2011) for example found no instance of social support at all, only interventions aimed at (regulation of) cognitive activities. De Smet et al. (2008) on the other hand found comparable low frequencies for focus on social activities, which increased at the end of the studied period. This contradicts our finding of low variability between lessons.

Many researchers have pointed out the importance of social activities within collaborating groups (Kreijns et al., 2003; Salmon, 2000). From our data it appears that the need to focus on social activities may have been small. A possible explanation for this is that in contrast to distant learning and asynchronous settings, the students in this study already knew each other very well. Therefore, the teacher only intervened when this was absolutely necessary. However, it could be argued that even when there are no apparent problems, it is important that a teacher focuses on the social processes during collaboration and makes students aware of their individual and group behavior (Phielix, Prins, & Kirschner, 2010).

Means

The results of this study show that the means most often used were feedback, explaining, and diagnosing. It must be noted that there was variation between lessons (see Figure 3) and that the general percentages are not an accurate depiction of this variation. Some time periods showed high percentages of open forms of interventions (lessons 2 and 3) while in others closed forms were used more often (for example lesson 6).

Again, it seems the teacher's interventions were group-specific. For example, in the first lessons there were a lot of explanations. It may have been expected that this was due to getting used to working in a CSCL setting, but it turned out that not so much the system, but the assignment needed more attention. In the Results section this was illustrated with some examples. Research has shown that if the task materials exert more control, that is, contain more detailed instructions for carrying out the task, the teacher might need less closed forms of interventions (Onrubia & Engel, 2012). In this case, the task might have exerted a very low amount of control, therefore eliciting more explanations from the teacher.

The change between lessons is consistent with other studies that have studied teacher interventions on multiple time points (albeit with fewer time points than in this study). Hsieh & Tsai (2012) found that means of teacher interventions varied during the three sessions students collaborated. However, the three sessions dealt with separate assignments, not an ongoing assignment like in our study. Gil et al. (2007) divided the data on teacher interventions in three phases; beginning, middle and end, and found differing distributions of means of intervening for each of the five teachers that were studied.

Implications

The combination of coding both focus and means has resulted in a nuanced description of teacher interventions. Although there were some indications of patterns, the general finding is that this teacher's interventions are foremost explained by sensitivity to group needs. Generalizations of teacher behavior into 'types' (Mazzolini & Maddison, 2003) or 'moderation profiles' (Asterhan, 2011) do not reflect the whole complexity of the teacher's task. It is clear from our data that the teacher's method of regulation was not the same throughout the whole timeframe. Instead, it seems the teacher's behavior was context-specific.

Some suggestions may be made about the type of supporting tools to be developed, based on the results presented here. The first suggestion would be to support teachers in monitoring the social activities of groups of students. The benefits of presenting visualizations of agreement and discussion processes to students have already been investigated (Janssen et al., 2007). Studying whether teacher access to such visualizations has an effect on teacher and student behavior would be an interesting step forwards.

A further suggestion for supporting tools is based on the high amount of Diagnosing interventions performed by the teacher in the current study. Interviews with teachers have shown that one of the main difficulties of guiding students during CSCL is keeping track of all information (De Smet, Van Keer, De Wever, & Valcke, 2009; Van Diggelen et al., 2008). Our results show that the frequency of Diagnosing increased as time progressed. This may have been due to the fact that in the last part of the assignment, the students started writing their final essay. It may have been harder for the teacher during this part to monitor the progress within the groups. This is supported by the fact that 20% of the interventions that the teacher sent to the whole class simultaneously were Diagnosing interventions. Therefore, teachers may benefit from tools that provide information on the status and quality of written products (Dillenbourg, Järvelä, & Fischer, 2009).

Limitations and Future Directions

An obvious limitation of this study is that only one teacher's interventions were analyzed. It is important that more studies are performed within this area, especially ones that have an experimental set-up. On the other hand, case studies do provide the opportunity to study the subject in great detail, which was exactly the aim of this study. Furthermore, in this particular article the conceptualization of teacher interventions also played a big role, which makes the number of participants a less important issue. Another limitation is that although we have described the complexity of studying the content of teacher interventions, there are many more aspects that play a role during CSCL. For example, we briefly stated that in our study, 49.1% of teacher interventions were initiated by a question from a student. Onrubio & Engel (2012) further expand on whether the teacher or the student initiates the conversation. A challenge for future research is to connect all aspects into a coherent image.

Conclusion

The ultimate conclusion that emerges from this study is that the teacher displayed a wide range of types of interventions. Some reasons for this variation were to be expected, such as the peak of explanations and the focus on regulation of cognitive activities at the start of the assignment, while others were highly group specific, such as the need to end a fight between students. We argue for studying teacher interventions as a reaction to the current situation, not as an expression of the teacher's "style". As Schwarz & Asterhan (2011) note, all teacher interventions are useful "as long as they are attuned to the needs of the group and its individual members at that time" (p. 436).

The starting point of this study was the premise that combining focus and means would lead to a more nuanced view of teacher interventions (Van de Pol et al., 2010) and the complex task that teachers face during CSCL. Each of the dimensions on its own indicated clear differences between lessons and between groups, but taken together a broader perspective emerged that pointed to several points of interest in the data.

It is important that a clear picture is obtained of the complexity of a teacher's task to regulate students' learning during CSCL, not only for theoretical purposes, but also for practical reasons. Ultimately, CSCL environments may be designed that not only support students, but also the teacher. By summarizing student activity, for example, the teacher's task of Diagnosing could become easier (Dillenbourg, Järvelä, & Fischer, 2009). Gaining a complete view of what a teacher does by using appropriate methods of analysis, is the first step towards this goal.

Endnotes

- (1) An extended version of this paper was published as: Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2013). Teacher interventions in a synchronous, co-located CSCL setting: analyzing focus, means, and temporality. *Computers in Human Behavior*, 29(4), 1377-1386. doi:10.1016/j.chb.2013.01.028

References

- Asterhan, C. (2011). Assessing e-moderation behavior from synchronous discussion protocols with a multi-dimensional methodology. *Computers in Human Behavior*, 27(1), 449-458.

- Asterhan, C., & Schwarz, B. (2010). Online moderation of synchronous e-argumentation. *International Journal of CSCL*, 5, 259-282.
- Azevedo, R., Cromley, J. G., Winters, F. I., Moos, D. C., & Greene, J. A. (2005). Adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instructional Science*, 33, 381-412.
- Chiu, M.M. (2004). Adapting teacher interventions to student needs during cooperative learning: how to improve student problem solving and time on-task. *American Ed Research Journal*, 41(2), 365-399.
- De Smet, M., Van Keer, H., Valcke, M. (2008). Blending asynchronous discussion groups and peer tutoring in higher education: An exploratory study of online peer tutoring behavior. *Computers & Education*, 50(1), 207-223.
- De Smet, M., Van Keer, H., De Wever, B., & Valcke, M. (2009). Studying thought processes of online peer tutors through stimulated-recall interviews. *Higher Education*, 59(5), 645-661.
- Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The Evolution of Research on Computer-Supported Collaborative Learning: From Design to Orchestration. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder, S. Barnes (Eds.), *Technology enhanced learning: Principles and products* (pp. 3-19).
- Erkens, G., Jaspers, J., Prangma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior*, 21(3), 463-486.
- Gil, J., Schwarz, B. B., & Asterhan, C. (2007). Intuitive moderation styles and beliefs of teachers in CSCLbased argumentation. In C.A. Chinn, G. Erkens & S. Puntambekar (Eds.), *Mice, minds and society. Proceedings of the CSCL Conference* (p. 219-228). International Society of the Learning Sciences, Inc.
- Greiffenhagen, C. (2012). Making rounds: The routine work of the teacher during collaborative learning with computers. *International Journal of CSCL*, 7(1), 11-42.
- Hsieh, Y. & Tsai, C. (2012). The effect of moderator's facilitative strategies on online synchronous discussions. *Computers in Human Behavior*, 28(5), 1708-1716.
- Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior*, 23, 1105-1125.
- Kapur, M. (2011). Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of CSCL*, 6(1), 39-56.
- Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments. *Computers in Human Behavior*, 19, 335-353.
- Lund, K. (2004). Human support in CSCL: What, for whom and by whom? In J.W. Strijbos, P.A. Kirschner, R.L. Martens, & P. Dillenbourg (Eds.), *What we know about CSCL and implementing it in higher education* (Vol. 3, pp. 167-198). Norwell, MA: Kluwer Academic Publishers.
- Mazzolini, M., & Maddison, S. (2003). Sage, guide or ghost? the effect of instructor intervention on student participation in online discussion forums. *Computers & Education*, 40, 237-253.
- McLaren, B. M., Scheuer, O., & Mikšátko, J. (2010). Supporting Collaborative Learning and E-Discussions Using Artificial Intelligence Techniques. *International Journal of AI in Education*, 20, 1-46.
- Molenaar, I., van Boxtel, C. & Slegers, P. J. C. (2011). Metacognitive scaffolding in an innovative learning arrangement. *Instructional Science*, 39, 785-803.
- Onrubia, J., & Engel, A. (2012). The role of teacher assistance on the effects of a macro-script in collaborative writing tasks. *International Journal of CSCL*, 7(1), 161-186.
- Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL environment: Effects of peer feedback and reflection. *Computers in Human Behavior*, 26, 151-161.
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of CSCL*, 4(3), 239-257.
- Salmon, G. (2000). *E-moderating. The key to teaching and learning online*. London: Kogan Page.
- Schwarz, B., & Asterhan, C. (2011). E-Moderation of Synchronous Discussions in Educational Settings: A Nascent Practice. *Journal of the Learning Sciences*, 20(3), 395-442.
- Soller, A., Martinez, A., Jermann, P., and Muehlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal of AI in Education*, 15, 261-290.
- Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, 22, 271-296.
- Van Diggelen, W., Janssen, J., & Overdijk, M. (2008). Analysing and presenting interaction data: a teacher, student and researcher perspective. ICLS'08 Proceedings of the 8th international conference on International conference for the learning sciences - Volume 3.
- Volman, M. (2005). A variety of roles for a new type of teacher: Educational technology and the teaching profession. *Teaching and Teacher Education*, 21, 15-31.

Learning to Argue in Mathematics: Effects of Heuristic Worked Examples and CSCL Scripts on Transactive Argumentation

Freydis Vogel¹, Elisabeth Reichersdorfer², Ingo Kollar¹, Stefan Ufer¹, Kristina Reiss², Frank Fischer¹

¹LMU Munich, Leopoldstr. 13, D-80802 Munich, Germany,

²Technische Universität München, TUM School of Education, Schellingstr. 33, D-80799 Munich, Germany,

Email: freydis.vogel@psy.lmu.de, elisabeth.reichersdorfer@tum.de, ingo.kollar@psy.lmu.de,

ufer@math.lmu.de, kristina.reiss@tum.de, frank.fischer@psy.lmu.de

Abstract: A previous study has shown that both CSCL scripts and heuristic worked examples implemented in a CSCL environment were effective to fostering students' acquisition of argumentation skills in the context of mathematical proof tasks (Kollar, et al. 2012). This paper investigates the extent to which transactive argumentation during the collaborative learning process can be evoked by both means of instructional support and to what extent transactive argumentation mediates their effects on students' knowledge about argumentation. We present process measures from a 2x2-factorial experiment with the factors CSCL script and heuristic worked examples conducted with $N=101$ prospective math teacher students. Results show that both means of instructional support induced transactive argumentation in the collaborative learning process. The self-generated transactive argumentation, but not the partner-generated transactive argumentation mediated the effects of both types of instructional support on students' development of argumentation knowledge. Nevertheless, the learning partners mutually influenced their transactive argumentation.

Introduction

Over the last decade CSCL research has focused on argumentation as a goal of educational interventions. To foster students' argumentation and to support them to develop the corresponding skills, various instructional approaches (e.g. CSCL scripts, representational guidance) have been designed and evaluated across various domains (for an overview, see Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2012). Argumentation skills are also important in mathematical discourse, particularly for working on mathematical proof problems. During the process of mathematical proof, argumentation skills are required at different points (Aberdein, 2009). Yet, students often have problems to construct arguments in general as well as in mathematical context. For example, Sadler (2004) summarized that students show serious difficulties in socioscientific argumentation (e.g.: they do not justify claims, they do not take any counter-argument into account, etc.). Within the mathematical domain, Heinze, Reiss, and Rudolph (2005) found that high school students were able to solve problems, requiring one single argument, but failed in producing logical chains of more than one argument in mathematical argumentation tasks. Thus, students' efforts to acquire argumentation skills within the mathematical context should be supported by using adequate instruction. The study presented in this paper is embedded in the context of a project that investigated the effectiveness of two kinds of instructional support on students' acquisition of argumentation skills in mathematical proof tasks. More specifically, a CSCL script adapted from Stegmann, Weinberger, and Fischer (2007) and heuristic worked examples (Reiss & Renkl, 2002) were applied. In previous analyses within this project, we showed that providing students with the CSCL script and heuristic worked examples both had positive effects on students' acquisition of argumentation skills (Kollar, et al., 2012). Yet, it is still not clear which collaborative learning processes led to these effects. Thus, the main purpose of this paper is to provide an analysis of the collaborative learning processes of the dyads in a mathematics learning environment. Especially, we investigate to what extent the two treatments caused transactive argumentation (i.e. learning partners mutually refer to each other using argumentative moves like criticizing). Further we explore if the induced transactive argumentation can explain the effectiveness of both scaffolds on students' development of knowledge about argumentation as a part of argumentation skills. We also analyse if there is a difference of the effectiveness between transactive argumentation the learners generated themselves and transactive argumentation that was generated by their respective learning partner and to what extent the frequency of transactive argumentation expressed by the learning partners mutually influenced each other.

CSCL scripts and heuristic worked examples for mathematical argumentation

Argumentation skills (i.e. the skills to engage in a social-discursive argumentative dialog in an effectual way) might intuitively rather be required in domains like politics or philosophy. Nevertheless, they are also necessary in mathematics (e.g. Schwarz & Linchevski, 2007), in particular when it comes to mathematical proof problems. According to Boero (1999), the proof process consists of the following steps: (1) generation of a conjecture, (2) formulation of a mathematical statement, (3) exploration of the mathematical statement, (4) selection of

adequate theorems to generate a proof draft, (5) construction of the proof draft, (6) formulation of the formal proof. At several points within this process, argumentation needs to be applied, e.g., when one has to find or evaluate a conjecture, to choose applicable arguments, or when a formal proof must be presented to a broader public (Aberdein, 2009). Thus, for argumentation in the mathematical domain formal patterns of the construction of single arguments and the social process of argumentation between dialog partners can be found, similar to other domains. For instance, Toulmin's (1958) argument schema is widely used for the evaluation of single arguments (van Eemeren & Grootendorst, 2004). Also, dialectical forms of argumentation - simplified as the cycle of 'argument', 'counter-argument', and 'synthesis' - might function as a common ground for social discursive activities where two or more dialogue partners are engaged in an argumentative discourse (Leitão, 2000). Recently, various kinds of CSCL instructions that support students' acquisition of argumentation skills have been investigated (Noroozi, et al., 2012). One instructional approach that has shown positive effects on the acquisition of argumentation skills is scripting (e.g. Stegmann, et al., 2007). In general, CSCL scripts distribute roles and activities among the learners and sequence activities and role changes to guide students through a collaborative learning process that is beneficial for their learning (King, 2007; Kollar, Fischer, & Hesse, 2006), both with respect to domain-specific knowledge and the internalization of the domain-general skills a script has learners to practice during learning (Fischer, Kollar, Stegmann, & Wecker, 2013). CSCL scripts that are designed for argumentation guide students through argumentative discourses by prompting them to fulfil adequate activities within each step of an argumentative discourse cycle (e.g. Hron, Hesse, Cress, & Giovis, 2000; Weinberger, Stegmann, & Fischer, 2010) or by distributing discussion roles among the learning partners (e.g. De Wever, Van Keer, Schellens, & Valcke, 2010). Studies about CSCL scripts for argumentation have shown positive effects on students' acquisition of domain-general argumentation skills. For instance, the study by Stegmann et al. (2007) could show that students learning with a CSCL script that was based upon the dialectical cycle of argument, counterargument and synthesis (Leitão, 2000) and Toulmin's argument schema (1958) outperformed students learning without collaboration support in developing argumentation skills. But there has not been systematic research on CSCL scripts for argumentation in the mathematical domain.

Yet, scaffolding collaborative learning processes may not be enough to help students acquire argumentation skills. A review by Vogel, Kollar, and Fischer (2012) revealed that the effectiveness of CSCL scripts on the acquisition of domain-specific knowledge can be advanced by combining them with additional instructional support that provides domain-specific content knowledge (e.g. content schema; Ertl, Kopp & Mandl, 2006). An improvement of the effectiveness of CSCL scripts through the simultaneous provision of domain-specific support may also be expected for the acquisition of domain-general skills (e.g. argumentation skills), since domain-specific instructional support provides content knowledge that can be more deeply elaborated when collaboration is guided by a script. When two kinds of instructional support are used in combination, at least an additive effect would be desirable, i.e. that the (positive) effects of two kinds of instructional support add up when applied together, but do not positively amplify each other. The optimum for the combination of two kinds of instructional support would be synergistic scaffolding (Tabak, 2004). Given the expectation of achieving synergistic scaffolding when combining a CSCL script with a domain-specific instructional support, *heuristic worked examples* were implemented as domain-specific instructional support for the present study. Generally, worked examples provide students with an elaborated worked out solution that is exemplary for solving the type of problem tasks assigned to the learners. While traditional worked examples (e.g. Atkinson, Derry, Renkl, & Wortham, 2000) have shown to be helpful for the acquisition of skills needed to solve rather well-defined problems, they lack of a flexible access to the heuristics strategies that underlie the process of solving rather complex problems (e.g. mathematical proof problems). To adapt the traditional worked examples to the needs of solving complex mathematical proof problems, heuristic worked examples have been developed (Reiss & Renkl, 2002) that describe an authentic solution process according to a process model, e.g. Boero's (1999) experts' model, and provide heuristic strategies. A study by Hilbert, Renkl, Kessler, and Reiss (2007) showed a positive effect of learning with heuristic worked examples compared to regular instruction on teacher students' geometry concepts and proof skills. For the study presented in this paper Boero's process model served as basis for the development of the used heuristic worked examples.

Transactive argumentation within the collaborative learning process

There are many types of discourse activities that contribute to individual learning. Chi (2009) differentiates between active, constructive and interactive learning processes. Active processes are observed when something is physically done with information by the learner. Constructive processes are characterized as the production of knowledge beyond the information the learner decodes from the learning material. Finally, interactive processes are characterized as collaborative processes in which the learners take each partner's contribution into account. When learning collaboratively, all three types of learning processes are possible to occur. However, what according to Chi (2009) really makes collaborative learning effective, are interactive processes. Others have called such processes "transactive" (Teasley, 1997; Weinberger & Fischer, 2006). Given the high potential ascribed to the interactive resp. transactive learning processes, in this paper we use the transactivity principle for

learning with CSCL scripts that was stated in the Script Theory of Guidance for Computer-Supported Collaborative Learning (Fischer, et al., 2013). According to this principle, CSCL scripts will be more beneficial for learning the more they induce a transactive learning process, i.e. the more they lead to the learning partners' mutually referring to each others contributions. Recent studies have shown the impact transactive CSCL scripts can have on learning (e.g. Noroozi, Teasley, Biemans, Weinberger, & Mulder, in press). In this study we specifically focus on transactivity of argumentation, i.e. the extent to which learners refer to their partners' contributions in an argumentative way, for example through criticizing or synthesizing the partners' arguments. Through actively building on each other's arguments, optimally a deep elaboration of both the argument and the underlying concepts is achieved. Thus, transactive argumentation should also lead to higher learning success. Furthermore, the repeated use of transactive argumentation during the collaborative learning process should lead to the internalization of a script, when it is designed to help students engage in transactive argumentation, which is the case for the CSCL script used in this study. One open issue is whether an individual's learning is dependent on his/her self-generated transactive argumentation, or whether it is (also) dependent on partner-generated transactive argumentation. It may be argued that, for the acquisition of argumentation skills it is more important for the learner to construct transactive argumentation her-/himself than to be exposed to the learning partner's transactive argumentation, because to actively generate transactive argumentation the learner must deal intensively with the partner's contribution, whereas the learner might not necessarily process the partner-generated transactive argumentation at all. Nevertheless, the learning partners might mutually influence each others' contributions by the transactive argumentation they express.

Research Questions

The research questions this paper tries to answer are:

(RQ1) What are the effects of a CSCL script, heuristic worked examples and the combination of both on students' use of transactive argumentation when collaboratively working on mathematical proof tasks? We expected a positive effect of the CSCL script compared to unscripted collaborative learning on the use of transactive argumentation, since CSCL scripts sequence the learners' collaborative learning activities by inducing advantageous activities (e.g. referring to the learning partner's contribution) at the appropriate point of time during the collaboration process. The heuristic worked examples provide domain-specific content that facilitates learners to construct arguments, especially when the construction of arguments is supported by the CSCL script (Sadler, 2004). Thus, for the heuristic worked examples compared to learning without heuristic worked examples we expected a positive effect as well, and for the combination of both scaffolds we expected to find a synergistic scaffolding effect (Tabak, 2004).

(RQ2a) To what extent are the effects of a CSCL script and heuristic worked examples on knowledge about argumentation mediated by self-generated transactive argumentation? We expected that self-generated transactive argumentation explains (i.e. mediates) the positive effect of the CSCL script on the acquisition of knowledge about argumentation. When learners carry out what a CSCL script suggests, this should lead to a more frequent use of transactive argumentation than in unscripted discussions and – mediated by transactive argumentation – to an internalization of knowledge about argumentation embedded in the script (Fischer, et al., 2013). Also, a mediation of the effect of learning with heuristic worked examples on the acquisition of knowledge about argumentation through self-generated transactive argumentation was expected. The heuristic worked examples provide domain-specific content students could use to repeatedly engage in an argumentative discourse. Again, this is expected to lead to a more frequent use of transactive argumentation and thus to an internalization of the script for argumentation.

(RQ2b) To what extent are the effects of a CSCL script and heuristic worked examples on knowledge about argumentation mediated by partner-generated transactive argumentation? The effects of both scaffolds on the acquisition of knowledge about argumentation might be mediated by the partner-generated transactive argumentation but not to the same extent as they are expected to be mediated by the self-generated transactive argumentation because the learners do not necessarily have to process the learning partner's contribution.

(RQ3) To what extent do the frequencies of transactive argumentation generated by each of the two learning partners reciprocally influence each other? By definition, the generation of transactive argumentation depends on the contributions the learning partner provides within the collaborative learning process. Therefore, we expected a positive relationship between self- and partner-generated transactive argumentation. Further we expected that a positive effect of the partner-generated transactive argumentation on one's own knowledge about argumentation would be mediated by the self-generated transactive argumentation.

Methodology

Participants and design

The study was conducted as part of a two weeks course for prospective math teacher students that were about to start their university education. Out of 162 students participating in the pre-test, 61 students missed more than

one treatment and/or did not show up at the posttest and thus had to be excluded for the purposes of this paper. After clearing for drop-outs, $N = 101$ math teacher students were included in the analyses presented in this paper. A 2x2-factorial experiment with the independent variables CSCL script (with vs. without) and heuristic worked examples (with vs. without) was established. Participants were randomly assigned to one of the four experimental conditions. The study took place on five consecutive days with pre- and post-test data collection on the first and fifth day. On the second through the fourth day the participants were exposed to one treatment session per day lasting 45 minutes. For each of the three treatment sessions, the learners were randomly assigned to new dyads to reduce the effect one specific participant might have on his or her learning partner.

Setting and learning environment

Students learned collaboratively in dyads in a CSCL environment (see Figure 1) on three different mathematical proof tasks (e.g.: “Take an uneven amount of consecutive numbers and add them up. Repeat this and try to find regularities. Formulate a conjecture and prove it.”). The learning partners were each equipped with one laptop and a graphic tablet and worked co-presently on the proof tasks. The laptops of both learning partners were linked to each other to distribute different interconnected prompts and material as well as to display a mirrored workspace where the learners could share their written communication and drawings (see Figure 1). Since the learning partners were allocated face-to-face they were able to speak to each other but they were requested to write their discussion about and progress on the mathematical proof task into the shared work space.

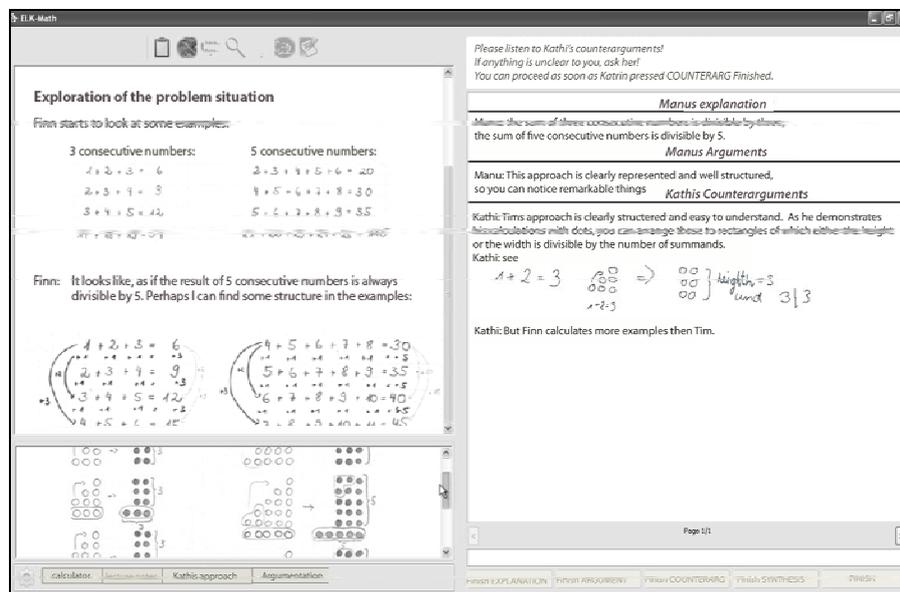


Figure 1. Screenshot of the computer program (left side of the screen: private work space including the problem to be solved resp. the heuristic worked example; right side: shared work space displaying script prompts or not).

The learning environment on the laptop screen was divided into two parts. On the left half of the screen, the learning environment provided the mathematical proof task, a calculator and domain-specific lecture notes (available in all conditions) as well as the heuristic worked example (in the conditions with heuristic worked examples only). On the right half of the screen, the students were able to share text and drawings by using the available text and graphic chat function (available in all conditions). The students had the opportunity to create any number of pages for their written communication and browse through them within the current treatment session. On the upper right side of the screen the script prompts were displayed that aimed at sequencing students' contribution types to the discussion (in the conditions with CSCL script only).

Independent variables

In all experimental conditions, students were requested to work on the proof task alternately individually and collaboratively by discussing their ideas. In the conditions *with CSCL script*, we adapted the script that was investigated by Stegmann et al. (2007) to the context of mathematical argumentation tasks. Thus, the collaborative discussion was sequenced into the three phases (1) argument, (2) counterargument, and (3) synthesis. When prompted to construct counterarguments or synthesize arguments, students were specifically asked to refer to their learning partner's contribution (Leitão, 2000). Also, students were encouraged to formulate sound arguments according to Toulmin's (1958) model of argument construction (including claims,

data and rebuttals). In the conditions *without CSCL script*, students discussed their ideas without receiving any guidance for their discussion.

In the conditions *with heuristic worked examples*, students received worked examples that split a possible solution of one proof task into steps of mathematical proof (adapted from Boero, 1999). The worked examples contained different heuristic strategies which an imaginary student applied to make progress within each of these steps. At each first, third and fourth step within the heuristic worked examples (i.e. at (1) generation of a conjecture, (3) exploration of the mathematical statement, (4) selection of adequate arguments to generate a proof draft), both students were provided with different example versions. After studying these specific steps individually, students were asked to present both versions of the step to each other and to discuss them. In the conditions *without heuristic worked examples*, students had to work on the mathematical proof tasks by problem solving only, i.e. without receiving guidance neither on the steps of mathematical proof nor on heuristic strategies. The students in these conditions were alternately asked to think about their ideas how to solve the problem individually and to present their ideas to each other and discuss them.

Dependent variables

For the pre- and post-test measure of *knowledge about argumentation*, students were requested to describe typical phases and activities they would expect to occur in a discussion about a science topic. A topic different from mathematics was chosen to investigate knowledge about argumentation that could be transferable to different domains. Students answers were coded for the amount of argumentative elements (e.g. pro-argumentation, counter-argumentation, etc.) they named correctly (for further information, e.g., on reliabilities, see Kollar, et al., 2012).

To measure *the frequency of transactive argumentation*, the written communication of the learning partners during the three treatment sessions were coded. As transactive argumentation, all statements were counted that built on partner's contribution in an argumentative way. To code learners' written communication, first the pages created in each of the three treatment sessions were segmented at the points where turn taking occurred. The resulting segments were taken as unit of analysis to code the written communication by coders that were trained as follows. The research assistant responsible for the study trained two student assistants to code the segments regarding transactive argumentation using a coding scheme with descriptions and examples of transactive argumentation (see Table 1). The training was conducted within eight weeks with alternately coding of the training material and discussing coding differences together with the research assistant to reach a more precise coding. The frequencies of segments containing transactive argumentation were then summed up for each learner separately. During the training the inter-rater reliability could be advanced from poor values ($ICC_{unjust} < .40$) to sufficient values ($ICC_{unjust} > .60$). After training, the two coders coded a sample of > 5% of the whole sample of written communication across all conditions and treatment sessions with sufficient inter-rater reliability ($ICC_{unjust} = .68$).

Table 1: Excerpt from the coding scheme with descriptions and examples for transactive argumentation.

Transactive argumentation	
Description	Examples
Criticizing: Comments that tackle the approach to solve the problem or the solution itself and contain counterargumentation and/or criticism directly referring to the learning partner's contribution	- "... but your description of the problem space is less helpful because not every kind of possible solutions can be displayed" - "... $2 + 3 = 5$ " (as counter-example to the claim that the sum of two consecutive numbers is always even)
Synthesizing: Comments that synthesize previous contributions with containing at least one contribution made by the learning partner	- "...the summary of the pros and cons we made is..." - "Taking your criticism into account we could agree on distinguishing between cases when the numbers are even and uneven"

Statistical analyses

To answer our research questions, we used univariate analysis of variance (to test the effects of the two treatments on the frequency of transactive argumentation) and linear regressions (to determine to what extent self- and partner-generated transactive argumentation would be a predictor for the knowledge about argumentation displayed in the posttest). To confirm the significance of mediating predictors in the linear regression models, we calculated Sobel tests (Sobel, 1982). For all tests the significance level was set to $\alpha = .05$. As measures of effect sizes, partial η^2 were used, with values between .01 and .05 being considered as weak effects, values between .06 and .14 as medium effects, and values of .14 and higher as large effects (Cohen, 1988).

Results

(RQ1) An ANOVA revealed a positive effect of learning with the CSCL script on the frequency of transactive argumentation during the collaboration process compared to learning without the CSCL script ($F(1,97) = 11.63$, $p = .001$, partial $\eta^2 = .11$). Also the effect of learning with heuristic worked examples on the frequency of transactive argumentation compared to learning without heuristic worked examples was positive ($F(1,97) = 28.41$, $p < .001$, partial $\eta^2 = .23$). The interaction effect was significant, ($F(1,97) = 18.24$, $p < .001$, partial $\eta^2 = .16$). Post-hoc comparisons of the four experimental groups showed that only the learners supported with both forms of instructional support at once achieved significantly higher frequencies of transactive argumentation than learners in the other three groups ($F(1,97) = 57.77$, $p < .001$, partial $\eta^2 = .37$).

(RQ2a) Learning with the CSCL script and learning with the heuristic worked examples positively predicted the acquisition of knowledge about argumentation measured between pre- and post-test (model 1; see Table 2 for exact β -values in the linear regression models; see also Kollar, et al., 2012). Both positive predictions disappeared when the frequency of self-generated transactive argumentation was integrated into the regression model (model 2a), while the frequency of self-generated transactive argumentation predicted significantly the acquisition of knowledge about argumentation. Sobel tests showed that self-generated transactive argumentation significantly mediated the effect of the CSCL script ($z = 2.26$, $p = .01$, one-tailed) and the effect of the heuristic worked examples ($z = 2.32$, $p = .01$, one-tailed) on students' acquisition of knowledge about argumentation.

(RQ2b) When the frequency of partner-generated transactive argumentation was included into the initial linear regression model, it did not serve as a significant predictor for students' development of knowledge about argumentation, and the CSCL script still positively predicted students' development of knowledge about argumentation significantly (model 2b). The Sobel test could also not confirm partner-generated transactive argumentation as mediator for the effect of the CSCL script on the acquisition of knowledge about argumentation ($z < 1$, ns). In contrast, heuristic worked example were no longer a significant predictor of knowledge about argumentation when partner-generated transactive argumentation was included into the initial linear regression model (model 2b). However, Sobel tests did not confirm that the effect of heuristic worked examples on knowledge about argumentation was mediated by the partner-generated transactive argumentation as the reduction of the β -value of heuristic worked examples between the model was not substantial ($z < 1$, ns).

Table 2: Summary of multiple regression models with predictors for knowledge about argumentation.

Variable	B	SE B	β
Model 1			
CSCL Script (C)	0.902	0.355	.332**
Heuristic Worked Examples (H)	0.664	0.375	.245*
C X H	-0.258	0.516	-.083
Model 2a			
CSCL Script (C)	0.416	0.398	.153
Heuristic Worked Examples (H)	0.058	0.440	.021
C X H	0.281	0.549	.091
Self-generated transactive argumentation per treatment session	0.549	0.222	.290**
Model 2b			
CSCL Script (C)	0.845	0.400	.311*
Heuristic Worked Examples (H)	0.598	0.432	.220
C X H	-0.199	0.553	-.064
Partner-generated transactive argumentation per treatment session	0.055	0.177	.036

* $< .05$, ** $< .01$, one-tailed

(RQ3) For the analysis of the extent learners of a dyad might have mutually influenced each other in their generation of transactive argumentation, a linear regression model revealed a significant positive relationship between the frequencies of the partner-generated and the self-generated transactive argumentation (stand. $\beta = .655$, $p < .001$). Further, the frequency of partner-generated transactive argumentation positively predicted the acquisition of knowledge about argumentation measured between pre- and post-test (model 3, see table 3 for exact β -values). The positive prediction for partner-generated transactive argumentation disappeared when the frequency of self-generated transactive argumentation was integrated into the regression model (model 4), while the frequency of self-generated transactive argumentation predicted significantly the acquisition of knowledge about argumentation. Sobel tests showed that self-generated transactive argumentation significantly mediated the positive effect of the frequency of partner-generated transactive argumentation ($z = 3.12$, $p = .002$, one-tailed).

Table 3: Summary of multiple regression models with predictors for knowledge about argumentation.

Variable	B	SE B	β
Model 3			
Partner-generated transactive argumentation per treatment session	0.304	0.152	.198*
Model 4			
Partner-generated transactive argumentation per treatment session	-0.113	0.191	-.073
Self-generated transactive argumentation per treatment session	0.783	0.235	.414**

* $p < .05$, ** $p < .01$, one-tailed

Conclusion and Discussion

The interaction effect between the CSCL script and the heuristic worked example indicates that both types of instructional support applied together produce synergistic effects (Tabak, 2004) on the use of transactive argumentation within the collaborative learning processes. Thus, the CSCL script and the heuristic worked examples amplified each other's effects on the collaborative process i.e. the effectiveness of the CSCL script was increased by the heuristic worked examples. Thus, when students were supported with heuristic worked examples, the CSCL script could induce transactive argumentation more effectively by guiding students through a sequence of argumentative discourse moves containing arguments, counterarguments and syntheses. This might have been caused by the richer content the heuristic worked examples provided on which students were better able to apply the prompts of the CSCL script that specifically aimed for a high transactivity during argumentation (Sadler, 2004). Further, the results underpin the importance of transactivity in collaborative learning processes (Fischer et al., 2013, Teasley, 1997) for individual skill acquisition. The self-generated transactive argumentation induced by the script and the heuristic worked examples significantly mediated the positive effects of both means of instructional support on students' advances in their knowledge about argumentation. Thus, it can be recommended to carefully design instructional interventions to foster students' knowledge about argumentation by focussing on ways that are likely to induce transactive argumentation. Interestingly, only self-generated transactive argumentation, but not the transactive argumentation generated by the learning partner was influential for the acquisition of knowledge about argumentation. Thus, for the acquisition of one's own knowledge about argumentation it might be more important that learners generate transactive argumentation by themselves than to be exposed to a learning partner who is generating transactive argumentation (Teasley, 1997). This makes sense, as learners have to be engaged in the partner's contribution when generating transactive argumentation while it is not necessary for them to process as deeply with the transactive argumentation which is generated by their learning partners. Nevertheless, the learning partner is still important for a beneficial learning process, since transactive argumentation needs the learning partner's contributions to refer to them. As the regression analyses with respect to RQ3 show, both learning partners mutually influenced their generation of transactive argumentation within the learning process. The importance of the learning partner is supported by results showing that both self-generated and partner-generated transactive argumentations are substantially positively related. Further, an indirect effect of partner-generated transactive argumentation on one's own development of knowledge about argumentation was found to be fully mediated by self-generated transactive argumentation. This means that it is not enough to just be exposed to transactive arguments but these arguments have to be transactively processed by the learners. Thus, as a further theoretical conclusion, the transactivity principle in the script theory of guidance (Fischer et al., 2013) has to be differentiated for the mathematical context that was used in this study. For the design of CSCL scripts to foster argumentation skills in mathematical context, it can be suggested that it should aim to induce transactive argumentation for each of the learning partners.

References

- Aberdein, A. (2009). Mathematics and argumentation. *Foundation of Science*, 14(1-2), 1-8.
- Atkinson, R.K., Derry, S.J., Renkl, A. & Wortham, D. (2000). Learning from Examples: Instructional Principles from the Worked Examples Research. *Review of Educational Research*, 70(2), 181-214.
- Boero, P. (1999). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof*, July/August 1999. <http://www.lettredelapreuve.it/>.
- Chi, M. T. H. (2009). Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science*, 1(1), 73-105.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- De Wever, B., van Keer, H., Schellens, T. & Valcke, M. (2010). Roles as a structuring tool in online discussion groups: The differential impact of different roles on social knowledge construction. *Computers in Human Behavior*, 26(4), 516-523.

- Ertl, B., Kopp, B., & Mandl, H. (2006). Fostering collaborative knowledge construction in case-based learning in videoconferencing. *Journal of Educational Computing Research*, 35(4), 377–397.
- Fischer, F., Kollar, I., Stegmann, K. & Wecker, C. (2013). Toward a Script Theory of Guidance in Computer-Supported Collaborative Learning. *Educational Psychologist*, 58(1), 56-66.
- Heinze, A., Reiss, K., & Rudolph, F. (2005). Mathematics achievement and interest from a differential perspective. *Zentralblatt für Didaktik der Mathematik*, 37(3), 212-220.
- Hilbert, T. S., Renkl, A., Kessler, S., & Reiss, K. (2008). Learning to prove in geometry: Learning from heuristic examples and how it can be supported. *Learning and Instruction*, 18(1), 54–65.
- Hron, A., Hesse, F. W., Cress, U., & Giovis, C. (2000). Implicit and explicit dialogue structuring in virtual learning groups. *British Journal of Educational Psychology*, 70, 53–64.
- King, A. (2007). Scripting collaborative learning processes: A cognitive perspective. In F. Fischer, I. Kollar, H. Mandl & J. M. Haake (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational, and educational perspectives* (pp. 13-37). New York: Springer.
- Kollar, I., Fischer, F., & Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learning and Instruction*, 17(6), 708–721.
- Kollar, I., Ufer, S., Lorenz, E., Vogel, F., Reiss, K., & Fischer, F. (2012). Using heuristic worked examples and collaboration scripts to help learners acquire mathematical argumentation skills. 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 1, Full Papers* (pp. 331–338). Sydney, Australia: ISLS.
- Leitão, S. (2000). The potential of argument in knowledge building. *Human Development*, 43, 332–360.
- Noroozi, O., Teasley, S. D., Biemans, H. J. A., Weinberger, A., & Mulder, M. (in press). Facilitating learning in multidisciplinary groups with transactive CSCL scripts. *International Journal of Computer-Supported Collaborative Learning*.
- Noroozi, O., Weinberger, A., Biemans, H. J.A., Mulder, M., & Chizari, M., (2012). Argumentation-Based Computer Supported Collaborative Learning (ABCSCCL): A synthesis of 15 years of research. *Educational Research Review*, 7(2), 79-106.
- Reiss, K. & Renkl, A. (2002). Learning to prove: The idea of heuristic examples. *Zentralblatt für Didaktik der Mathematik*, 34(1), 29-35.
- Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*, 41(5), 513–536.
- Schwarz, B. B., & Linchevski, L. (2007). The role of task design and argumentation in cognitive development during peer interaction: The case of proportional reasoning. *Learning and Instruction*, 17(5), 510–531.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13, 290–312.
- Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, 2(4), 421-447.
- Tabak, I. (2004). Synergy: A Complement to Emerging Patterns of Distributed Scaffolding. *The Journal of the Learning Sciences*, 13(3), 305–335.
- Teasley, S. D. (1997). Talking about reasoning: How important is the peer in peer collaborations? In C. O'Malley (Ed.), *Discourse, Tools, and Reasoning: Situated Cognition and Technologically Supported Environments* (pp. 361–384). Berlin: Springer.
- Toulmin, S. E. (1958). *The uses of argument*. Cambridge: Cambridge University Press.
- Van Eemeren, F. H. & Grootendorst, R. (2004). *A Systematic Theory of Argumentation: The pragma-dialectical approach*. Cambridge: Cambridge University Press.
- Vogel, F., Kollar, I., & Fischer, F. (2012). Effects of computer-supported collaboration scripts on domain-specific and domain-general learning outcomes: A meta-analysis. 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* (pp. 446–450). Sydney, Australia: ISLS.
- Weinberger, A., & Fischer, F. (2006). A Framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers and Education*, 46(1), 71–95.
- Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior*. 26, 506-515.

Acknowledgments

The research reported here was funded by the Deutsche Forschungsgemeinschaft (DFG).

Relationships between Listening and Speaking in Online Discussions: An Empirical Investigation

Alyssa Friend Wise, Simone Nicole Hausknecht, Yuting Zhao
Simon Fraser University, 250-13450 102 Avenue, Surrey BC, Canada, V3T 0A3
Email: afw3@sfu.ca, shauskne@sfu.ca, yza174@sfu.ca

Abstract: This study investigated relationships between how students “listen” (access existing posts) and “speak” (contribute posts) in asynchronous online discussions. Ten variables indexing four dimensions of students’ listening (breadth, depth, temporal contiguity and revisitation) and five variables indexing three dimensions of students’ speaking (discursiveness, depth of content and reflectivity) were calculated for 31 students participating in six week-long online discussions as part of an undergraduate educational psychology course. Multi-level mixed-model linear regressions indicated that responsiveness of students’ posts was positively predicted by how often they revisited previously read peer posts, and negatively related to a greater number of posts in the discussion overall. Post content quality was predicted by the percentage of posts viewed that students actually read (as opposed to scan). Put together, results suggest that when students take the time to read and re-read their peers’ posts there are related benefits in the quality of the posts they contribute.

Introduction

Asynchronous discussions are often seen as a powerful venue for knowledge construction due to their affordances for thoughtful commentary and reflective responses (Lipponen, 2002). The core premise is that learners build their ideas collectively and individually through dialogue; thus well-designed and supported online discussions can contribute to learning. Various mechanisms have been proposed to explain such learning including articulating one’s ideas, receiving feedback on these, socio-cognitive conflict caused by exposure to divergent views, the taking of multiple perspectives into account, and the internalization of collaborative activity (Stahl 2005; Lipponen, 2002). In common, all depend on two basic interrelated processes that learners must engage in: “speaking” (contributing posts to the discussion); and “listening” (accessing existing posts) [Wise et al., 2013]. When learning discussions are truly collaborative, these two activities are intimately related and inform each other. In contrast, if learners do not attend to others’ posts (or do so in an incoherent way) the “discussion” that results is more akin to a series of parallel monologues, explaining the findings of shallow and disjointed conversations noted by various researchers (e.g. Thomas, 2002; Webb et al 2004).

Previous research has informed our understanding about the ways in which learners engage in the activities of speaking and listening in online discussions (e.g. Pena-Shaff & Nicholls, 2004; Hewitt, 2003; Ho & Swan, 2007; Wise et al., 2013) but heretofore not connected the two. This is important because the interrelationship (or lack thereof) between the activities of speaking and listening may be an important factor contributing to the extent to which asynchronous discussions live up to or fall short of their possibilities for supporting knowledge construction. In this study we bridge this gap by examining how students’ listening and speaking behaviors relate to each other. This work connects to the conference theme of learning across levels as the interdependent processes of individuals within the context of groups are examined over time.

Theoretical Framework

While the metaphorical language of speaking and listening refers to real-time spoken conversations; in online discussions these activities take a different form. In “listening” learners engage with the text-based expressions of others’ ideas at their discretion and on their own timeline (Jonassen & Kwon, 2001); in “speaking” they exercise decisive control over the timeline for composing their thoughts, and when and where in the conversation they contribute. This temporal flexibility and decoupling of participation timelines changes the dynamics of discussions, allowing time for reflection and the opportunity to revisit comments already “heard,” but also creating challenges for managing discussions that proliferate in one’s absence (Lipponen, 2002; Peters & Hewitt, 2010). New categories with which to characterize these behaviors also emerge.

Conceptualizing Dimensions of “Listening” in Online Discussions

Previous work (Wise et al., 2012a; 2012c; 2013) has conceptualized different dimensions to characterize the ways in which students attend to the posts of others and has explored different approaches students take in such online listening. At a basic level, students can differ in the breadth and depth with which they view their classmate’s contributions. The *breadth* with which students attend to others’ posts is important in terms of the diversity of ideas that they are exposed to and their ability to respond to the discussion as a whole, while the *depth* with which they attend to these posts is an indication of the degree to which they are considering others’

ideas. For example some students “cover” the conversation by opening a large percentage of their classmate’s posts but spending little time on them, while others attend to only a portion of discussion posts but focus on them deeply (Wise et al., 2012a). Because students are in control of their own timelines of participation (Jonassen & Kwon, 2001) it is also necessary to consider the *temporal contiguity* of students’ listening; for example the degree to which they disperse or concentrate their participation. Finally, online discussions permit *revisitation* activities in which students can choose to return to posts (made by themselves and others) that they have attended to previously. Students’ bias towards reading new posts (Hewitt, 2003), however, may limit the frequency with which such revisitation actually occurs. Generally richer listening behaviors along each of these dimensions (greater breadth, depth, contiguity and revisitation) are thought to be more desirable for interactive dialogue in online discussions; however, such connections need to be examined empirically. To theorize and test potential relationships with specificity we must first conceptualize qualities of speaking along several dimensions as well.

Conceptualizing Aspects of “Speaking” Quality in Online Discussions

The different characteristics, functions and qualities of posts in online discussions have been theorized by many researchers (for a selected overview see reviews by De Wever et al, 2006 and Hew et al., 2010). Examined together, three common dimensions can be seen as important in almost all models: discursiveness (that learners’ comments refer to each other in meaningful ways); content (that the learning material is thoughtfully considered); and reflectivity (that the learning process itself is taken as an object for examination). Each of these dimensions is expanded on below.

First, for discussions to function as interactive dialogues rather than a series of parallel monologues (Boulos & Wheeler, 2007), posts need to contain *discursive* elements through which participants link their comments to each other. These elements can be responsive (e.g. expressions of social support, proposing consensus) or elicitive (e.g. asking questions). Responsiveness itself can take many forms; at a basic level a simple act such as acknowledging others may create the social support required for individuals to build trust to take risks within a discussion (Cheung et al., 2008). At a deeper level, when students respond to the ideas in a post they may expand or challenge that student’s (and others’) existing thinking, and when they respond to multiple ideas synthetically they can initiate a process of developing collective understanding (Gunawardena, et al., 1997). Similarly by eliciting responses from others, students contribute to the interactivity of the dialogue.

In addition to discursiveness, the depth with which academic *content* is discussed is central to the learning. A common way to assess this across multiple discussion topics draws on the argumentation literature and looks at the degree to which students make claims, and use reasoning, evidence, and theory to support them (e.g. Lin et al. 2012; Weinberger & Fischer, 2006). The underlying notion is that richer argumentation structures (more content-related claims and the greater use of supporting evidence, and theory) indicate deeper consideration of the learning material. Finally, the opportunity for *reflectivity* has been cited as a particular advantage of asynchronous online discussion since time-unlimited review of earlier parts of the discussion is possible (Harasim, 2000; Knowlton, 2005). Within a discussion, a student may consider the process of the group’s knowledge construction (Knowlton, 2005), but also the development of their own ideas on a topic. Thus target of reflection may be the learning process at either the group or individual level.

Put together, these five elements (responsiveness, elicitation, argumentation, individual reflection, and group reflection) across three dimensions (discursiveness, content, and reflectivity) provide a useful framework with which to examine the contributions a post makes to a discussion. In the next section we describe theoretically predicted relationships between these aspects of speaking quality and the dimensions of listening described above.

Connecting Speaking and Listening Activity

Theoretically, speaking and listening are intimately interrelated activities in the process of constructing knowledge though online discussions; however such connections have not yet been examined empirically. In this section we explicate the logic by which we would conceptually expect them to relate using the dimensions of speaking and listening outlined previously. Considering *breadth* of listening, as students attended to a greater proportion of their peer’s posts, we would expect them to be more *discursive* in their own comments, responding to and eliciting ideas from them. In addition as they become aware of a greater number of perspectives and views on the discussion topic, they are also likely to create posts with more sophisticated *argumentation* that supports, and perhaps qualifies, their position with respect to these other views. *Depth* of listening would also be expected to support *discursiveness* and *argumentation* as a richer understanding of peers’ ideas would support more thoughtful responses and questions, as well as lead to richer content as students carefully support or qualify their ideas based on this understanding. Turning to *revisitation*, rereading of already viewed peer posts suggests additional consideration of the ideas contributed by others, and thus would be expected to further support *discursiveness* and *argumentation* in the ways described above. Returning to ideas considered (or contributed) previously also can support the process of *reflection* on both group and individual

learning processes. *Temporal dispersion* may also support *reflection*; if students distribute their participation over a greater period of time and number of sessions they may be more likely to notice changes in their own and others' views. In contrast the *temporal contiguity* of conducting listening and speaking actions in the same session may be needed as a foundation for relationships between the two activities to be established.

Research Questions

1. What listening behaviors are associated with the discursiveness of a student's post in terms of responsiveness and elicitation?
2. What listening behaviors are associated with the depth of content of a student's post in terms of argumentation?
3. What listening behaviors are associated with the reflectivity of a student's post in terms of individual and group reflection?

Methods

Learning Environment and Participants

Students in a fully online undergraduate course on educational psychology participated in six week-long small-group discussions with 8-10 classmates. Discussions were conducted in three two-week sets (weeks 3/4, 8/9 and 11/12); the instructor gave students both group and individual feedback after the first two discussions (worth 5% of the course grade) and the latter four (worth 20%). Each week, students were asked to discuss contrasting perspectives on an authentic educational controversy and come to a collective position with rationale. Students were required to contribute at least two posts per topic and given guidelines for expectations of post quality in line with the dimensions of discursiveness, content, and reflectivity discussed above. Thirty-one of 52 students enrolled in the course consented to have data on their discussion participation collected for this study.

Data Extraction and Variable Calculations

Listening Variables

Clickstream (log-file) data was collected on all actions students took in the system to assess listening activity; action types were "view" (opening others' posts), "post" (creating a post), or "review" (revisiting previously read posts). Times between subsequent actions were subtracted to calculate duration, actions were divided into sessions-of-use, and views were subcategorized as scans or reads based on a maximum reading speed of 6.5 words per second (wps) [see Hewitt et al., 2007]. Ten variables were calculated for the different listening dimensions (see Table 1).

Table 1. Summary of ten listening variables along four dimensions

Dimension	Variable	Definition
Breadth	Percentage of others' posts viewed	# of unique posts made by others that a student viewed* divided by the total # of posts made by others
	Percentage of others' posts read	# of unique posts made by others that a student read* divided by the total # of posts made by others
Depth	Percentage of real reads	# of times a student read others' posts divided by their total # of views
	Average length of real reads (min)	Total time a student spent reading posts, divided by the number of reads
Temporal Contiguity	Number of sessions	# of times a student logged-in to the discussion
	Percentage of sessions with posts	# of sessions in which a student made a post, divided by their total # of sessions
	Participation range (days)	# of days between when a student first and last logged-in
Revisitation	Reviews of own posts	# of times a student reread posts they made
	Reviews of instructors' posts	# of times a student reread posts made by the instructor
	Reviews of other' posts	# of times a student reread posts made by others they had viewed previously

*Views include all accessing of others posts. Reads include only posts viewed slower than 6.5 wps

Speaking Variables

All 479 posts made by participants were extracted from the discussion tool and coded by two researchers for the five speaking variables (along three dimensions) described previously to evaluate post quality. The post was used as the unit of analysis for both theoretical and practical reasons as this was the unit through which students expressed their ideas in interaction with others and it presented an unambiguous basis for segmentation. The discursive dimension was assessed through the degree of students' Responsiveness to others' posts [$\kappa = 0.71$] and Elicitation that invited responses from others [$\kappa = 0.91$] to capture the degree to which a posts attempted to connect to preceding and subsequent discussion. The content dimension was assessed as richness of Argumentation [$\kappa = 0.74$], which captures the depth with which the academic content was considered. Finally the reflectivity dimension was evaluated through the degree to which students exhibited Reflection on their Individual Learning Process [$\kappa = 0.83$] and Reflection on the Group's Learning Process [$\kappa = 0.75$] to assess the degree to which students considered the process of knowledge construction either as an individual or for the group. Coding was based on a combination and adaptation of prior schemes and models by Hara et al. (2000), Knowlton (2005), Pena-Shaff & Nicholls (2004), Weinberger & Fischer (2006), and Wise et al. (2012d); see Table 2 for an overview of the scheme used.

Table 2. Overview of coding scheme for speaking variables

Discursiveness	
<i>Responsiveness</i>	<i>Elicitation</i>
0 None	0 None
1 Acknowledging	1 Questions not clearly directed to anyone
2 Responding to an idea	2 Questions directed to one person
3 Responding to multiple ideas	3 Questions directed to the group
Content	
<i>Argumentation</i>	
0 No argumentation	
1 Unsupported argumentation (Position only)	
2 Simple argumentation (Position + Reasoning)	
3 Complex argumentation (Position + Reasoning + Qualifier/preemptive rebuttal)	
Reflectivity	
<i>Reflection on Individual Learning Process</i>	<i>Reflection on Group Learning Process</i>
0 No individual reflection	0 No group reflection
1 Shallow individual reflection	1 Shallow group reflection
2 Deep individual reflection	2 Deep group reflection

Statistical Analysis

Multi-level mixed-model linear regressions for each speaking variable on predicted relevant listening variables were conducted to examine relationships. Because students' discussion behaviors may change across a series of discussions, aggregating data across the entire semester could obscure relationships between listening and speaking behaviors. Thus, models were based on variable averages calculated for each discussion week, the unit of activity in the course. For each model, the explanatory variables of interest were included as fixed effects (see Table 3) while effects of group-membership, discussion-week, group-by-week interactions (operationalized as # of posts per group), students-nested-within-groups and student-by-week interactions (operationalized as # of posts per student) were included as random effects. Backwards elimination was used to iteratively remove explanatory variables and refit equations until all remaining variables had $p < .10$. The two post-count variables remained in the model regardless of their significance. The alpha level used for interpretation was .05.

Results

Summary Statistics

There was great diversity in listening and speaking behaviors in the discussions. Although all students logged in the forum at least once each discussion set, some engaged in minimal participation with no posting and little attention to the posts of others, while others logged-in multiple times and read every post in the discussion (see Table 4). The number of posts in each discussion ranged from 13 to 52. The average level of responsiveness was at the mid-point of the scale, while elicitation was low and argumentation was high, though all varied substantially. Reflection on both individual learning and group processes was consistently low.

Table 3. Listening Variables Included in Regression of Each Speaking Variable

<i>Listening Dimensions</i>	Speaking Variables			Individual Reflection	Group Reflection
	Responsiveness	Elicitation	Argumentation		
<i>Breadth</i>	% of others' posts viewed	% of others' posts viewed	% of others' posts viewed		
	% of others' posts read	% of others' posts read	% of others' posts read		
<i>Depth</i>	% of real reads	% of real reads	% of real reads		
	Av. length of real reads	Av. length of real reads	Av. length of real reads		
<i>Temporal Contiguity</i>				# of sessions	# of sessions
	% of sessions with posts	% of sessions with posts	% of sessions with posts	Participation range	Participation range
<i>Revisitation</i>	# of reviews of:	# of reviews of:			
	-other students' posts	-other students' posts	-other students' posts	-own posts -instructors' posts -other students' posts	-own posts -instructors' posts -other students' posts

Table 4. Summary Statistics for Data Aggregated by Student and Discussion

Variable	Mean	S. D.	Min	Max
Speaking Quality Variables				
Responsiveness	1.51	0.78	0.00	3.00
Elicitation	0.52	0.72	0.00	3.00
Argumentation	2.17	0.81	0.00	3.00
Reflection on Individual Learning	0.27	0.34	0.00	1.67
Reflection on Group Process	0.33	0.33	0.00	1.50
Speaking Quantity Variables				
# of posts made (by group)	29.60	8.46	13	52
# of posts made (by student)	2.57	1.60	0	10
Listening Variables				
Percentage of others' posts viewed	0.72	0.31	0.00	1.00
Percentage of others' posts read	0.50	0.28	0.00	1.00
Percentage of real reads (not scans)	0.44	0.21	0.00	1.00
Average length of real reads (in min)	3.85	3.21	0.00	17.35
Number of sessions	6.96	5.23	0	29
Percentage of sessions with posts	0.40	0.26	0.00	1.00
Participation range (days)	4.08	1.87	0	7
Number of reviews of own posts	2.56	3.21	0	18
Number of reviews of instructors' posts	10.30	11.23	0	93
Number of reviews of other students' posts	10.67	11.17	0	55

Multi-level Regressions

Modeling results indicated that the responsiveness of students' posts was related both to the number of reviews of other students' posts and the total number of posts made by the group in a particular discussion week. Number of reviews of other students' posts was a positive predictor (greater reviewing of others' posts in a discussion week was associated with the making of more responsive posts) while the total number of posts made by the group was a negative predictor (a greater number of posts made by a group in a week was associated with lower average responsiveness in group members' posts). The level of elicitation in students' posts was also predicted by the number of reviews of other students' posts; however in this case the relationship was negative (more elicitative posts by a student in a discussion week was associated with less reviewing of others' posts). Richness of argumentation was predicted only by the percentage of posts viewed that students actually read (as opposed to scanned). This relationship was positive (a greater percentage of reading in a discussion week was associated with richer argumentation in the posts made). Neither individual nor group reflection was predicted by any of the listening variables.

Table 5. Summary of fixed effects standardized regression coefficients for models of five speaking variables

		Estimate	Standard Error	t value
<i>Responsiveness</i>	# of Posts per group	-0.018	0.009	-2.06*
	# of Posts per student	0.021	0.031	0.68
	Reviews of other' posts	0.013	0.005	2.50*
<i>Elicitation</i>	# of Posts per group	-0.001	0.007	-0.19
	# of Posts per student	0.047	0.035	1.34
	Reviews of other' posts	-0.016	0.006	-2.65*
<i>Argumentation</i>	# of Posts per group	-0.003	0.009	-0.33
	# of Posts per student	-0.041	0.024	-1.71
	Percentage of real reads	0.522	0.257	2.03*
<i>Individual Reflection</i>	# of Posts per group	-0.000	0.004	-0.10
	# of Posts per student	-0.024	0.017	-1.40
	Reviews of other' posts	0.005	0.003	1.66
<i>Group Reflection</i>	# of Posts per group	-0.003	0.005	-0.51
	# of Posts per student	-0.017	0.015	-1.17

* $p < .05$

Discussion

The major finding of this study is the relationship between listening (in terms of depth and revisitation of others posts) and quality of speaking (in terms of discursiveness and argumentative quality). While a connection between listening behaviors and speaking quality has been proposed theoretically (Wise et al., 2012a; 2013) and is implicit in much research on online discussions, this is the first work we are aware of that provides direct empirical evidence to support the connection. Below we discuss the specific relationships found, contextualizing them in the larger framework of prior research on online discussions.

Discursiveness is an important dimension of speaking in online discussions because it is what links individual comments together as a dialogue. Responsiveness can vary from simply social acknowledgements to building on or challenging individual ideas to synthetically integrating multiple perspectives (Gunawardena et al., 1997). The positive relationship found between revisiting others' posts and responsiveness suggests that the richer end of this spectrum tends to occur when posts are attended to multiple times. Examples of such behavior have been found in previous research on online discussion we conducted using a microanalytic case-study approach. In one study (Wise et al., 2012a) we found that a student characterized as interactive in her discussion participation always spent substantial time reading and re-reading others' posts before making her own, highly responsive, posts. In another extreme example (Wise et al., 2012c) a student who often built on others posts and synthesized the group discussion always located her post as a reply to a post she had viewed at least three times already.

Put together, this research suggests an important role for reviewing previously read posts in effective discussion participation. It is reasonable that students may need to read others' posts multiple times to make sense of them in the context of the discussion before being able to respond to the ideas with a complex and thoughtful response. However, prior research has documented students' tendency to do just the opposite; that is focus on only new posts (Hewitt, 2003). Recent work that has attempted to address this problem of new post

bias through the design of a discussion forum interface that encourages students to read and re-read posts in a connected fashion (Marbouti, 2012) may thus prove particularly valuable.

In contrast to the positive relationship found between revisitation and responsiveness, a negative association was found between revisitation and elicitation. This can be interpreted in several different ways. It is possible that rereading previously viewed peer posts helps students clarify some of the questions or doubts they had when they viewed those posts the first time, leading them to ask fewer questions. However, elicitation relates not only to clarification but also raising wonderings to the group. Thus another possible interpretation is that when learners repeatedly set questions to the group, they were more likely to focus their energies on the new responses to these, rather than posts they had read previously. Finally, it is important to note that overall levels of elicitation in the discussions studied were low; thus it is possible that this finding is due to the actions of a small sub-set of the population and of limited generalizability. This is clearly an area that requires further investigation.

The final relationship found for discursiveness was that responsiveness was negatively predicted by a greater amount of posts in the overall discussion. This is consistent with previous findings that a large amount of posts in a discussion lead students to feel overwhelmed (Peters & Hewitt, 2010) and suggests that it may be beneficial to make groups small, thus keeping discussions at a manageable size which allows students to be responsive as part of an interactive dialogue.

Considering argumentation, previous work has questioned whether it is breadth of listening, depth of listening, or a combination of the two that is important to support the richness of post content (Wise et al., 2013). The finding here of a relationship between the percent of posts read (not scanned) and richness of argumentation clearly indicates depth as the more relevant dimension. This aligns with the finding of a relationship between rich responsiveness and post revisitation since returning to a previously viewed post to consider it again could be considered as deep listening in conceptual sense. Logically it makes sense that deep attention to peers' posts can support a richer understanding of meaning and thus stronger argumentation as this understanding drives students to consider and support or qualify their own ideas more deeply. This may help explain part of the mechanism by which conscientious design of online discussion forums can encourage rich argumentation (Lin et al, 2012). In combination with the lack of findings for listening breadth it also provides empirical evidence to support our previous assertion that listening deeply to some of a discussion may be preferable to listening shallowly to all of it (Wise et al., 2013).

Unfortunately, here again research shows us that students tend to do the opposite of what is beneficial, focusing on breadth rather than depth. For example in one of the prior set of cases studies mentioned above, we found evidence of two students who viewed almost all the posts in their discussion, but without drawing on their ideas in their own posts (Wise et al., 2012a). On deeper inspection it seemed that their purpose for viewing posts differed, for one student it appeared to be an effort to "cover" the discussion content, while for the other it was to be able to acknowledge others' posts in a social manner. Additionally, in a larger cluster analysis of student listening behaviors we found that the apparent difference in breadth between two clusters of students was somewhat dissipated by a follow-up analysis revealing that while the "broad" listeners did view all the posts, they spent the majority of their listening efforts in concentrated activity examining a smaller number of posts in depth (Wise et al., 2013). These findings, along with others indicating that students often use widespread scanning as a strategy for coping with high-volume discussions (Wise et al., 2012b; Peters & Hewitt 2010), suggest that students do not instinctively know how to listen effectively in online discussion, and that it is important to provide them with explicit instruction about how to do so.

Conclusion and Study Significance

This study provides empirical evidence to support a relationship between listening behaviors (depth and revisitation) and quality of speaking (discursiveness and depth of content) in online discussions. This is an important area for research because speaking and listening are two interrelated aspects of participating in an online discussion, but previous research has not examined this connection. As shown in this study, patterns in listening can help explain and predict patterns in speaking. Specifically, when students take the time to read and re-read some number of their peers' posts, there are related benefits in the quality of the posts they contribute. The connection between these listening behaviors and post qualities is particularly important given past studies suggesting weak student listening behaviors (Thomas, 2002) and tendencies to focus on reading only new posts or using scanning as a strategy for coping with high-volume discussions (Wise et al., 2012b; Peters & Hewitt 2010). In addition, while the bulk of guidance for students' participation in online discussions focuses on how to post, understanding what listening behaviors are associated with what speaking ones suggests new ways to support students in effective discussion participation. Future work will test the efficacy of providing students with listening guidance to support their speaking quality and expand this work to examine listening/speaking relationships in other kinds of discussion contexts.

References

- Boulos, M. N., & Wheeler, S. (2007). The emerging web 2.0 social software: An enabling suite of sociable technologies in health and health care education. *Health Information and Libraries Journal*, 24(1), 2-23.
- Cheung, W. S., Hew, K. F., & Ling Ng, C. S. (2008). Toward an understanding of why students contribute in asynchronous online discussions. *Journal of Educational Computing Research*, 38(1), 29-50.
- De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, 46(1), 6-28.
- 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.
- Hara, N., Bonk, C. J. & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional Science*, 28(2), 115-152.
- Harasim, L. (2000). Shift happens: Online education as a new paradigm in learning. *The Internet and Higher Education*, 3(1/2), 41-61.
- Hew, K., Cheung, W. & Ng, C. (2010). Student contribution in asynchronous online discussion: A review of the research and empirical exploration. *Instructional Science*, 38(6), 571-606.
- Hewitt, J. (2003). How habitual online practices affect the development of asynchronous discussion threads. *Journal Educational Computing Research*, 28(1), 31-45.
- Hewitt, J., Brett, C., & Peters, V. (2007). Scan rate: A new metric for the analysis of reading behaviors in asynchronous computer conferencing environments. *American Journal of Distance Education*, 21(4), 215-231.
- Ho, C.H. & Swan, K. (2007). Evaluating online conversation in an asynchronous learning environment: An application of Grice's cooperative principle. *Internet and Higher Education*, 10(1), 3-14.
- Jonassen, D. H., & Kwon, H. (2001). Communication patterns in computer mediated versus face-to-face group problem solving. *Educational Technology Research and Development*, 49(1), 35-51.
- Knowlton, D. S. (2005). A taxonomy of learning through asynchronous discussion. *Journal of Interactive Learning Research*, 16(2), 155-177.
- Lin, H., Hong, Z., & Lawrenz, F. (2012). Promoting and scaffolding argumentation through reflective asynchronous discussions. *Computers & Education*, 59(2), 378-384.
- Lipponen, L. (2002). Exploring foundations for computer-supported collaborative learning. In Stahl, G. (Ed.) *Proceedings of CSCL 2002* (pp. 72-81). Boulder, CO: ISLS.
- Marbouti, F. (2012). Design, implementation and testing of a visual discussion forum to address new post bias. Unpublished masters thesis. Burnaby, CA: Simon Fraser University.
- Pena-Shaff, J. B. & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education*, 42(3), 243-265.
- Peters, V. & Hewitt, J. (2010). An investigation of student practices in asynchronous computer conferencing courses. *Computers & Education*, 54(4), 951 - 961.
- Stahl, G. (2005). Group cognition in computer-assisted collaborative learning. *Journal of Computer Assisted Learning*, 21(2), 79-90.
- Thomas, M. J. W. (2002). Learning within incoherent structures: The space of online discussion forums. *Journal of Computer Assisted Learning*, 18(3), 351-366.
- Webb, E., Jones, A., Barker, P., & van Schaik, P. (2004). Using e-learning dialogues in higher education. *Innovations in Education and Teaching International*, 41(1), 93-103.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, 46(1), 71-95.
- Wise, A. F., Hsiao, Y. T., Marbouti, F., Speer, J. & Perera, N. (2012a). Initial validation of "listening" behavior typologies for online discussions using microanalytic case studies. In J. van Aalst, J., K. Thompson, K., M. Jacobson, & P. Reimann (Eds.) *Proceedings of ICLS 2012* (pp. 56-63). Sydney, Australia: ISLS.
- Wise, A. F., Marbouti, F., Hsiao, Y. & Hausknecht, S. (2012b). A survey of factors contributing to learners' "listening" behaviors in asynchronous discussions. *Journal of Educational Computing Research*, 47(4), 461-480.
- Wise, A. F., Perera, N., Hsiao, Y., Speer, J. & Marbouti, F. (2012c). Microanalytic case studies of individual participation patterns in an asynchronous online discussion in an undergraduate blended course. *Internet and Higher Education*, 15(2), 108-117.
- Wise, A. F., Saghafian, M. & Padmanabhan, P. (2012d). Towards more precise design guidance: Specifying and testing the functions of assigned student roles in online discussions. *Educational Technology Research and Development*, 60(1), 55-82.
- Wise, A. F., Speer, J., Marbouti, F. & Hsiao, Y. (2013). Broadening the notion of participation in online discussions: Examining patterns in learners' online listening behaviors. *Instructional Science*. 41(2), 323-343.

Influence of Epistemological beliefs and Goal Orientation on Learning Performance in CSCL

Kui Xie, Ohio State University, Columbus, OH 43210, xie.359@osu.edu
Kun Huang, University of North Texas Health Science Center, Fort Worth, TX 76107,
kun.huang@unthsc.edu

Abstract: Epistemological beliefs were found to affect students' cognitive engagement, study strategies, as well as motivation in classroom settings. However, research on the relationship between epistemological beliefs, motivation, and student learning in online settings has been under-studied and under-theorized. This study investigated the relationships among epistemological beliefs, achievement goals, and learning performance in online discussions. 124 students participated in discussion activities in an online instructional technology course. The results from correlation analysis and structural equation modeling indicate that epistemological beliefs have significant effects on students' learning performance, through the mediation of achievement goals.

Introduction

Epistemological beliefs (EB) were found to affect students' cognitive engagement, study strategies, as well as motivation. However, research on the interplay between EB, motivation, and student participation and perception in the context of online learning has been under-studied and under-theorized. In order to fill this gap in the literature, this study built and tested a model among EB, motivation, and student participation and perception in asynchronous online discussion activities in a college-level online class.

Theoretical Framework

EB refers to individual's beliefs about knowledge and knowing (Hofer & Pintrich, 2002). Studies on EB generally focus on five dimensions: the *structure of knowledge* (ranging from the belief that knowledge can be described as isolated pieces to the belief that knowledge structure is complex and highly interrelated), the *certainty of knowledge* (ranging from the belief that knowledge is certain and unchanging to the belief that knowledge is tentative and evolving), the *nature of ability to learn* (ranging from the belief that ability to learn is innate to the belief that learning ability can be acquired with effort), the *speed of learning* (ranging from the belief that learning takes place quickly or not at all to the belief that learning is a gradual process), and the *source of knowledge* (ranging from the belief that knowledge comes from omniscient authorities to the belief that knowledge emerges from personal construction) (Schommer, 1990; Schraw, Bendixen & Dunkle, 2002; Braten & Stromso, 2005). Past research suggested that EB could influence academic performance (e.g., Qian & Alvermann, 1995; Schommer-Aikins, Duell, & Hutter, 2005; Schommer-Aikins & Easter, 2006; Schommer, 1993; Windschitl, 1997), cognitive engagement (e.g., DeBacker & Crowson, 2006; Ravindran, Greene, & DeBacker, 2005), and study strategies (e.g., Kardash & Howell, 2000; Schommer, Crouse, & Rhodes, 1992).

Students' motivation in achievement settings is often examined through the lens of goal orientations, which focus on the reasons for students' engagement in academic activities. According to achievement motivation theories, there are three types of goals: mastery goals, performance-approach goals, and performance-avoidance goals (A. Elliot, 1999; A. Elliot & Church, 1997; E. Elliot & Harackiewicz, 1996). Studies have reliably shown the connection between achievement goals and learning. For example, individuals with mastery goals are likely to choose challenging tasks, show interest, persistence and effort, demonstrate self-regulated learning, adopt meaningful study strategies, and achieve better learning outcomes (A. Elliot, McGregor, & Gable, 1999; E. Elliot & Dweck, 1988; Eppler & Harju, 1997; Graham & Golan, 1991; Licht & Dweck, 1984). Self-efficacy is another theory that examines students' motivation. Defined as "the conviction that one can successfully execute the behavior required to produce the outcomes," (Bandura, 1977, p. 193), self-efficacy is a person's inter belief of his or her own competency. Compared with individuals with low self-efficacy, those with higher levels of self-efficacy are likely to initiate effort when faced with challenging tasks, show more persistence and effort, adopt meaningful learning strategies, and achieve better academic learning outcomes (Greene & Miller, 1996; Pajares & Miller, 1994; Walker, Greene, & Mansell, 2006; Zeldin & Pajares, 2000; Zimmerman & Bandura, 1994). Recent studies investigated the relationship between EB and students' motivation in classroom settings. The findings generally suggested that students who held less mature beliefs were less likely to adopt mastery goals but more likely to adopt performance goals, and those who held more mature beliefs were on the contrary (e.g., Braten & Strømsø, 2004; DeBacker & Crowson, 2006; Schutz, Pintrich, & Young, 1993).

While studies have demonstrated the individual relationships between EB, motivation and learning, there is a need to understand how these variables interact with each other within a broader network. Existing

few attempts examined these networked relationships in classroom settings (Chen & Pajares, 2010; Kizilgunes, Tekkaya, & Sungur, 2009; Ravindran, et al., 2005). As today's learning environment is shifting from classroom to online settings, it becomes necessary to investigate the interrelationships among these constructs in an online learning environment.

Further, recent research indicated that students' actual participation and learning behavior can greatly reflect their motivation (Fredricks, Blumenfeld, & Paris, 2004; Reeve, 2006). However, studies on the relationship between EB and learning have been mainly relying on students' self-reported cognitive engagement (e.g., DeBacker & Crowson, 2006; Ravindran, et al., 2005) or study strategies (e.g., Kardash & Howell, 2000; Schommer, et al., 1992), while the relationship between EB and students' learning behaviors is largely unknown. Technologies like learning management systems (e.g., WebCT©) make it possible to track a variety of students' participation records in online learning activities (Xie, 2012), which is a challenging tasks in face-to-face settings (Dennen, 2008). The online participation data enable researchers to develop better insights into students' online learning behaviors and their relationship with motivation.

This study investigated the interplay among EB, motivation, and student participation and perception in asynchronous online learning. Based on the existing literature, we hypothesized that EB would influence students' motivation (i.e., achievement goals and self-efficacy), which in turn affected students' participation in as well as perception of online learning. We further hypothesized that EB not only influence online learning indirectly through motivation, but also had direct impact on online learning.

Methodology

Participants

124 students (33 males and 91 females) from an online instructional technology class at a large Southeast university participated in this study. The course focused on... The ages were ranged from 19 to 61. One of the major assignments in this course was online discussions on selected topics of instructional technology. The online discussions were facilitated by iDiscuss, a system that is capable of tracking students' participation record. 116 participants (94%) rated their confidence level as high in the use of technology to complete the coursework.

Measurement

Five groups of variables were measured in this study (1) EB, (2) goal orientation, (3) perceived competence, (4) perceived learning performance, and (5) online participation.

Epistemological beliefs inventory (EBI) was used to assess students' EB (32 items) (Schraw, Bendixen & Dunkle, 2002). The EBI has five subscales: (1) simple knowledge, (2) certain knowledge, (3) fixed ability, (4) quick learning and (5) omniscient authority. A recent validation study documented that the subscale α of those measuring variables ranged from 0.50–0.60s across two administrations, and test–retest correlations ranged from 0.62 to 0.81 (see Schraw et al, 2002). Questionnaire on approaches to learning was used to assess students' three achievement goal orientations: (1) mastery, (2) performance-approach, and (3) performance-avoidance. The subscale α 's typically ranging from 0.60–0.90s (see Greene, Dillon & Crynes, 2003; Miller, Greene, Montalvo, Ravindran & Nichols, 1996). Perceived competence scale is composed of 6 questions asking students to rate their perceived learning competence in online discussions. Perceived learning scale has 10 questions measuring students' perceptions of their learning. Students' online participation was measured by the quantitative data collected from the iDiscuss system, which includes students' posting behaviors (total number of posts) and non-posting behaviors (total topics read, number of times logged in).

Procedure

Students worked in small groups of 8-10 to participate daily in online discussion activities in a 16-week session. Each student moderated a chapter discussion for a designated week within his or her own group. During the course, student peers were invited to share information and contribute to knowledge constructions in the chapter discussions. The instructor monitored students' discussion activities and supported the discussion when needed. Participants were invited to complete the above-mentioned set of instruments during the semester.

Results

To test the relationships among EB, perceived competence, achievement goals, and learning performance (including perceived learning performance and actual online participations) in a CSCL setting, three steps of analyses were performed. The first step analyzed the sample items, descriptive statistics, and reliability coefficient for the items measuring each of the variables. The second step generated a correlation matrix to examine the relationships among variables of interest. The third step involved the utilization of Structural Equation Modeling (SEM), using maximum likelihood techniques with AMOS version 18. SEM allows

researchers to identify the latent variables and explore the pattern of inter-relationships between variables in a structured framework.

Table 1 shows the sample items of measured variables, their descriptive statistics, and the reliability coefficient among the items within each of the variables. As can be found in Table 1, students scored higher in the *mastery goals* (mean = 5.96) and *perceived learning* (mean = 5.26). They scored lower in two of the EB variables: *quick learning* (mean = 2.46) and *certain knowledge* (mean = 3.07).

Table 1. Sample items, descriptive statistics, and reliability coefficient.

Variable: Sample Item	Mean	SD	Min-Max	α
Omniscient authority: People who question authority are troublemakers.	4.91	.91	2.40-6.80	.61
Quick learning: Students who learn things quickly are the most successful.	2.46	.95	1.00-5.40	.69
Simple knowledge: Too many theories just complicate things.	4.09	.75	1.86-6.14	.62
Fixed ability: How well you do in school depends on how smart you are.	3.80	.92	1.71-6.00	.71
Certain knowledge: What is true today will be true tomorrow.	3.07	.82	1.50-5.25	.58
Mastery goals: I want to learn as much as possible from this class.	5.96	.90	2.50-7.00	.88
Performance-approach goals: It is important for me to do better than the other students.	4.38	1.49	1.00-7.00	.91
Performance-avoidance goals: My goal for this class is to avoid performing poorly.	4.61	1.24	1.00-7.00	.82
Self-efficacy: I feel that I am pretty competent in the online discussions.	5.20	1.03	1.67-7.00	.86
Perceived learning: The discussions helped me to think more deeply.	5.26	1.31	1.00-7.00	.95

Note. SD – standard deviation; α – Cronbach α coefficient.

Correlation among variables

An examination of zero-order correlations, shown in Table 2, provides the validity of composite variables in EB, perceived competence, goals, perceived learning performance and online participation.

Table 2. Pearson correlation among epistemological beliefs, goals, online behavior, perception variables.

	1	2	3	4	5	6	7	8	9	10	11	12
1. Omniscient authority	1.00	.08	.24**	.22*	.33**	.16	.06	.22*	-.13	.13	.05	.09
2. Quick learning		1.00	.33**	.55**	.26**	-.22*	.15	.22*	-.05	-.11	-.08	-.11
3. Simple knowledge			1.00	.32**	.18*	-.01	.21*	.22*	-.04	-.06	-.13	.01
4. Fixed ability				1.00	.18*	-.04	.16	.16	-.15	-.12	-.15	-.09
5. Certain knowledge					1.00	-.11	-.00	.12	-.35**	-.06	-.05	-.06
6. Mastery goals						1.00	.21*	-.05	.42**	.25**	.23*	.35**
7. Performance-approach goals							1.00	.24**	.18*	-.05	-.02	.11
8. Performance-avoidance goals								1.00	-.27**	-.12	-.17	-.04
9. Self-efficacy									1.00	.25**	.27**	.21*
10. Posting participation										1.00	.62**	.29**
11. Non-posting participation											1.00	.22**
12. Perceived learning												1.00

* $p < .05$ level (two-tailed).

** $p < .01$ level (two-tailed).

Zero-order correlation analysis results (Table 2) suggested that EB variables are correlated with each other, except for the relationship between *authority* and *quick learning*. Among the goal variables, *mastery goal* has a significant correlation with *performance-approach goal*, which significantly correlates with *performance-avoidance goal*. The effects of EB on achievement goals are partially supported by the correlation between them. In addition, *self-efficacy* has a positive correlation with *mastery goal*, *performance-approach goal*, and all of online learning variables. However, it has a negative correlation with *performance-avoidance goal*.

Structural Equation Models

For the purpose of this study, one hypothetical model was built representing EB's direct influence on learning performance, as well as indirect influence through the mediation of motivation variables. In the final model,

Structural Equation Modeling did not find significant direct relationship between EB and learning performance, suggesting that EB did not have direct effects on online learning performance. The results corresponded to the correlation results in Table 2, which suggested no significant correlation between EB and learning performance. The significant relationships among variables in the final model supported EB's significant influence on online learning through the mediation of motivation variables. The model had a good fit of the data as indicated by the Chi-square goodness-of-fit indices ($\chi^2 = 2828$, $\chi^2/df = 1.693$) being less than 2, indices GFI (.911) being greater than 0.9, and the RMSEA (.049) falling below 0.05.

Out of the five EB variables, both omniscient authority ($\beta = -.33$, $p = .009$) and quick learning ($\beta = -.51$, $p = .026$) had significantly negative effects on mastery goal. Fixed ability significantly predicted both performance-approach goal ($\beta = .28$, $p = .043$) and performance-avoidance goal ($\beta = .12$, $p = .029$). No other relationship between EB and goals was detected.

The relationships between self-efficacy and goals were significant in the final model. Self-efficacy had significant positive effects on both mastery goal ($\beta = .39$, $p < .001$) and performance-approach goal ($\beta = .23$, $p = .041$). It also negatively influenced performance-avoidance goal ($\beta = -.42$, $p < .001$). The results were echoed by the correlation results, which denoted that self-efficacy was positively correlated with mastery goal and performance-approach goal while negatively correlated with performance-avoidance goal. The relationships between self-efficacy and online learning participation were also significant. Self-efficacy had significantly positive effects on both posting participation ($\beta = .36$, $p = .05$) and non-posting participation ($\beta = .50$, $p = .018$). Self-efficacy did not show significant relationship with perceived learning and class engagement.

With respect to the relationships between goals and online learning, several significant paths were found in the resulted model. Mastery goal significantly influenced all of the online learning variables: perceived learning ($\beta = .43$, $p < .001$), online posting participation ($\beta = .51$, $p = .004$), and online non-posting participation ($\beta = .46$, $p = .020$). All the effects were positive. In addition, performance-avoidance goal had a significantly negative relationship with perceived learning ($\beta = -.14$, $p = .041$) and non-posting behavior ($\beta = -.22$, $p = .044$).

Discussions and Implications

In this study, we investigated the influence of EB on students' learning and motivation in online settings. SEM method was used to examine the relationships among these variables.

First, the results indicated that achievement goals serve as mediators to bridge the relationship between EB and students' learning performance in CSCL settings. The SEM results indicated no direct relationship between EB and online learning variables, but suggested that EB had an indirect effect on online learning participation and perception which was channeled via achievement goals. These results were supported by previous studies, which suggested that EB may function as implicit theories that influence the adoption of goals for learning, and these goals can also mediate the relations between EB and cognition and learning performance (Hofer & Pintrich, 1997), as well as learning approach and achievement (Kizilgunes, et al., 2009).

Second, the results indicated that different EB systems and students' perceived competence predicted students' goal orientations in CSCL settings. Particularly, students' beliefs about the speed of learning and omniscient authority appeared to affect students' mastery goal adoption. The model suggested that students who believed that learning occurs quickly were less likely to adopt mastery goals, which is inline with previous studies that found beliefs in quick learning negatively predict mastery goals (e.g., Bråten & Strømsø, 2004). On the other hand, the model also indicated that those who believed that knowledge came from authority were more likely to adopt mastery goals. This result contradicts the finding of Bråten and Strømsø's (2004) study, which found that students who conceived of knowledge as stable and given were less likely to adopt mastery goals. Besides omniscient authority and quick learning, this study found that perceived competence might function as an antecedent for all three achievement goals. Furthermore, the results indicated that students' beliefs about fixed ability to learn influenced their performance goal adoption. Those who believe that the ability to learn was fixed at birth were more likely to adopt performance goals. Specifically, when students believe their learning ability is innate and stable, they will be less likely to emphasize competence development and less likely to foster mastery goals. This finding is in agreement with achievement motivation theories, which suggest that a mastery goal is characterized by the belief that academic effort will result in an achievement, and in contrast, a performance goal reflects the belief that ability alone leads to success (Topping & Ehly, 1998). As a result, if students' perceived competence is high, they tend to adopt performance approach goals and demonstrate their competence in front of others; if their perceived competence is low, they tend to adopt performance avoidance goals try to avoid the demonstration of their incompetency in front of others.

Another result in this study shows that achievement goals played a significant role in predicting students' perceived learning and their actual learning behaviors in CSCL settings. Mastery goals positively predicted all the learning behavior and perception variables, showing that students with mastery goals are motivated to learn and engage in CSCL. Performance avoidance goals, however, negatively predicted students' non-posting behavior and perceived learning in CSCL setting. This is because students with performance avoidance goals tend to avoid failure rather than acquire competence. Instead, they focused more on completing

the minimum class requirements and were less likely to participate in those non-posting online learning activities (e.g., reading, evaluation, etc.) that cannot be observed by teachers and peers.

This study makes contributions to the literature by moving beyond the self-reported measures traditionally adopted in motivation research and incorporating students' actual participation data in the modeling. This study also makes contributions to the literature by moving beyond the face-to-face setting and investigating the relationship among these critical constructs in online learning settings.

The findings of this study can help education practitioners to understand the lacuna of their current teaching activities through proper understanding of the effects of EB on students' motivation and online learning participation and perception. In addition, the mediating effects of goals between EB and learning performance can provide insights to educators that different teaching strategies or technologies should be developed to facilitate the study of students with different learning goals.

References

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215.
- Braten, I., & Strømsø, H. (2004). Epistemological beliefs and implicit theories of intelligence as predictors of achievement goals. *Contemporary Educational Psychology*, 29(3), 371-388.
- Chen, J., & Pajares, F. (2010). Implicit theories of ability of Grade 6 science students: Relation to epistemological beliefs and academic motivation and achievement in science. *Contemporary Educational Psychology*, 35, 75-87.
- DeBacker, T., & Crowson, M. (2006). Influences on cognitive engagement: Epistemological beliefs and need for closure. *British Journal of Educational Psychology*, 76, 535-551.
- Dennen, V. P. (2008). Looking for evidence of learning: Assessment and analysis methods for online discourse. *Computers in Human Behavior*, 24(2), 205-219.
- Dweck, C. S. (1999). *Self-Theories: Their role in motivation, personality, and development*. Philadelphia, PA: The Psychology Press.
- Elliot, A. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34(3), 169-189.
- Elliot, A., & Church, M. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, 72(1), 218-232.
- Elliot, A., McGregor, H., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology*, 91(3), 549-563.
- Elliot, E., & Dweck, C. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), 5-12.
- Elliot, E., & Harackiewicz, J. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, 70, 461-475.
- Eppler, M., & Harju, B. (1997). Achievement motivation goals in relation to academic performance in traditional and nontraditional college students. *Research in Higher Education*, 38(5), 557-573.
- Fredericks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59-109.
- Graham, S., & Golan, S. (1991). Motivational influences on cognition: Task involvement, ego involvement, and depth of information processing. *Journal of Educational Psychology*, 83, 187-194.
- Greene, B., & Miller, R. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology*, 21, 181-192.
- Hofer, B., & Pintrich, P. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88-140.
- Hofer, B., & Pintrich, P. (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kardash, C., & Howell, K. (2000). Effects of epistemological beliefs and topic-specific beliefs on undergraduates' cognitive and strategic processing of dual-positional text. *Journal of Educational Psychology*, 92(3), 524-535.
- Kizilgunes, B., Tekkaya, C., & Sungur, S. (2009). Modeling the relations among students' epistemological beliefs, motivation, learning approach, and achievement. *The Journal of Educational Research*, 102(4), 243-255.
- Licht, B., & Dweck, C. (1984). Determinants of academic achievement: The interaction of children's achievement orientations with skill area. *Developmental Psychology*, 20, 628-636.
- Pajares, F., & Miller, D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology*, 86, 193-203.
- Qian, G., & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology*, 87(2), 282-292.

- Ravindran, B., Greene, B., & DeBacker, T. (2005). Predicting preservice teachers' cognitive engagement with goals and epistemological beliefs. *The Journal of Educational Research, 98*(4), 222-232.
- Reeve, J. (2006). Extrinsic rewards and inner motivation. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 645-664). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Schommer-Aikins, M., Duell, O., & Hutter, R. (2005). Epistemological beliefs, mathematical problem-solving beliefs, and academic performance of middle school students. *The Elementary School Journal, 105*(3), 289-303.
- Schommer-Aikins, M., & Easter, M. (2006). Ways of knowing and epistemological beliefs: Combined effect on academic performance. *Educational Psychology, 26*(3), 411-423.
- Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology, 85*, 406-411.
- Schommer, M., Crouse, A., & Rhodes, N. (1992). Epistemological beliefs and mathematical text comprehension: Believing it is simple does not make it so. *Journal of Educational Psychology, 84*(4), 435-443.
- Schraw, G., Bendixen, L. D., & Dunkle, M. E. (2002). Development and validation of the Epistemic Belief Inventory. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 103-118). Mahwah, NJ: Erlbaum.
- Schutz, P. A., Pintrich, P., & Young, A. J. (1993). *Epistemological beliefs, motivation, and student learning*. Paper presented at the annual meeting of the American Educational Research Association.
- Walker, C., Greene, B., & Mansell, R. (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learning and Individual Differences, 16*, 1-12.
- Windschitl, M. (1997). Student epistemological beliefs and conceptual change activities: How do pair members affect each other? *Journal of Science Education and Technology, 6*(1), 37-47.
- Wu, D., & Hiltz, S. R. (2004). Predicting learning from asynchronous online discussions. *Journal of Asynchronous Learning Networks, 8*(2), 139-152.
- Zeldin, A., & Pajares, F. (2000). Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal, 37*(1), 215-246.
- Zimmerman, B., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal, 31*, 845-862.
- Xie, K. (2012). What do the numbers say? The influence of motivation and peer feedback on students' behavior in online discussions. *British Journal of Educational Technology*. DOI: 10.1111/j.1467-8535.2012.01291.x

Volume 1

Symposia

Embedding Participatory Design into Designs for Learning: An Untapped Interdisciplinary Resource?

Elizabeth Bonsignore, June Ahn, Tamara Clegg, Mona Leigh Guha, Jason C. Yip, Allison Druin (*Discussant*)
University of Maryland, Human-Computer Interaction Lab, 2117 Hornbake Bldg, College Park, MD 20742
Juan Pablo Hourcade, Dept. of Computer Science, University of Iowa, 14 MacLean Hall Iowa City, IA 52242
ebonsign@umd.edu, juneahn@umd.edu, tclegg@umd.edu, mona@cs.umd.edu, juanpablo-hourcade@uiowa.edu, jasonyip@umd.edu, allisond@umiacs.umd.edu

Abstract: Given the rapid evolution of social networks and online communities, interest in participatory cultures—online and offline social spaces with low barriers to entry and support for creating and sharing knowledge—is increasing. Design-based research (DBR) that invites children to share in the process of designing the technologies that support their learning is a natural extension of this participatory cultures movement. In this symposium, we establish a rationale for using Participatory Design (PD) techniques that can inform and enrich the process of designing technologies that support collaborative learning. We provide empirical examples from our own research of the ways in which PD can be incorporated into learner-centered technology designs. Our experiences demonstrate that PD is not only a key contributor in the design of learning technologies themselves, it can also be a valuable resource that sheds light on the learning processes of the children who use them.

Introduction

Participatory Design (PD) is an array of theories, practices, and research methods whose core philosophy is to include end-users as active participants in the technology design process (Muller, 2008; Schuler & Namioka, 1993). Our symposium's integrating theme is the notion that PD can be a valuable resource for researchers engaged in design-based research (Hoadley, 2002) throughout the learning sciences: one that may enrich the process of designing technologies that support collaborative, socially constructed learning, and also benefit the learners who participate in their design. Yet, a review of current research across the learning sciences reveals that PD is a largely untapped resource in design-based research (DBR) studies. A handful of researchers across the learning sciences and human-computer interaction (HCI) have been engaged in design-based approaches that closely resemble each other in that they invite the children being studied to be partners throughout the iterative design process (e.g., Ahn, Gubbels, Kim & Wu, 2012; Barab, Thomas, Dodge, Carteaux & Tuzun, 2005). Although these pockets of potential exist, we have not yet enjoyed cross-pollination among these parallel research tracks. Our symposium is a first step in building a deeper dialogue among researchers in CSSL and HCI that can lead to an integrated, interdisciplinary research agenda that incorporates PD into broader DBR frameworks. In particular, we focus on the potential benefits of enlisting children as full partners with adult researchers and designers, throughout the design process of a learning program or system. Our goals are to explore 1) why PD methods hold such promise to augment existing DBR approaches, and 2) how PD techniques can be integrated into DBR projects.

Background

PD grew out of the Scandinavian trade union democracy movement of the 1970s, where its initial contextual focus was the workplace, with the goal of giving union workers a greater voice in the design of computers that they were required to use (Bødker, Ehn, Sjögren, & Sundblad, 2000). The primary motivation for PD is the democratic ideal that the people who are affected by a decision or event should be given the opportunity to influence it. A key corollary is that the goal of technology design is not to “automate the skills of human workers,” but “to give workers better tools for doing their jobs” (Schuler & Namioka, 1993, p. xi). PD is sometimes used synonymously with the term, co-design (cooperative/collaborative design); however, co-design emphasizes a more in-depth, and often equal, partnership between designer and end-user, and enlists user participation at earlier stages in the design process (Bødker et al., 2000; Walsh, Foss, Yip, & Druin, 2013).

While PD was created initially in a sociopolitical context, its use has been extended to many other user populations, such as children (Druin, 1999, 2002), individuals with disabilities (Frauenberger, Good, & Keay-Bright, 2011), and older adults (Ellis & Kurniawan, 2000). As the diversity of user populations engaged in PD has grown, its methodologies have also been extended to design-based research contexts outside the workplace, such as urban planning and policy (Friedman et al., 2008), social media (Hagen & Robertson, 2010), and education (Barab et al., 2005). In general, PD approaches such as Contextual Inquiry (Holtzblatt & Jones, 1993) evolved from working with adults during technology design processes. Similarly, Cooperative Inquiry adapts and extends PD techniques; however, the users who partner with design researchers in Cooperative Inquiry are children, not adults (Druin, 1999, 2002; Guha, Druin, & Fails, 2013).

In PD projects that develop technology for children, children can assume various roles, usually at specific points in the design process (Druin, 2002). For example, a child can evaluate a system as a user and tester (e.g., near or just after design completion), or as an informant (e.g., early in the design cycle). In Cooperative Inquiry, however, children act as full partners with adult designers throughout the design process (Druin, 2002). Child designers, typically between the ages of 7-11, are actively involved in technology design from conception to completion, sharing ideas and evaluations equally with the adult members of the team. Likewise, adults play active roles throughout the process by first helping child designers generate and articulate ideas and then synthesizing those ideas into manageable designs. Cooperative Inquiry enriches PD research with techniques that enable teams of children and adults to share ideas in ways that maximize idea elaboration yet minimize differences in age, ability, and communication styles (Druin, 1999). For example, “Bags-of-Stuff” is a Cooperative Inquiry technique in which art supplies are used by intergenerational design teams to build low-tech prototypes of the technologies they envision (Walsh et al., 2013). Because of Cooperative Inquiry’s emphasis on equal partnership, it is often referred to as co-design. However, it is important to note that Cooperative Inquiry is a specific co-design method whose distinction is its focus on child design partners. All of the learning design projects presented in this symposium were informed by Cooperative Inquiry approaches.

Participatory Approaches in the design of Learning Environments

In many ways, the user-centered core of PD approaches is closely related to the ideals underpinning learner-centered theories and methods developed in the CSCL research community. Both are founded on the principle that their target populations (end-users and learners) are best served when they are given a high degree of agency in the process under investigation (technology design or learning). Each embodies the participatory cultures movement, emphasizing the cultivation of knowledge communities in which content and expertise are co-created (Jenkins Clinton, Purushotma, Robinson, & Weigel, 2006). Still, few CSCL studies seem to have integrated PD techniques (Barab et al., 2005). Existing studies also enlist learners in limited design roles (e.g., informant and tester), rather than the full spectrum of PD roles available—especially that of equal partner. Others focus on how adult educators, not child learners, can be co-opted into the design process (e.g., Hernandez-Leo et al., 2006).

Even the limited research that has employed PD promotes the value of such techniques for learning. For example, Stanton, Neale, and Bayon (2002) combined the educational goal of learning collaboration skills with the PD philosophy of giving children as much control as possible during their extended development work with KidPad, a drawing application adapted for use with multiple mice and tangible interfaces (Druin, Stewart, Proft, Bederson, & Hollan, 1997). KidPad enjoyed sustained successful use in primary schools in the United Kingdom for at least three years (Stanton et al., 2002). Quest Atlantis (QA), a multi-user virtual environment (MUVE) developed by Barab et al. (2005) to better understand the role of online community and games in supporting academic content learning, combined elements from methodological approaches in critical ethnography, participatory action research, and PD to develop a “critical design” philosophy (Barab, Dodge, Thomas, Jackson, & Tuzun, 2007). The QA research team attributed their success to their cultivation of “rich relationships with [children] we came to regard not simply as ‘participants’ but as ‘collaborators’” (Barab et al., 2007, p. 280), which also resulted in a collaborative interface that was effective in supporting academic goals. More recently, PD-like approaches with adult educators have shown promise in fostering the design of novel participatory assessment systems (Itow & Hickey, 2012).

The collaborative and cognitive benefits of design thinking have also been touted in learning contexts. Research in game-based learning, game design, and e-textiles design has demonstrated the problem-solving expertise and increased agency that children can derive from their active participation in these projects (Kafai, 1996; Kafai, Fields, & Searle, 2012; Squire, 2011). Design thinking is an integral element of these approaches, just as design thinking is promoted in Cooperative Inquiry approaches. This wealth of evidence in the learning sciences begs the question, why limit the approach to game design or physical computing? Why not include learners in the process of designing technologies they are using to learn, as well?

Perhaps one of the obstacles to building more convergent paths across the respective disciplines is related to the paradox of informed participation: individuals cannot really be informed unless they participate; yet they cannot really participate unless they are informed (Eden, 2002). For CSCL research, this challenge is reflected in the question, how can learners actively and effectively design technologies for concepts that they have not yet learned? PD research faces a similar conundrum: in a rapidly evolving, increasingly ubiquitous technological landscape, how can users contribute to ideation and concept development for contexts-in-use that they have not yet experienced? These parallel, but nearly equivalent challenges serve to underscore the potential value of increasing collaboration and cross-pollination across CSCL and HCI.

Complimentary Perspectives: Embedding PD into Designs for Learning

The presentations included in this symposium offer different lenses for considering PD approaches in learning contexts. Our first presentation explores the potential cognitive and social benefits that PD techniques

hold for the child designers who engage in them. Our next two presentations review the ways in which PD techniques enhanced collaborative technologies that promote digital literacies (e.g., multimedia, storytelling) and STEM learning. Our fourth presentation addresses the tension between content expertise and design expertise, offering a comparison of PD results from children experienced in design with children learning in STEM environments. Our final presentation explores opportunities and challenges facing CSCL and HCI researchers who seek to include learners with special needs, such as those with autistic spectrum disorder, in the design process.

From a theoretical perspective, our exploratory studies reveal that PD techniques can be both an effective means for improving technology designs, and a valuable resource in DBR frameworks. From a practical perspective, we demonstrate various PD approaches we took that may inform future DBR studies across CSCL and HCI that seek to incorporate PD into their methodologies (e.g., Guha, Druin, & Fails, 2013).

Presentation 1: The Cognitive and Social Experiences of Children Involved in PD

Mona Leigh Guha

Over the past several years, researchers and designers across academia and industry have partnered with children to design new technologies (Gibson, Newall, & Gregor, 2003; Takach & Varnhagen, 2002). While many studies have explored the effects that co-design with children has on the resulting technology, few have investigated the impact of co-design participation on the child designers themselves. Specifically, there has been little documentation on the impact that becoming an equal partner in co-design with adults may have on the participating child designers. Much of the literature regarding children's involvement in technology design includes only incidental mentions of the potential benefits to child design partners. Formal studies focusing on the child design partner's experiences are scarce. My study represents an initial foray into discovering if there are indeed specific social and cognitive experiences available to child design partners (Guha, 2010).

First, I will introduce co-design with children as a set of PD techniques that involve children as equal partners throughout the technology design process, from conception to completion. I will include a discussion of various co-design techniques, including the points during the design process in which they may be used most effectively. Next, I will describe a qualitative case study that explored the social and cognitive experiences of children involved in co-design with adults. The children who participated in the study were members of an intergenerational team of adult researchers and child designers, known collectively as Kidsteam, from the Human-Computer Interaction Lab (HCIL). In this case study, I followed eight child design partners over the course of a year, collecting data such as observational notes and artifacts, as well as interviews with the children and their parents. The data was inductively coded into categories to arrive at a model of the social and cognitive experiences of child design partners. In particular, the child design partners had experiences in the cognitive domain in skills and content, in the social domain in relationships, enjoyment, and confidence, and in the overlapping social and cognitive domains of communication and collaboration (Guha, 2010). I will detail the conceptual framework that describes these social and cognitive experiences of child design partners, and also propose opportunities for future collaborative, interdisciplinary work among HCI and CSCL researchers.

Children across a range of ages, not just 7-11 year olds, can be included successfully in the co-design process. At the HCIL over the past decade, we have found that even very young children (4-5 years old) have life experiences that they can draw upon to contribute to a learning technology design process, given flexible design techniques and equal footing with their adult design partners (Guha et al., 2004). At the other end of the childhood age spectrum, we have also found ways to use co-design with pre-teens (Knudston et al., 2003) and teenagers (Yip, Foss, & Guha, 2012). Each of these co-design efforts shows promise for their incorporation into the design process for learning technologies as well.

Presentation 2 – Participatory Design, Mobile Storytelling, and New Media Literacies

Elizabeth Bonsignore

I will relate the impact of PD in the development of an online digital library for children, the International Children's Digital Library (ICDL), and a mobile storytelling application (StoryKit). The two projects highlight the ways in which PD processes resulted in collaborative, creative technologies that were highly intuitive and engaging for children. Their intuitive, child-friendly features prompted their international proliferation not only in family/social contexts, but also in a variety of learning contexts.

The ICDL (childrenslibrary.org) is a freely available, online collection of children's literature from around the world. The ICDL collection includes over 4500 books in 61 languages, representing 65 countries. Over 6.5 million children and adults from 228 countries have visited the ICDL since its official launch in November 2002, with an average of 100,000 visitors logged per month. The design goals for the ICDL were to inspire intercultural awareness in children (ages 3-13) by providing broad online access to an international collection of children's literature; to create new technologies that were age appropriate and engaging; and to expand existing PD methods by involving children in the design process (Druin, 2005). The HCIL's Kidsteam engaged in iterative design cycles from 2000-2005, testing the viability of various search/browse/reading

interfaces (Druin, 2005; Hutchinson, Druin, & Bederson, 2006). Kidsteam's co-design approaches offered rich insights into how children want to browse and search for books (and information in general), and how they prefer to read digital texts. For example, co-design efforts with Kidsteam confirmed that children prefer to search for books in ways that reflect their physical, image-based features (e.g., colors on the cover) or affective elements (e.g., a "happy" or "sad" book), which resulted in an innovative and popular search interface as well as new metadata categories for library cataloguers to consider (Druin, 2005). I will also share the ways in which the ICDL design has been shown to increase children's motivation to read, expand the variety of books they choose, and support their interest in exploring different cultures (Druin, Weeks, Massey, & Bederson, 2007).

One of the design findings from the ICDL project was that children not only wanted to read beautiful books from around the world; they also wanted to create and share their own stories. Furthermore, they wanted to be able to carry and share their stories with them in the same "mobile" ways that they used print books: in the lap of a grandfather, or on a bus with a friend. StoryKit is a mobile application (app) that enables the creation of multimedia stories on iOS devices. Within StoryKit's integrated interface, children can create original stories, or modify sample ICDL stories, using their own photos, drawings, text, and audio. StoryKit authors can also share their stories with friends and family via the Internet. Like the ICDL, an intergenerational design team of adults and children (ages 7-77) designed StoryKit. Since its launch in the Apple iTunes App Store in September 2009, StoryKit has been used over 2 million times by over 385,000 distinct users in 175 countries and in 40 languages/dialects. I will show how the design features that were implemented as a result of the co-design process with Kidsteam resulted in an intuitive, integrated interface whose use has skyrocketed in classrooms around the world. For example, almost 60% of shared stories have been created in formal education settings (i.e., school). The audio tool that is integrated into StoryKit's interface as a result of Kidsteam design sessions has also proven very effective in supporting children in early primary grades (ages 5-7) by allowing them to tell their stories with confidence orally, even as they are learning to read and write (Bonsignore, Quinn, Druin, & Bederson, in press). Moreover, educators have found that the saving and sharing process, which creates a digital artifact of their primary and elementary level writers' progress in literacy activities (ages 6-8), has also helped them to scaffold their efforts in the art of reflection and revision (Bonsignore et al., in press).

Presentation 3 – Participatory Design and Social Media for STEM Learning

June Ahn

I will describe the role of PD in two projects that explore the development of social media platforms for youths and STEM learning. The two cases will highlight how PD yielded new insights about both 1) the design of more engaging and sociable CSCL platforms, and 2) the learning process of children themselves as it relates to aspects of STEM education. In these projects, children and youths engaged in participatory STEM learning practices as they co-designed technologies with the research team. I argue that PD is not only salient for the actual design of technology tools, but is also a valuable contributor to broader DBR frameworks in the learning sciences. Researchers using PD can not only learn about better design techniques, but also glean insight into the learning processes of children that can then be more tightly integrated into the design process.

The first case, Sci-identity, is an NSF funded Cyberlearning project that engages urban, inner-city youths in science fiction storytelling using diverse digital media, as a means for enhancing their identification with STEM ideas. A major component of the project is the design of a social media platform (sci-identity.org) where youths can share, comment on, and remix their stories. We used PD techniques with two groups: the HCIL's Kidsteam and the middle school students who were themselves a part of the program. I will outline how Kidsteam provided deep insight into the learning process required when children attempt the complex activity of writing science-infused narratives. Kidsteam child design partners were particularly insightful about the potential obstacles children face when attempting such a complex literacy practice, and their insights led to the direct design of particular social media features on the Sci-identity.org site, such as the "Brain," which became a shared repository of science knowledge that authors could use as inspirations for their stories. In our design work with middle school adolescents (twens), we gained deeper insight into the controversial issue of remix practices. Their concerns about remix directly resulted in the redesign of remix functionality in the social media platform. The students in this project simultaneously learned about remix behavior while actively designing their technological environment in ways that reflected their values. We found that: "Working from a youth perspective allows one to recognize the underlying mechanisms for sharing, credit, and permission, and design these functions in ways that align with the perspective of youths" (Ahn, Subramaniam, et al., 2012, p. 7).

For the second case, I will outline the development of SINQ, a social media platform that leverages features seen in popular sites such as Instagram and Reddit, in ways that promote collaborative learning of Scientific INquiry (Ahn, Gubbels, et al., 2012). In this project, we embarked on a series of development "sprints" that involved programming work by the research team coupled with co-design sessions with Kidsteam. The design narrative of this experience sheds light on how specific PD techniques in each of the development sprints allowed us to 1) glean insight into how children acquire scientific inquiry skills and dispositions in their everyday thinking, and incorporate these learning processes into the design of SINQ; and 2) observe the ways in

which children interact with media, and online communities, in ways that helped us create a more engaging, usable, and sociable CSCL tool. Each of our development sprints could also be conceptualized as cycles of DBR in the learning sciences. Our experience creating SINQ illuminates how children and youths can act as both learners who critically reflect on their STEM learning practices and co-designers who provide direct design recommendations to the developers of the social media platforms used in these projects.

Presentation 4 – Design Expertise and Content Expertise: Complementary Perspectives in the Design of Technologies for Science Learning

Tamara Clegg and Jason Yip

Our project originates from learning sciences research aimed at helping learners to see themselves more scientifically by facilitating their engagement in personally meaningful science experiences. We approached this goal by designing a life-relevant learning (LRL) environment that 1) engages children in science in a context relevant to their everyday lives (cooking) and 2) helps them begin to use science to achieve their own personal goals. This LRL environment, Kitchen Chemistry (KC), engages elementary and middle school children in scientific inquiry through making and perfecting dishes. We draw upon technology to support and scaffold learners' scientific inquiry as they cook.

In previous work, we found that the learning environment helped children begin to engage more deeply in science and to identify themselves scientifically. We also found that the technology we used supported the cognitive aspects of learners' experiences, but not their personally meaningful aspects. Consequently, technology use needed to be prompted and heavily supported by facilitators (Clegg, Gardner, & Kolodner, 2011). To enhance the potential for technologies that are used in LRL environments to contribute to learners' personally meaningful experiences, we have used Cooperative Inquiry-based PD techniques to re-design technology support for LRL. We began these efforts by using StoryKit, a technology that had been designed through Cooperative Inquiry with Kidsteam. We found that learners were highly engaged when they used StoryKit, and that the app supported their scientific inquiry processes (Clegg et al., 2012) and decision-making practices (Yip, Clegg, et al., 2012) quite naturally. We are currently extending StoryKit's affordances to design new mobile technologies that specifically scaffold learners' scientific inquiry in their daily lives. We remain committed to taking PD approaches to address the challenge of engaging learners in the design of learning technologies as they are still learning the concepts and practices the technology is aimed at supporting.

We will also discuss our efforts in addressing this challenge with two different co-design teams whose goal was to re-design early iterations of KC technology. Specifically, we will outline the results from two case studies of design sessions with members from Kidsteam and participants of two different iterations of KC (Yip, Clegg, et al., in press). Kidsteam children are well versed in many co-design techniques and consistently work together with adults each week (Druin, 1999). In contrast, KC children are experts in the learning context and subject domain knowledge. KC learners become specialists in developing food science investigations, integrating cooking and observation techniques, and using technology for collaborative learning. Because we are developing technology for learning environments, it was important for us to design with children who had expertise in co-design and the subtle contextual knowledge of KC. We explored three research questions on co-designing learning technologies with children: 1) What are the affordances and constraints of designing learning technologies with children who have subject expertise; 2) with children who have design expertise? and 3) How can the results of designing with the two groups be combined to inform design practice that involves either group? (Yip et al., in press). We advocate that comparing two sets of co-designers allows for triangulation and insight into complementary design ideas that can be optimized for the design of learning technologies.

Presentation 5 – Enhancing the Social Skills of Children with Autism Spectrum Disorders with Multitouch Applications and Activities, supported by Participatory Design

Juan Pablo Hourcade

Children in the autism spectrum have been gaining a greater amount of attention recently, in great part due to increasing rates of diagnosis (CDC, 2012). Early diagnosis and therapy can make a significant positive difference in improving children's social skills. However, even a majority of those who receive early diagnosis and treatment do not grow up to live independently as adults (Eaves & Ho, 2008; Howlin, Goode, Hutton & Rutter, 2004). This means that in addition to emphasizing early diagnosis, there is a need to develop new interventions that can improve the likelihood that children diagnosed with autism will grow up with the skills necessary to live independently. I will outline my research team's efforts to add to our understanding of interventions that can support this special, under-represented population, highlighting the ways in which PD techniques enhanced our design activities and outcomes.

When the first multi-touch tablets became commercially available in the summer of 2009, we saw an opportunity to provide a novel intervention to children in the autism spectrum. While children in the spectrum vary significantly in their needs and ability, in our experience, most of them have a strong affinity for computers. This may be due to the predictability of computers and the ability to control them, something that

does not occur with face-to-face interactions. We saw the potential for multi-touch technologies to provide the ability for children to engage in face-to-face interactions while doing something they enjoy.

As we began to consider how to proceed, we contacted local parent support groups, observed children, and also visited with a local group of adults with Asperger's syndrome. As we engaged with these groups we developed a set of principles that guided our work, which we refer to as APPS: Access, Participation, Personalization, and Sustainability. In terms of access, we emphasized designing and working on technologies that become widely accessible, given the important and immediate needs of the community. We engaged with all stakeholders including the children, their parents, teachers, and special education staff. In terms of personalization, we addressed the high variability in this population by designing several simple, open-ended apps together with activities that could be conducted with them. The combination of apps and activities gave us a lot of flexibility in personalizing what children experienced, and made it more likely we could find an activity and app that could work for a specific set of children (Hourcade, Bullock-Rest, & Hansen, 2012). Finally, to support the sustainability of the project, we made all the code open source and activity guides freely available online (openautismsoftware.org) to ensure others could pick up the project and expand it on their own.

We used a variety of techniques to engage the children in PD activities. Children who could speak without much trouble were able to participate in the same type of activities in which typically developing children participate. If anything, they were less likely to hold back criticism. We had to be more creative with lower-functioning children who spoke little or did not speak at all. To get feedback from them about apps or design ideas, we had to use picture cards, or simply write “yes” and “no” on sticky notes and ask them to point at one of them when asking them whether they liked a particular aspect of an activity or app. We also learned that some of the children really disliked changes in the user interface if they had gotten used to an earlier version. In one case it caused a great deal of frustration in a child, so we recommend that interaction designers in similar situations keep previous versions of apps handy (Hourcade et al., 2012).

We have conducted an evaluation of the impact of the apps and activities on children's behavior, and found in a recent study that children had more verbal and physical interactions when using the apps than when conducting similar activities without apps (Hourcade, Williams, Miller, Huebner, & Liang, 2013). We also found some of the apps led to more supportive comments from children than other apps or activities without apps. Our experiences point to the feasibility and value of engaging all stakeholders including children with autism in PD activities, as well as the positive effects of our particular approach to engaging children in face-to-face activities to improve their social skills.

Significance

The aim of our symposium is to advance the practice of DBR across multiple disciplines through collaborative investigations into PD techniques that enhance the design of learning technologies as well as benefit the learners who participate in their design. The integrative nature of our goal is significant to the conference theme of “learning across levels of space, time, and scale,” precisely because it seeks to reduce disciplinary boundaries and amplify commonalities across design methods. Each of the studies detailed here offers a metaphorical grain of sand that, taken together, builds a worldview demonstrating the potential for PD techniques to shape future learner-centered systems. From the outset, CSCL has been an interdisciplinary field. Deepening the dialogue between HCI and CSCL researchers who are using participatory approaches in learning contexts can only serve to strengthen CSCL's interdisciplinary tradition.

References

- Ahn, J., Gubbels, M., Kim, J. & Wu, J. (2012). SINQ: Scientific INquiry learning using social media. In *Extended Abstracts on Human Factors in Computing Systems (CHI EA 2012)*, Austin, TX (pp. 2081 – 2086). NY, USA: ACM.
- Ahn, J., Subramaniam, M., Fleischmann, K. R., Waugh, A., Walsh, G., & Druin, A. (2012). Youth identities as remixers in an online community of storytellers: Attitudes, strategies, and values. *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–10.
- Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Development*, 53(1), 86–107.
- Barab, S., Dodge, T., Thomas, M. K., Jackson, C., & Tuzun, H. (2007). Our designs and the social agendas they carry. *Journal of the Learning Sciences*, 16(2), 263–305.
- Bødker, S., Ehn, P., Sjögren, D., & Sundblad, Y. (2000). Co-operative Design - perspectives on 20 years with “the Scandinavian IT Design Model.” In *Proceedings NordiCHI 2000*, Stockholm, Sweden, 1–9. Retrieved from: http://cid.nada.kth.se/pdf/cid_104.pdf
- Bonsignore, E., Quinn, A. J., Druin, A., & Bederson, B. B. (in press). Sharing stories in the wild: a mobile storytelling case study. *Transactions of Computer-Human Interaction (ToCHI)*.

- Centers for Disease Control and Prevention (CDC). (2012, March 29). *CDC - Division of Birth Defects, National Center on Birth Defects and Developmental Disabilities - Data and Statistics, Autism Spectrum Disorders*. Retrieved from <http://www.cdc.gov/ncbddd/autism/data.html>
- Clegg, T., Bonsignore, E., Yip, J., Gelderblom, H., Kuhn, A., Valenstein, T., Lewittes, B., et al. (2012). Technology for promoting scientific practice and personal meaning in life-relevant learning. In *Proceedings of the 11th International Conference on Interaction Design and Children (IDC 2012)*, Bremen, Germany, (pp. 152-161). NY, USA: ACM.
- Clegg, T. L., Gardner, C. M. & Kolodner, J. L. (2011). Technology for supporting learners in out-of-school learning environments. (2011). In H. Spada, G. Stahl, N. Miyake, N. Law (Eds.), *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL 2011 Conference Proceedings, Vol. 1* (pp. 248-255). International Society of the Learning Sciences.
- Druin, A., Stewart, J., Proft, D., Bederson, B., & Hollan, J. (1997). KidPad: a design collaboration between children, technologists, and educators. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '97)*, Atlanta, GA (pp. 463-470). NY, USA: ACM.
- Druin, A. (1999). Cooperative inquiry: developing new technologies for children with children. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*, Pittsburgh, PA, (pp. 592-599). NY, USA: ACM.
- Druin, A. (2002). The role of children in the design of new technology. *Behaviour & Information Technology*, 21(1), 1-25.
- Druin, A. (2005). What children can teach us: developing digital libraries for children with children. *The Library Quarterly*, 75(1), 20-41.
- Druin, A., Weeks, A., Massey, S., & Bederson, B. B. (2007). Children's interests and concerns when using the international children's digital library: a four-country case study. In *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '07)*, Vancouver, BC (pp. 167-176). NY, USA: ACM.
- Eaves, L.C. & Ho, H.H. (2008). Young adult outcome of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38(4), 739-47.
- Eden, H. (2002). Getting in on the (inter)action: exploring affordances for collaborative learning in a context of informed participation. In G. Stahl (Ed.), *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community (CSCL '02)* (pp. 399-407). International Society of the Learning Sciences.
- Ellis, R. D., & Kurniawan, S. H. (2000). Increasing the Usability of Online Information for Older Users: A Case Study in Participatory Design. *International Journal of Human-Computer Interaction*, 12(2), 263-276.
- Frauenberger, C., Good, J., & Keay-Bright, W. (2011). Designing technology for children with special needs: bridging perspectives through participatory design. *CoDesign*, 7(1), 1-28.
- Friedman, B., Borning, A., Davis, J. L., Gill, B. T., Kahn Jr., P. H., Kriplean, T., & Lin, P. (2008). Laying the foundations for public participation and value advocacy: interaction design for a large-scale urban simulation. In *Proceedings 2008 International Conference on Digital Government Research* (pp. 305-314). Montreal, Canada: Digital Government Society of North America.
- Gibson, L., Newall, F., & Gregor, P. (2003). Developing a web authoring tool that promotes accessibility in children's designs. In *Proceedings of the 2003 Conference on Interaction Design and Children (IDC '03)*, Preston, England (pp. 23-30). NY, USA: ACM.
- Guha, M. L. (2010). *Understanding the social and cognitive experiences of children involved in technology design processes* (3443561). University of Maryland, College Park. *ProQuest Dissertations & Theses (PQDT)*. (855814381).
- Guha, M. L., Druin, A., Chipman, G., Fails, J. A., Simms, S., & Farber, A. (2004). Mixing ideas: a new technique for working with young children as design partners. In *Proceedings of the 2004 Conference on Interaction Design and Children (IDC '04)*, College Park, Maryland (pp. 35-42). NY, USA: ACM.
- Guha, M.L., Druin, A., and Fails, J. (2013). Cooperative Inquiry revisited: Reflections of the past and guidelines for the future of intergenerational co-design. *International Journal of Child-Computer Interaction*, 1(1), 14 - 23.
- Hagen, P., & Robertson, T. (2010). Social technologies: challenges and opportunities for participation. In *Proceedings of the 11th Biennial Participatory Design Conference*, (pp. 31-40). NY, USA: ACM.
- Hernandez-Leo, D., Villasclaras-Frenandez, E. D., Asensio-Perez, J. I., Dimitriadis, Y., Jorin-Abellan, I. M., Ruiz-Requies, I., & Rubia-Avi, B. (2006). COLLAGE: A collaborative Learning Design editor based on patterns. *Educational Technology & Society*, 9(1), 58-71.
- Hoadley, C. P. (2002). Creating context: design-based research in creating and understanding. In G. Stahl (Ed.), *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community (CSCL '02)*, Boulder, CO (pp. 453-462). International Society of the Learning Sciences.

- Holtzblatt, K., & Jones, S. (1993). Contextual Inquiry: A Participatory Technique for System Design. In D. Schuler & A. Namioka (Eds.), *Participatory design: principles and practices* (pp. 177–210). Hillsdale, NJ, USA: Erlbaum.
- Hourcade, J.P., Bullock-Rest, N.E., & Hansen, T.E. (2012). Multitouch tablet applications and activities to enhance the social skills of children with autism spectrum disorders. *Personal and Ubiquitous Computing*, 16(2), 157-168.
- Hourcade, J.P., Williams, S.R., Miller, E.A., Huebner, K.E., & Liang, L.J. (2013). Evaluation of tablet apps to encourage social interaction in children with autism spectrum disorders. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. NY, USA: ACM.
- Hutchinson, H. B., Bederson, B. B., & Druin, A. (2006). The evolution of the international children's digital library searching and browsing interface. In *Proceedings of the 2006 Conference on Interaction Design and Children (IDC '06)*, Tampere, Finland (pp. 105–112). NY, USA: ACM.
- Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry*, 45(2), 212–29.
- Itow, R. C., & Hickey, D. T. (2012). Design-Based Implementation of Spreadable Educational Practices within the Participatory Learning and Assessment Network (PLANet). In J. van Aalst, K. Thompson, M.J. Jacobson, and P. Reimann, (Eds.), *Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012)*, Vol. 2 (pp. 553-554). International Society of the Learning Sciences.
- Jenkins, H., Clinton, K., Purushotma, R., Robinson, A. J., & Weigel, M. (2006). Confronting the challenges of participatory culture: media education for the 21st century. Chicago, IL: MacArthur Foundation.
- Kafai, Y. B. (1996). Learning design by making games. In Yasmin B. Kafai & M. Resnick (Eds.), *Constructionism in practice: designing, thinking and learning in a digital world* (pp. 71–96). Mahwah, NJ: Erlbaum.
- Kafai, Y. B., Fields, D. A., & Searle, K. A. (2012). Making Technology Visible: Connecting the Learning of Crafts, Circuitry and Coding in Youth e-Textile Designs. In J. van Aalst, K. Thompson, M.J. Jacobson, and P. Reimann, (Eds.), *Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012)*, Vol. 1 (pp. 188-195). International Society of the Learning Sciences.
- Knudtzon, K., Druin, A., Kaplan, N., Summers, K., Chisik, Y., Kulkarni, R., et al. (2003). Starting an intergenerational technology design team: a case study. In *Proceedings of the 2003 Conference on Interaction Design and Children (IDC '03)*, Preston, England (pp. 51–58). NY, USA: ACM.
- Muller, M. (2008). Participatory Design: The Third Space in HCI. In A. Sears & J. Jacko (Eds.), *The human-computer interaction handbook* (2nd ed., pp. 1061–1082). NY, USA: Taylor & Francis.
- Schuler, D., & Namioka, A. (1993). *Participatory Design: Principles and Practices*. Hillsdale, NJ: Erlbaum.
- Squire, K. (2011). *Video Games and Learning: Teaching and Participatory Culture in the Digital Age*. NY, USA: Teachers College Press.
- Stanton, D., Neale, H., & Bayon, V. (2002). Interfaces to support children's co-present collaboration: multiple mice and tangible technologies. In G. Stahl (Ed.), *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community (CSCL '02)*, Boulder, CO (pp. 342–351). International Society of the Learning Sciences.
- Takach, B. S., & Varnhagen, C. (2002). Partnering with children to develop an interactive encyclopedia. In M. M. Bekker, P. Markopoulos, & M. Kersten-Tsikalkina (Eds.), *Proceedings of the International Workshop "Interaction Design and Children"*, Eindhoven, The Netherlands (pp. 129–137). Shaker.
- Walsh, G., Foss, E., Yip, J., Druin, A. (2013). FACIT PD: Framework for Analysis and Creation of Intergenerational Techniques for Participatory Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2013)*, Paris, France. NY, USA: ACM.
- Yip, J.C., Clegg, T.L., Bonsignore, E., Gelderblom, H., Lewites, B., Guha, M.L., & Druin, A. (2012). Kitchen Chemistry: Supporting learners' decisions in science. In J. van Aalst, K. Thompson, M.J. Jacobson, and P. Reimann, (Eds.), *Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012)*, Vol. 1 (pp. 103 – 110). International Society of the Learning Sciences.
- Yip, J. C., Clegg, T. L., Bonsignore, E., Gelderblom, H., Rhodes, E., & Druin, A. (in press). Brownies or Bags-of-Stuff? Domain Expertise in Cooperative Inquiry with Children. In *Proceedings of the 12th International Conference on Interaction Design and Children (IDC 2013)*. NY, USA: ACM.
- Yip, J.C., Foss, E., & Guha, M.L. (2012). Co-designing with adolescents. Paper presented at the Designing Interactive Technology for Teens Workshop, NordiCHI, Copenhagen, Denmark. Retrieved from: <http://www.chici.org/ditt2012/papers.html>

Acknowledgments

For each of the studies described in our presentations, we thank the participants, their parents, and the local school communities that partnered with us. For the work undertaken at the HCIL, we also thank Kidsteam child and adult designers who have successfully co-designed a host of novel technologies over the past decade.

Mass Collaboration – an Emerging Field for CSCL Research

Organizer

Ulrike Cress, Knowledge Media Research Center (KMRC), 72076 Tuebingen, Schleichstr. 6, Germany,
u.cress@iwm-kmrc.de

Presenters

Brigid Barron, School of Education, Stanford University, barronbj@stanford.edu
Gerhard Fischer, Center for Lifelong Learning and Design, University of Colorado, Boulder,
gerhard@colorado.edu
Ulrike Cress, Iassen Halatchlyiski, Aileen Oeberst, Knowledge Construction Lab, KMRC, Tuebingen
Email: u.cress@iwm-kmrc.de, i.halatchlyiski@iwm-kmrc.de, a.oeberst@iwm-kmrc.de
Andrea Forte, College of Information Science and Technology, Drexel University, aforte@drexel.edu
Mitchel Resnick, Media Lab, MIT, mres@media.mit.edu

Discussant

Allan Collins, collins@bbn.com

Abstract: Mass collaboration is a present-day Internet practice with far-reaching implications for education and for a knowledge society in general. The goal of this symposium is to establish the concept of mass collaboration as a relevant topic in CSCL by presenting related research conducted with multifaceted perspectives from different labs. Several presentations will provide insight into the current approaches to the complex and large-scale phenomenon. They will address a range of theories and methodologies to identify and approach the major aspects of learning and knowledge development in the informal context of the present-day social web. The presenters will ground the analyses of mass collaboration processes and outcomes on a variety of examples of effective Web 2.0 settings and technological environments. The anticipated discussion will map out what unique insights can be gained from studying this form of collaboration compared to more formalized small-scale settings, as well as possible directions for future research in the area.

Introduction

Whereas in former times collaboration was mostly bound to smaller groups, the Internet tools of today provide various possibilities for the collaboration of masses of users. There is an almost unlimited variety of online communities where users share personal stories, experiences, or anything that can be expressed digitally. In wikis thousands of users collaboratively gather and organize knowledge. With social tagging systems users annotate and share online resources. The participants in such communities are not just a mass of learning individuals or passive consumers; they actively produce meaningful content and act as “prosumers” (O’Reilly, 2006; Tapscott & Williams, 2006). Most importantly, their activity develops outside the formal educational system (National Research Council, 2009).

The symposium will demonstrate the relevance of mass collaboration for CSCL research by bringing together the latest theoretical and empirical endeavors in this research area. It is organized by the Knowledge Construction Lab at the KMRC, Tuebingen, where the *Co-Evolution Model of Individual Learning and Collaborative Knowledge Building* (Cress & Kimmerle, 2008) serves as a systemic framework for describing and analyzing mass collaboration. The symposium presents results from this line of research together with relevant work being done in other institutes. Together the presentations want to open a dialogue about those aspects of mass collaboration that have potential for education and learning and are thus of interest for the learning sciences (Cress, in press; Cress & Fischer, subm.; Fischer, 2011). The discussion will build on the symposium on long-tail learning at the CSCL conference in 2009 (Brown & Adler, 2008; Collins, et al 2009) and will extend the view to various forms of online interaction when masses of people learn together, collaborate and create new knowledge. The presentations will introduce relevant theoretical approaches to the process of mass collaboration; empirical studies will detail out different analysis methodologies, and examples of communities and Web 2.0 environments in formal and informal settings will ground the discussion of practical issues.

In sum, the aim of the symposium is threefold:

- (1) to establish the concept of mass collaboration as a relevant topic of CSCL
- (2) to present and interconnect existing research on the subject, and
- (3) to give prototypical examples of learning communities and mass collaboration platforms

In the following we give a short introduction to the concept of mass collaboration, providing the background for the presentations in the symposium.

CSCL and Mass Collaboration

Since the emergence of CSCL as a research field the predominant approach has been to study small groups of students in a neatly arranged situation: The students engage in synchronous discourse around a problem-solving task, and the sequence of their interactions represents a major research interest with regard to meaning making and learning outcomes. *Collaboration* has mainly been used in the sense of Roschelle and Teasley, (1995; p. 70) who defined it as “a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem”. Whereas in its early times CSCL mainly dealt with synchronous small-group or classroom settings, currently, CSCL research faces a much broader range of real-life situations, where people take part in decentralized communities, act asynchronously and do not necessarily all come to a shared conception of a problem. In fact, complex knowledge phenomena involve long periods of time, larger and changing numbers of people, and fuzzy-structured settings. In this spirit, any human achievement can be seen as a collaborative accomplishment – in terms of *dwarves standing on the shoulders of giants*. Extending the view on collaboration beyond small groups leads to a macro approach that considers the complexity of knowledge development across space, time, and collectives of people. CSCL research is just starting to address this global level of human learning and knowledge creation (Kafai & Peppler, 2011; Collins et al, 2009). The increasing number of Web 2.0 communities and tools now bring this perspective to the foreground.

Artifacts as Mediators of Mass Collaboration

The large-scale perspective raises the question of how intersubjective understanding and collaborative activities of a vast number of people are coordinated. Bearing in mind that most of the participants cannot interact directly, an efficient form of coordination is needed. All the more so when the process is collaborative and fulfills the conditions that individuals act consciously following a common direction; that they take the perspective of the other participants into account; and that they contribute by building on the accomplishments of others. The solution to the coordination problem resides in *artifacts* that support the collaborative process by mediating common understanding (Bruner, 1996). Beyond making artifacts accessible for a large number of people, social software environments afford their collaborative creation, revision and remixing in a mass collaborative process. Collaborative artifacts represent crystallized knowledge that is preserved from past interactive situations and that can be built on in future to produce phenomena like scientific understanding, social practices or social norms (Engeström & Sannino, 2010; Paavola, Lipponen & Hakkarainen, 2004). An artifact is a means to an end and at the same time it is a desired product whose creation can even be the main goal of the collaboration process (see for example Kafai & Resnick, 2000). Theoretical and methodological approaches to mass collaboration need to be centered around artifacts as major elements in the complex process across space, time, and collectives of people.

Learning through Mass Collaboration

Whereas formal learning mostly takes place organized in smaller groups and classroom settings, mass collaboration by definition happens “in the wild”, and thus, mainly outside the educational system. Most of the Web 2.0 communities that share and create knowledge maintain an informal context. Activities in these communities induce individual learning at the same time as they demonstrate collective *knowledge creation* (Paavola, Lipponen & Hakkarainen 2004). In order to interact, people have to create or refer to artifacts. In the process of artifact construction and use people deepen their understanding and enhance their individual knowledge (Kafai & Resnick, 2000). Thus, learning and knowledge creation in a collaborative setting are one continuous process of internalization and externalization (Nonaka & Takeuchi, 1995; Vygotsky, 1978).

During mass collaboration it is obvious that learning occurs not just at the *level of individuals*, but also at the *community level*: A community may enhance its knowledge base, and may deepen its understanding of a given topic. Through cumulated reciprocal referencing during a discourse some views in the group may become more pronounced than others. Some ideas stand the test of time, others fade away. This is obvious in the example of wikis: Some ideas remain in the collaborative text, while others are revised or deleted soon after they have been contributed. These processes are not just random. They show that a community represents a self-organized, autopoietic system (Maturana & Varela, 1987). Knowledge development within this system is an emergent process. It depends on individuals’ activities but is not reducible to the knowledge of independent individuals. The “learning” of a community and the learning of the participating individuals are intertwined and suggest a systemic understanding of mass collaboration (Cress & Kimmerle, 2008).

Mass Collaboration Environments

There are many examples of mass collaboration environments for learning and knowledge creation. Perhaps the most remarkable example is the online encyclopedia *Wikipedia*. It offers a unique field for studying large-scale open-ended knowledge processes from large amounts of data on the history of articles and authors' contributions. Wikipedia is a dynamic knowledge artifact of interconnected articles deliberately produced by a complex system of individual collaborative activities. The explicit written content mediates shared understanding on a specific topic amalgamating perspectives and styles of expression of a multitude of authors into a coherent exposition. Although Wikipedia is not aimed at developing new knowledge or at providing a learning environment for the contributors, the processes that unfold there share some essential features of scientific and knowledge-building discourse (Forte & Bruckman, 2006; Swarts, 2009). Thus there is a vibrant field of research concerned with data from Wikipedia, and several presentations in this symposium (those by Aileen Oeberst, Iassen Halatchlyiski and Andrea Forte) will be concerned with it.

While Wikipedia has emerged as a self-organized community, a lot of mass collaboration platforms have been deliberately designed to support learning. One prominent example is the *Scratch online community* (presentation by Michael Resnick), where children playfully create visual artifacts like animations, stories, and games. Learning and knowledge creation are again supported by and organized around user-generated artifacts. The environment affords users not only the possibility to share and comment on their projects, but also to use parts of the existing projects to create new ones (Resnick et al., 2009; Brennan, Resnick & Monroy-Hernandez, 2010; Monroy-Hernandez, 2012). Thus, learning and knowledge development in the Scratch community is sustained by a culture of remixing of existing content. Analogous to Wikipedia the individual contributions may be interlinked, revised and rebuilt. Contributors use others' contributions to create new artifacts, which then in turn represent the basis for future developments.

The third example of a mass collaboration environment in this symposium is Citizen Science (presentation by Brigid Barron). It stands not for one specific technology, but subsumes various projects where amateurs or nonprofessional scientists participate in scientific activities. Amateurs, for example, collect data or test natural phenomena. Here it is obvious that their activities lead to new knowledge. The main goal of the projects is not the learning of the individuals, but the creation of new knowledge by using scientific method.

Presentation 1: Cultures of Participation — Fundamental Transformations of Learning, Working, and Collaborating

The first presentation is an introduction to the topic, pointing out the current and societal relevance of mass collaboration. **Gerhard Fischer**, director of the Center for Lifelong Learning and Design (University of Colorado, Boulder), explores with mass collaboration in the context of *cultures of participation* (Fischer, 2011; Jenkins, 2009). He states that social media enable a shift from consumer cultures, which are specialized in producing finished artifacts to be consumed passively, to *cultures of participation*, in which all people are provided with the means to participate and to actively contribute in *personally meaningful* problems. The participatory web supports moving away from a world in which a small number of people create artifacts, define rules, make decisions concerning many others towards a world in which everyone would be able to actively participate following their own interests and to make their voices heard. This shift has introduced unique and fundamental opportunities, challenges, and transformative changes for innovative research in CSCL. Fischer will explore in his presentation the theoretical foundations and system developments for understanding, fostering, and supporting cultures of participation. His framework is centered on the following aspects:

- *Meta-design* defines and creates social and technical infrastructures for cultures of participation in which new forms of collaborative learning and design can arise.
- *Social creativity* transcends the individual human mind by making sense of the *variety of voices*, in order to frame and solve complex problems. Shared artifacts enable relevant transdisciplinary collaborations.
- *Rich ecologies of participation* will emerge based on by different levels of participation, expertise, interests, and motivation.
- *Idiosyncratic interests* and unique contributions by self-directed learners will lead to long-tail distributions of knowledge.
- *Drawbacks* of cultures of participation can be seen in aspects of “do-it-yourself” societies, in fragmented cultures in which people live in their own “filter bubbles”, and in accumulation of irrelevant information.

The framework is grounded in a variety of different application contexts (including: open source software, urban planning, assistive technology, energy sustainability, and formal education). The presentation will include initial design guidelines and explore the implications of these developments for future research and innovations in technology enhanced learning.

Presentation 2: The Co-Evolution Model as a Theoretical Framework for Describing Mass Collaboration

The second presentation will introduce research conducted at the KMRC in Tuebingen. It consists of three parts with three different presenters: **Ulrike Cress** will present the *Co-Evolution Model of Individual Learning and Collaborative Knowledge Building* as a research framework for describing and analyzing learning processes in mass collaboration (Cress & Kimmerle, 2008; Kimmerle, Cress & Held, 2010). The model describes individuals as cognitive systems, who externalize their knowledge by creating artifacts. The community deals with the individuals' contributions by interlinking, revising or even rejecting them. The community acts as a social system that deals with the information according to its own rules. Knowledge is an emergent product of this process. Individual learning takes place when users consume, that is, internalize existing knowledge, adopt the rules of a community, and participate by contributing artifacts. Thus, on the one hand, a system makes use of an individual's knowledge to create a collective knowledge base; on the other hand, an individual develops her or his own knowledge by actively participating in a community. This means that mass collaboration should be analyzed not only at the level of individuals, but also at the level of the whole community as a social system that shapes the individual's activities.

The following two presentations give an example of how the relevant system level processes can be analyzed. The first study is a qualitative in-depth analysis of a distinctive Wikipedia article, while the second is a quantitative large-scale analysis basing on thousands of interlinked pages.

Aileen Oeberst will present a study of the development of the Wikipedia article about the nuclear power plant in Fukushima during the first nine days after the nuclear disaster on March 11, 2011. She will show how the rules of the social system Wikipedia guided individual contributions to a topic where—at that time—no verified knowledge existed. People had to deal with a flood of novel, highly specific and mainly uncertain incoming information and they had to construction meaning out of different information sources. Many authors were involved in this process. Their activities and coordination were only mediated by Wikipedia's norms, that valuable and thus accepted contribution should be verifiable and written from a neutral point of view. This norm effectively shaped individuals' edits and enabled authors without much domain expertise to construct an article, which nuclear experts judged to be of high quality. These findings bring not the individuals as relevant contributors but the rules of the knowledge system to the fore. The system's definition of knowledge implicitly coordinated the knowledge construction process, and shaped the activities of its contributors. In sum, the presentation shows that mass collaboration is much more than sharing and accumulating knowledge across individuals. It is an emergent process, where the system makes use of individuals to create system-accepted knowledge.

Iassen Halatchliyski will present an approach to studying mass collaboration that can encompass large-scale and long-term characteristics: the macro level of the phenomenon (Halatchliyski, Moskaliuk, Kimmerle & Cress, 2010). In the example of Wikipedia this macro level can be perceived using the concept of knowledge domains in analogy to scientific fields that govern scientific production. New knowledge is essentially situated (Lave, 1988) and complies with the specific context of its production. This means that the characteristics of a knowledge domain are a major factor of the future development of knowledge—not only in this domain, but also in the adjacent domains. Such influence is expected to stem from the already accumulated knowledge base and from the active participants in the domain. In Wikipedia, this reasoning raises two interrelated questions: what are the important articles and who are the important authors in a knowledge domain? Halatchliyski uses the term *pivotal knowledge* referring to those Wikipedia articles that are an important part of the structure of a knowledge domain, and thus, influence the future development of knowledge. In this presentation he will show how to identify pivotal knowledge by adapting the concept of a network to a knowledge base of thousands of interconnected Wikipedia articles.

The network concept has already been used to describe knowledge organization at different levels such as the semantic memory of individuals (e.g., Collins & Loftus, 1975), or the meaning-making process within a group discourse (Stahl, 2006). Using a social network analysis methodology, the complex, system-level patterns of knowledge creation during mass collaboration are now becoming an accessible focus of CSCL research.

Presentation 3: Large-Scale Collaboration and Cultures of Accountability

The third presentation deals not only with the aspect of production, but also with the aspect of consumption of knowledge and information provided in mass collaboration environments. **Andrea Forte** from Drexel University's iSchool will talk about the intersection of critical consumption and production of participatory information resources. She will discuss her research group's ongoing efforts to understand how contributing to online information resources can not only yield content knowledge, but also expose learners to cultures of accountability and equip them to engage critically with information sources.

Research on collaborative environments involves understanding who learns what, when, and how. In Wikipedia and other wiki systems, collaborative writing supports editors in learning *about* a wide range of topics and the characteristics of this learning have been the topic of many studies, including Forte's and others in this Symposium. However, there are other kinds of learning taking place as people become socialized in communities of collaborative authorship.

As the world's largest and most widely accessed collection of reference works, quality is a critical concern for Wikipedia projects. Each language edition has developed complex policies and social norms that help contributors construct high-quality artifacts and that guide discourse on the site (Butler et. al., 2008; Kriplean et. al., 2007); this has bred a culture of accountability among contributors. Analogously, when asked to contribute to a public information resource as part of their classwork, high school students report a sense of responsibility to their potential readership for the quality of their written work (Forte & Bruckman, 2009). Even when they work individually, this sense of responsibility is manifest in student strategies for citing, organizing content, and looking for information. Moreover, when they have experience producing information online, some students begin to leverage these experiences to assess the quality online information resources they use. Forte's research group is expanding this work to better understand how experiences with large-scale online collaboration affect people's understanding and assessment of online resources both on Wikipedia and in more specialized collaboratively produced resources like Ancestry.com or Findagrave. Early findings suggest that experience in contributing to participatory information sources yield sophisticated strategies for engaging with online sources.

Presentation 4: Learning through Remixing

Mitchel Resnick from the MIT Media Lab will discuss his group's research on mass collaboration in the context of *Scratch* (<http://scratch.mit.edu>), an online community in which young people (ages 8 and up) program interactive stories, games, animations, and simulations – and share their creations with one another online (Resnick et al., 2009). Since the launch of Scratch in 2007, young people around the world have shared roughly 3 million projects in the online community, adding (on average) two new projects every minute. The collection of projects is incredibly diverse: interactive newsletters, science simulations, virtual tours, animated dance contests, interactive tutorials, and many others, all programmed with Scratch's graphical programming blocks.

An important aspect of participation in the Scratch online community is the ability to remix other people's projects. Members of the community not only interact with and comment on one another's projects, they can also modify, extend, and repurpose the programming code and media elements underlying the projects. All projects in the Scratch community are covered by Creative Commons Share-Alike license, so community members are free to build upon the sprites, scripts, images, and sounds of other people's projects. Roughly one-third of all projects in the Scratch community are remixes of other projects.

In his presentation, Resnick will analyze and discuss:

- Different approaches to remixing within the Scratch community
- What people learn (and don't learn) as they remix one another's projects
- How ideas spread through the community through remixing
- Community attitudes towards remixing – and factors underlying those attitudes
- Design strategies for supporting and encouraging a culture of remixing

Many of the ideas and examples in the presentation will draw upon the research of Andres Monroy-Hernandez, who recently finished his PhD dissertation focusing on remixing in the Scratch community (Monroy-Hernandez, 2012).

Presentation 5: Long-tail Learning and Access to External Resources

Brigid Barron and Caitlin Martin from Stanford University will report findings from a study of a genre of cyber-enabled massively collaborative activity known as *Citizen Science*. Citizen Science projects capitalize on the interest and efforts of non-scientist collaborators who join forces to contribute data that helps address problems of concern. Networked technologies have dramatically changed the potential of such projects. With mobile GPS enabled data collection devices, data contributions can easily be shared. Applications like Google Maps make it easy to share location based information and online databases allow data contributions to be reviewed by professional scientists and community members. Our research project was based in the premise that we need to understand Cyberlearning as a human-technical system and that to advance design relevant knowledge we need to attend to both the social community and the ways that the technology supports learning within the community. We chose to study a Citizen Science project with a significant focus on education called *Vital Signs*. Vital Signs is a citizen science networked system located in the state of Maine, linked statewide to schools and accessible not only to teachers but to anyone who want to learn. Because the state has a longstanding laptop program started in 2002, all middle school students have access to their own iBook, which they use at home and at school. Middle school teachers were also provided with laptops, technical assistance, and professional development. We will present findings from our research organized to address three main questions:

- 1) What patterns of personalized learning can we identify among a diverse group of teachers/learners who vary in prior experiences and community socioeconomic status?

- 2) When do Vital Signs opportunities sustain engagement beyond the classroom (for example by sparking family-based learning activities through a school-based project thereby bridging across formal and informal settings)?
- 3) How can member contributions to networked communities be mined and harvested for formative assessment data that designers and teachers can use to improve learning processes and outcomes?

Discussant

As discussant **Allan Collins** will consider the presented projects to elaborate on a number of volatile questions: Are people becoming smarter and what are they learning by working in a mass collaborative environment? What are the other gains and losses from participation in such environments? Are people thinking more deeply about important matters or are they flitting from topic to topic in the shallows that Nicolas Carr (2010) bemoans? With the anticipated controversial discussion, the symposium will offer an opportunity to reflect on the direction and forces driving contemporary CSCL research.

References

- Brennan, K., Resnick, M., & Monroy-Hernandez, A. (2010). Making projects, making friends: Online community as catalyst for interactive media creation. *New Directions for Youth Development*, vol. 128, pp. 75-83.
- Brennan, K. (2009). Scratch-Ed: an online community for Scratch educators. In C. O'Malley, D. Suthers, P. Reimann, A. Dimitracopoulou (Eds.), *Computer Supported Collaborative Learning Practices: CSCL2009 Conference Proceedings* (Vol. II, 76-78). Rhodes, Greece: International Society of the Learning Sciences.
- Brown, J. S., & Adler, R. P. (2008). Minds on Fire: Open Education, the Long Tail, and Learning 2.0. *EDUCAUSE Review*, 43(1), 16-32. available at <http://www.educause.edu/ir/library/pdf/ERM0811.pdf>
- Bruner, J. (1996). *The Culture of Education*, Harvard University Press, Cambridge, MA.
- Butler, B., Joyce, E., & Pike, J. (2008). Don't look now, but we've created a bureaucracy: the nature and roles of policies and rules in Wikipedia. In Proceeding of the twenty-sixth conference on *Human Factors in Computing Systems (CHI)* (pp. 1101–1110). New York, NY, USA: ACM, 2008.
- Carr, N. (2010). *The Shallows: What the Internet Is Doing to Our Brains*. New-York: Norton.
- Collins, A., Fischer, G., Barron, B., Liu, C., & Spada, H. (2009). Long-Tail Learning: A Unique Opportunity for CSCL? In C. O'Malley, D. Suthers, P. Reimann, A. Dimitracopoulou (Eds.), *Computer Supported Collaborative Learning Practices: CSCL2009 Conference Proceedings* (Vol. II, 22-24). Rhodes, Greece: International Society of the Learning Sciences.
- Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407-428.
- Cress, U. (in press). Mass collaboration and learning. To appear in R. Luckin, Goodyear, P., Grabowski, B. Puntambekaer, B. Underwood, & Winters, N. (Eds.). *Handbook on Design in Educational Computing*. Taylor and Francis.
- Cress, U. & Fischer, G. (in press). Mass Collaboration with Social Software in TEL. In TEL-Reader.
- Cress, U., & Kimmerle, J. (2008). A Systemic and Cognitive view on Collaborative Knowledge Building with Wikis. *International Journal of Computer-Supported Collaborative Learning*, 3(2), 105-122. <http://www.springerlink.com/content/g509739lp56gk040/fulltext.pdf>
- Engeström, Y., & Sannino, A. (2010). "Studies of Expansive Learning: Foundations, Findings and Future Challenges," *Educational Research Review*, 5(1), 1-24.
- Fischer, G. (2011). "Understanding, Fostering, and Supporting Cultures of Participation," *ACM Interactions XVIII.3* (May + June 2011), pp. 42-53. <http://l3d.cs.colorado.edu/~gerhard/papers/2011/interactions-coverstory.pdf>
- Forte, A. & Bruckman, A. (2006) From Wikipedia to the classroom: exploring online publication and learning. *Proceedings of the International Conference of the Learning Sciences*, Vol 1. Bloomington, IN, pp. 182-188.
- Forte, A. & Bruckman, A. (2009). Citing, writing and participatory media: wikis as learning environments in the high school classroom. *International Journal of Learning and Media*, 1, 4, 23-44.
- Halatchliyski, I., Moskaliuk, J., Kimmerle, J., & Cress, U. (2010). Who integrates the networks of knowledge in Wikipedia? *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*. New York: ACM Press.
- Jenkins, H. (2009) *Confronting the Challenges of Participatory Cultures: Media Education for the 21st Century*, MIT Press, Cambridge, MA.
- Kafai, Y. & Resnick, M. (2000). *Constructionism in Practice: Designing, Thinking, and Learning in a Digital World*. Mahwah, NJ: Erlbaum.

- Kafai, Y. & Peppler, K. (2011). Beyond Small Groups: New Opportunities for Research in Computer-Supported Collective Learning. Paper published in the proceedings of the 2011 Computer-Supported Collaborative Learning (CSCL) Conference.
- Kimmerle, J., Cress, U., & Held, C. (2010). The interplay between individual and collective knowledge: Technologies for organisational learning and knowledge building. *Knowledge Management Research & Practice*, 8, 33-44.
- Kriplean, T., Beschastnikh, I., McDonald, D.W., and Golder, S. (2007) Community, Consensus, Coercion, Control: CS*W or how policy mediates mass participation. in *Proc. GROUP: ACM Conference on Supporting Group Work*.
- Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life (learning in doing)*. Cambridge: Cambridge University Press.
- Maturana, H. R, & Varela, F. J. (1987). *The tree of knowledge: The biological roots of human understanding*. Boston, MA, US: New Science Library/Shambhala Publications.
- Monroy-Hernandez, A. (2012). *Designing for Remixing: Supporting an Online Community of Amateur Creators*. PhD dissertation, MIT Media Lab.
- National-Research-Council (2009) *Learning Science in Informal Environments — People, Places, and Pursuits*, National Academy Press, Washington, DC.
- Nonaka, I. & Takeuchi, H. (1995). *The knowledge creating company: how Japanese companies create the dynamics of innovation*. New York: Oxford University Press
- O'Reilly, T. (2006) *What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software*, available at <http://oreilly.com/web2/archive/what-is-web-20.html>.
- Paavola, S., Lipponen, L. & Hakkarainen, K (2004). Models of Innovative Knowledge Communities and three metaphors of learning. *Review of Educational Research*, 74, 557-576. <http://rer.sagepub.com/content/74/4/557>
- Resnick, M., Maloney, J., Monroy-Hernandez, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for All. *Communications of the ACM*, vol. 52, no. 11, pp. 60-67.
- Roschelle, J. & Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In O'Malley, C.E., (ed.), *Computer Supported Collaborative Learning* (69-97). Springer-Verlag, Heidelberg. (1995)
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.
- Swarts, J. (2009), The collaborative construction of "fact" on Wikipedia., in Brad Mehlenbacher; Aristidis Protopsaltis; Ashley Williams & Shaun Slattery, ed., 'SIGDOC', ACM, pp. 281-288 .
- Tapscott, D., & Williams, A. D. (2006) *Wikinomics: How Mass Collaboration Changes Everything*, Portofolio, Penguin Group, New York, NY.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Scripting and Orchestration: Recent Theoretical Advances

Organizers

Frank Fischer, Ludwig Maximilian University, Munich, Germany Frank.Fischer@psy.lmu.de
 Jim Slotta, University of Toronto, Toronto, Canada. jslotta@gmail.com

Discussant

Clark Chinn, Rutgers University, clark.chinn@gse.rutgers.edu

Abstract: This symposium brings together four research groups that have been working on advancing theoretical models of guidance for CSCL. The models share the emphasis on scripting and orchestration but vary in terms of their specific focus, grain size of collaboration, and the nature of learning activities they address. The goals of this symposium are, first, to present recent advances in theorizing guidance from a scripting and orchestration perspective, and, second, to discuss commonalities, differences and future trajectories for theory development on guidance for CSCL. The audience will be interactively involved by using technologies for knowledge building. These technologies allow the audience to contribute challenging cases, questions, and ideas for studies during and after the presentations via laptops and smart phones. To approach the second goal, the discussant will draw from these contributions and involve the audience and the presenters in refining and synthesizing the ideas in a final synthesis discussion.

Introduction

In computer-supported collaborative learning, learners are often working together on complex problems requiring them to conduct inquiries and design complex artifacts. There is a stockpile of empirical evidence that activities such as problem-solving, inquiry and design support deep learning, particularly when the learner is scaffolded (Hmelo, Holton & Kolodner, 2000; Linn & Eylon, 2006; Quintana, Reiser, Davis et al., 2004; Schauble, Glaser, Schulze & John, 1995). Without guidance, many learning designs would transpire in a way that diverges from the design, making it difficult to test theoretical conjectures, or build constructive models. Guidance can be provided by different sources (e.g., teacher, computer, peer), on different social levels (individuals in groups, small groups, class, communities) and with different types of scaffolding (channeling, prompting, hinting, etc). There is now an impressive number of conceptual and empirical papers on the topic of guidance in CSCL. Still, apart from the grand theories of Vygotskij and Piaget, there has been limited progress in the development of new theories and theoretical models of guidance in CSCL (i.e., to enable and support collaborative learning).

In this symposium, we bring together 4 research groups that have been working intensively on developing theories or theoretical models of guidance for CSCL. The work from these groups varies in terms of its grain size of collaboration, the nature of learning activities it investigates, and to the extent to which they explicitly conceptualize the role of the teacher. Indeed, not all contributors would say they are explicitly advancing a “theory,” although all are advancing explicit theoretical positions in the form of models or frameworks. The goals of this symposium are (1) to present recent advances in theorizing guidance for CSCL from a scripting and orchestration perspective, (2) to discuss commonalities and differences between the approaches and future trajectories for theory development on guidance for CSCL.

Overview of contributions

The symposium includes 4 contributions, each elaborating on one theoretical approach of scripting. They have in common that they are aiming at explaining and improving guidance for CSCL. The Dillenbourg paper on orchestration offers a historical perspective on instructional guidance in CSCL that identify the intellectual roots of the research field to help locating the current approaches to scripting and orchestration in a bigger picture. Moreover, the paper sketches a model of orchestration in which instructional guidance is seen as an organic and multi-level process to optimize the classroom beyond cognitive learning, with respect to multiple constraints that exist in real world classrooms. In this process, external collaboration scripts are only one of several aspects and the teacher plays an eminent role in aligning the aspects before and during the lesson. The paper by Fischer et al. on the script theory of guidance addresses how internal collaboration scripts (i.e. knowledge on collaboration) develop through participation in CSCL practices, and how these practices are, in turn, shaped by the internal collaboration scripts of the participating group members. With respect to guidance, external collaboration scripts (i.e. instructional support for collaboration) are seen as a set of scaffolds facilitating participation in CSCL practices through activation of internal collaboration script components that would not have been spontaneously transferred into the CSCL practice. The Tchounikine paper on appropriation focuses on the dangers but also the chances that come with the phenomenon that learners are not interacting with external scripts and tools as they are designed and intended by the designer but as the learner understands them and as they increasingly

integrate them in their existing cognitive tool kit, based on their goals and on their prior experiences. The Slotta paper on the knowledge community and inquiry model has its main focus on the dynamic interplay of individual cognition and a developing collective knowledge base. With respect to instructional guidance, he advances a scripting perspective to connect the individual and small group level to the level of collective epistemic activities.

Our discussant, Clark Chinn, will identify overarching issues and important lines of future research in CSCL, presenting an interpretation of the presented approaches.

Interaction with audience

An interesting challenge for the CSCL conference, posed by Jeremy Roschelle, Roy Pea and others, is to show better progress within our conference meeting of the kinds of interactions we would advocate for collaborative learning in our research. Over time, our community should offer a wider range of experiences to the conference audience (both those who attend and those who may need to participate from elsewhere). Using “computer supported” technologies -- from audience response systems to wikis, to ubiquitous computing and augmented reality -- we should enable real time contributions and rich discussions or “knowledge work” to be performed. This is not an easy transition for a research community to make, particularly when it involves new forms of interaction that are only beginning to solidify. However, if there is any community that should rightfully explore such interactions it is CSCL. To that end, we will integrate a new form of knowledge building technology developed by Slotta and his team at the University of Toronto, to support audience engagement and dynamic discussion. Anyone in the audience with a laptop, smart phone or tablet (iPad, etc) will be able to log on during and after the presentations and contribute challenging cases, questions, and ideas for further research. The discussant will draw from these contributions and involve the audience and the presenters in refining the ideas in a final synthesis discussion.

Paper 1. Mediating, Scripting, Orchestrating: the Evolution of Instructional Design in CSCL

Pierre Dillenbourg, Ecole Polytechnique Federale de Lausanne, Switzerland, pierre.dillenbourg@epfl.ch

We rarely use the terms “instructional design” in CSCL, even though much of our work is about designing learning activities. I argue that, through its evolution, CSCL enriched instructional design with several concepts. I illustrate this evolution with the design of argumentation activities.

The foundational idea of CSCL is actually that collaboration can be ‘designed’ (Roschelle, 1990): communication tools do not simply transport messages between learners but shape their interactions. This mediation effect is obvious when learners seat in front of the same computer, i.e. when computers are not justified by data communication. For instance, CSCL scholars designed graphical editors providing learners with argumentation primitives such as ‘hypothesis’, ‘evidence’, etc. (Suthers et al, 1995). The design of these tools indeed constitutes a piece of instructional design since the interactions induced by the interface are expected to trigger specific cognitive processes, which is the craft of instructional design. This type of design nonetheless constitutes an expansion of what was understood as instructional design in the eighties; it is a more subtle or indirect way to induce targeted interactions, closer to the way ‘design’ is understood outside education. In the next decade, CSCL produced stronger forms of interactions shaping, namely ‘scripts’: the learners had to follow a sequence of interactions steps considered as - to reuse the same example- necessary for a well-formed argumentation (Weinberger et al, 2002). If Suzanne produces a claim, John would be prompted to provide counter-evidence. This approach is closer to the constraining flavor of instructional design but brings a novelty: the script is designed to be internalized by learners, which is not the case for the lesson plans produced by instructional design.

In the same decade, a third approach emerged. For triggering argumentation, one can ask peers to read texts that would provide them with conflicting evidence (‘jigsaw’ method). Alternatively, one may identify learners who have opposite opinions and then pair them for argumentation activities. These pedagogical scenarios are closer to what is produced by instructional design (Jermann & Dillenbourg, 1999). They nonetheless expand instructional design by integrating individual, collaborative and class-wide activities into a consistent workflow. Confusingly, these methods were also called ‘scripts’. We renamed macro-scripts or classroom scripts, while the scripts-to-be-internalized were renamed ‘micro-scripts’.

In the last decade, micro- and macro- scripts left research labs, they did not enter into classrooms without difficulty. The need for flexibility emerged as a key issue: since scripts introduce structure in teamwork, it may occur that this structure is too rigid or inappropriate and has to be changed on the fly. Adaptation has always been a central concept in instructional design, usually referring to pre-defined changes of learning activities based on learners’ behavior. In ‘adaptive instruction’, the design process includes the design of what can be changed during the enactment of the lesson plan. In CSCL, the need for adaptation cannot completely be anticipated. Macro-scripts are not closed environments; they stretch over multiple activities and tools (Dillenbourg &

Hong, 2008). Adaptations may be required by any event that populates the classroom ecosystem: students arriving late, discipline problems, lack of time, ... Therefore, the term ‘flexibility’ as a different flavor that ‘adaptation’: it is about empowering users, teachers and learners, in changing themselves the lesson plan. Flexibility is nonetheless an instructional design concept: since workflows are per definition rigid structures, new forms of flexibility had to be invented such as ‘team jokers’ or ‘orchestration cards’ (Dillenbourg & Tchounikine, 2007). Flexibility is still a key challenge for instructional design, both technically and pedagogically.

During this last decade, a less visible but more fundamental change has appeared. It concerns the result of instructional design. In school-oriented CSCL research, what is designed is not anymore a single activity or a piece of software but a richer set of activities or tools to be orchestrated by a teacher. In this case, instructional design includes designing the ways in which the teachers will handle complexity. The terms ‘design for classroom orchestration’ (Dillenbourg, to appear) include design concerns that were not salient before in instructional design such as minimizing the cognitive load induced by monitoring the learners, facilitating what has traditionally been referred to as ‘classroom management’, maximizing curriculum relevance, optimizing teaching time and teacher’s energy, etc. This vision stretches instructional design over new concerns such the physicality of the classroom and of the hardware. It pays attention to elements that what would in the past be disregarded as belonging to the logistics of education (e.g. time to move tables). This broader view of design constitutes a contribution from CSCL to instructional design (if one does not understand ‘instructional’ in a narrow sense). What is designed is not a single entity with intrinsic educational affordances. The design produces a set of elements that integrate into classroom to form a distributed system. In this organic view of design, the goal is not only to maximize learning gains but also to satisfy classroom constraints (time, discipline, curriculum,...) with a minimal orchestration load.

Orchestration hence addresses at guidance from a complementary angle; it is not about how much guidance or structure the teams need, but how much guidance the teacher is able provide, given his or her heavy load. It takes into consideration My ‘petite theory’ of orchestration includes two models. The first model describes pedagogical scenarios or scripts as multi-plane and multi-layer activity graphs. These vertices of the graph are the learning activities. The term “multi-plane” describe the fact that these activities may occur at six levels, named “planes” by Vygostky: individual, team, class, periphery, community and world. The edges of the graph are links between activities. The weight of a link $[a_i a_j]$ is the probability that an a_j can be conducted despite the fact that learners have not completed a_i . Examples of edges are dataflow, prerequisites, motivation didactic elicitation, ... The flexibility of a graph depends of the link weights and can be defined as the effort necessary to transform an activity graph on the fly. A notation system has been developed to represent these graphs (Figure 1).

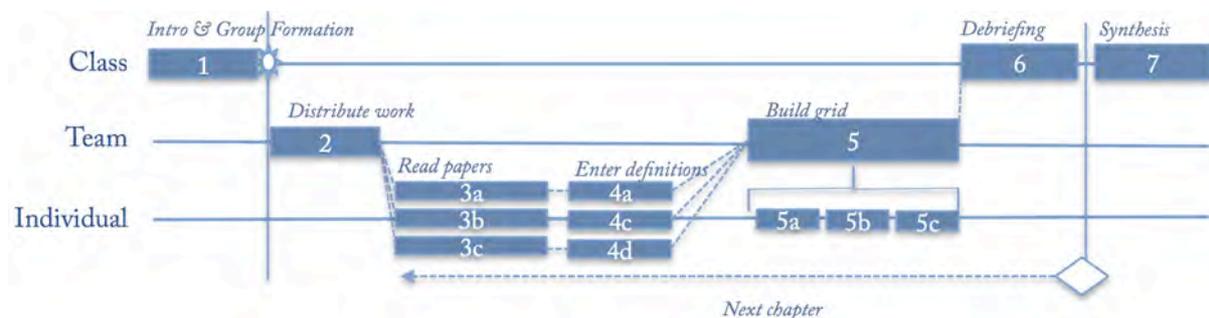


Figure 1. Representing the activity graph of a macro-script

The term ‘multi-layer’ refers to the fact that graph does not only exist as a digital structure but also as physical structure. I will present some paper-based interfaces in which the teacher may adapt the graph by handling paper cards. The great flexibility of paper reminds the study from Hutchins on ‘how a cockpit remember its speed’. Surprisingly, few authors reuse his notion of distributed systems to model a classroom, while it has a great potential.

The second model borrows ideas from information theory in order to the flows of information in a classroom. For instance, when a teacher visually scans the faces of the students in the classroom, he builds a representation of their attention level, with a certain degree of uncertainty. If he or she writes on the backboard for 20 seconds, he does not update his model and the uncertainty increases. Entropy does only decrease without external energy and this energy mostly come from the teacher, who will ask for instance new questions to reduce the uncertainty of his model. This barely legal reuse of information theory inspires the design of new type of learning technologies, namely buffers, for reducing the information load that emerged from monitoring and scaffolding learners.

Paper 2. Advances in the Development of a Script Theory of Guidance for Computer-Supported Collaborative Learning

Frank Fischer, Ingo Kollar, Karsten Stegmann & Christof Wecker, Ludwig Maximilian University, Munich, Germany, frank.fischer@psy.lmu.de

In this contribution, we will present an outline of a script theory of guidance for computer-supported collaborative learning that builds on the basic distinction of internal scripts (understood as individual memory structures) and external scripts (as instructional interventions that structure collaborative learning processes). The script theory of guidance builds on two quite distinct theoretical traditions. First, it refers to recent schema-theoretic accounts that have broken down the initially quite rigid script concept into multiple components of a dynamic memory. These components are considered to be flexibly combined in response to characteristics of the situation and the individual's goals (e.g., Kintsch, 1998; Schank, 1999). Second, it builds on socio-cultural theory with its assumptions on the relation of discourse on the social plane and the development of complex cognitive skills (e.g., Vygotsky, 1978).

In the script theory of guidance, internal collaboration scripts represent knowledge on collaboration that enable learners to effectively understand and act in recurring CSCL practices (Kollar, Fischer & Hesse, 2006). CSCL practices are, in turn, shaped (but not determined) by the internal collaboration scripts of the participating individuals. In recent conceptions of the theory internal collaboration scripts are considered to be composed of the hierarchically organized components play, scenes, roles and scriptlets (Fischer, Kollar, Stegmann & Wecker, 2013). Internal collaboration scripts may change dynamically if they do not lead to successful understanding or acting in a CSCL practice. New internal collaboration script components develop as re-configurations of existing components that have been instrumental in reaching the learner's goals.

External collaboration scripts consist of scaffolds that stimulate functional or inhibit dysfunctional internal script components (Fischer et al., in press): play scaffolds, scene scaffolds, role scaffolds, scriptlet scaffolds that are designed to support the learning of internal script components and of subject matter knowledge. As a basic tenet of the theory, internal script components that already exist in the learners' memory need to be taken into account in order to foster collaboration and learning in CSCL. The script theory of guidance for CSCL can inform the design of external collaboration scripts on when external collaboration scripts can effectively be used, which type of scaffold is likely to support learners in employing functional internal script components, and how these scaffolds can be faded to increase a self-directed configuration and re-configuration of internal script components (e.g., Wecker & Fischer, 2011). The script theory of guidance for CSCL specifies seven principles on (1) how CSCL practices are shaped through situational constraints and the internal collaboration scripts of the participating learners, (2) how internal collaboration scripts develop through participation in these CSCL practices, and (3) how external collaboration scripts can support the development and application of internal collaboration script components through scaffolding of different component levels. For example, the transactivity principle states that the more a given CSCL practice requires the transactive application of knowledge, the better this knowledge is learned through participation in this CSCL practice. The optimal scripting level principle states that external collaboration scripts should have the largest effects on learning if the scaffolding is targeted at the highest hierarchical level of the internal collaboration script (e.g., the play), where components on the subordinate levels (e.g., scenes, scriptlets) are already available to the learner. Empirical studies in CSCL provide evidence for the script theory of guidance. For example, they show that well-designed external collaboration scripts enable learners to engage in CSCL practices on a level beyond what they would be able to achieve spontaneously (e.g., Schellens, De Wever, van Kehr & Valcke, 2007; Schoonenboom, 2008). Other studies found evidence for example supporting the transactivity principle (e.g., Stegmann, Weinberger & Fischer, 2011). However, more specific investigations on the validity of the seven principles are needed.

Paper 3. Towards Theoretical Bases to Acknowledge Appropriation Phenomena in CSCL

Pierre Tchounikine, Université Joseph Fourier, Grenoble, France, pierre.tchounikine@imag.fr

Many CSCL settings are based on the hypothesis that the technological enactment framework provided to students will have an influence on their behavioral and cognitive activity. This is particularly the case for scripting approaches, within which the instructional design is reified by both the script (the instructions) and the technology (the system interface or data/work-flow).

However, the activity that will emerge from the confrontation of students with the task and the technological setting may be influenced by different dimensions, and is subject to different contingences (Tchounikine 2008, Tchounikine 2011). Teachers set tasks and students interpret the specifications of the task, their subsequent activity being a more or less rational response to the task (Goodyear 2001). Activity is related to the task but also to other dimensions (e.g., students' effective motivations or perception that is developed by the students

of the script and the provided technological setting) that evolve in time, and are interrelated within systemic relations. Typically, technical functionalities and properties are not received raw but actively constituted on receipt by the user (Jones et al. 2006). The characteristics of the technological setting will be interpreted in different ways by learners, who will appropriate them, in context, according to their purposes and in terms of their own current interests or needs. Unexpected usages of educational software are frequent and are not a matter of “good” or “bad” design only (as examples: usage of chat as a means of perception for mutual presence or actions; usage of a function meant to edit a result as a “support for thinking” or, vice versa, editing of a result (elaborated via other means) with tools meant to elaborate the result, thought of and considered as a “support for thinking” and a vector for the targeted learning; change in the way the environment is used due to the evolution of motivations (and, thus, effective activity), for example from “playing the game of the pedagogic contract and using the platform to meet the teacher’s demand” to “deal with urgency and produce the expected result (whatever the means are)”. Basically, learners do not use the provided software to solve the set task, they consider some task (that may only correspond more or less to that set by the teacher) and take advantage of the means that seem best adapted to them (which may correspond to some usage of the provided software) in the context of their activity.

In contexts within which the learning setting designed artifacts (the script, the technological enactment framework) are supposed to influence the students’ activity, the fact they may be appropriated in different ways is of importance. CSCL cannot escape the theoretical issue of understanding how actors (teachers, students) appropriate these artifacts.

In line with activity-oriented works (Kaptelinin & Nardi 2006), it may be hypothesized that the constructive nature of activity impacts users’ use and appropriation of software and scripts. Designers create artifacts based on how they imagine their future usage, yet artifacts only become instruments for users in the context of these users’ activities, i.e. when and through the way they allow these users to achieve the tasks they consider in the way they consider them. When designing artifacts, we need to take care that when a user adapts an artifact as a way to adapt it to his/her activity, this adaptation may present a situated dimension, an interpretative dimension, a constructive dimension and/or an oriented dimension.

In CSCL settings the script, the teachers’ activity and the technologies (when designed as flexible technologies (Dillenbourg & Tchounikine 2007, Sobreira & Tchounikine 2012) may be adapted. The way the script and/or the technical framework may be adapted to comply with appropriation issues while remaining coherent with the script’s design rationale (i.e., why and how it is meant to support learning) must be studied in relation with an understanding of the scripting principles and effects. Moreover, adaptation issues must also be studied in relation with how teachers conducting the session may be empowered to understand and orchestrate the script’s enactment. The development of a theoretical perspective to appropriation must thus be conducted in relation with theoretical developments related to scripting and orchestration.

Paper 4. The Knowledge Community and Inquiry Model: Scaffolding individual, Collaborative, and Collective Activities

Jim Slotta, University of Toronto, 252 Bloor St. West, Toronto, Canada, jslotta@gmail.com

This paper advances Knowledge Community and Inquiry (KCI) as a pedagogical model that guides the design of complex inquiry curriculum that includes individual, collaborative (i.e., small group) and collective (whole class or multiple class) activities. In KCI (see Figure below), all individuals and groups work together to produce a collective knowledge base that serves as a resource for subsequent inquiry activities. The curriculum is several weeks or months in duration, and includes technology-enhanced materials, tools, and virtual environments that scaffold the various designed interactions.

In contrast with other “knowledge community” approaches (e.g., Scardamalia & Bereiter, 2006), KCI includes scripted inquiry that is carefully designed to address the science learning goals. Thus, while KCI is theoretically committed to the notion of collective epistemology, it has one foot planted firmly in the theoretical space of scaffolded inquiry (Slotta & Peters, 2008). The model is inspired partly by “Web 2.0” environments, including content communities (e.g., Wikipedia, YouTube) and social networks (Facebook), which are increasingly familiar to students, teachers and researchers (Slotta & Najafi, 2012, Figure 2). So, while KCI attempts to get students working together (“one for all”), it also makes the advantages of that collective achievement accessible to individual or small group inquiry (“all for one”).

KCI curriculum is developed through a sustained co-design effort including teachers, researchers, technology developers, and interaction designers. The designed artifact constitutes a “script” that includes real-time decisions or assignments made by intelligent agents, and involves student-contributed content, social tagging, learning in ubiquitous and distributed contexts, and a wide range of individual and collaborative (i.e., small group) scripted activities. The script must be “orchestrated” by the teacher, who is greatly enabled by the technology-environment. Below, I describe a set of design principles that guide our creation of individual, collaborative, cooperative and collective inquiry activities, and how those activities are scaffolded.

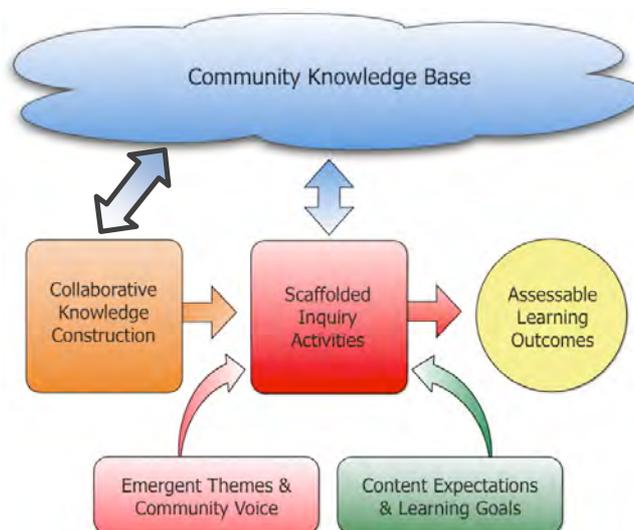


Figure 2. Knowledge Community and Inquiry Model: Basic Processes and Constraints

Principle 1. Students work collectively as a knowledge community, creating a knowledge base that serves as a resource for their ongoing inquiry within a specific science domain. Students are scaffolded to work collectively (i.e., all students working in parallel, building on one another’s contributions, as in wiki editing). It is not an ill-defined task, however, for students to build any knowledge base that appeals to them (i.e., as in discovery oriented inquiry); rather, the science content expectations are used as an explicit framework or index that scaffolds the collaborative construction of a relevant and accessible knowledge base. In a recent curriculum on climate change, we established a knowledge base in the form of a wiki, with blank page “templates,” where each new page would be a major climate change issue. Students were responsible for coming up with the issues, but each issue page included 5 sub-headers that were not negotiable, corresponding to the 5 scientific concepts of the curriculum unit: carbon sinks/sources; greenhouse gases; energy currents in ocean and atmosphere; scientific models and forecasts; and remediation efforts

Principle 2. The knowledge base is accessible for use as a resource as well as for editing and improvement by all members. In the climate change unit described above, students came to understand that the knowledge base could serve as a collectively constructed and validated resource for their community. Notably, 5 class sections (n= 121) of a high school biology course collaborated in creating a single common wiki, using our blank template pages. The result was 14 major issue pages, each averaging 3509 words, with 90 page revisions made by 8 authors. This is a rarely observed level of collaborative knowledge production, and demonstrated to researchers and teachers alike that students enjoyed the task and excelled at it.

Principle 3. Collaborative Inquiry activities are designed to address the targeted science learning goals, including assessable outcomes. Inquiry is seen as a process where individual learners build on their existing ideas to develop scientific understandings (e.g., Linn & Eylon, 2006). Students work individually and in small groups, making use of the community knowledge base as a key resource. We designed a “climate change remediation” assignment, where groups (size n=2, 3 or 4) collaboratively designed a remediation that addressed at least 3 of the issues in the knowledge base – targeting specific geographical regions of Canada. One requirement was to make explicit connections to each of the 5 major scientific elements listed above (ie, assessable outcomes). Just as the collaborative knowledge construction phase was scaffolded through the use of wiki page templates, this small group inquiry project was carefully scaffolded through the design of structured page templates and instructions (e.g., “make sure that you refer to at least 3 of the biodiversity issues”).

Principle 4. The teacher's role must be clearly specified within the inquiry script, but also include a general orchestration role. In KCI, the teacher's role is that of an expert collaborator or mentor, responding to student ideas as they emerge, and orchestrating the pedagogical flow of activities. Teachers are not just a vague "guide on the side" – an image or assignment that often paralyzes teachers or leaves them "sidestreamed" by the overly scaffolded learning environments, which often include no explicit role for the teacher. KCI curriculum includes specific, scripted interactions with students or responses to materials, such as providing feedback and making "consequential" orchestrational decisions based on the content of student interactions and artifacts. For example, in a recent physics activity, we engaged students in solving ill structured problems in a smart classroom setting, where they worked collaboratively at various stations within the room (e.g., inputting their tags and votes on personal tablet computers, then using collaborative Smartboard (i.e., interactive whiteboard) activities to negotiate a consensus). After setting up the problems for solution, the group "submitted" their readiness, and the teacher's his own tablet) was notified that a group was ready for a debriefing encounter. The teacher then walked over to that group's station and consulted with them about their problem set-up. If (and only if) they had done a sufficiently detailed and accurate job, the teacher touched a "go ahead" button in his own tablet, and the student tablets were all refreshed with new tasks for the next scaffolded activity in the script.

The final example above illustrates the current focus of our research, with regard to scaffolding technologies as well as inquiry designs. Through painstaking design (which can take up to a year's time) we develop substantive curricula – typically whole semester or multi-week courses, carefully developing the script according to the KCI principles above. Once we have the pedagogical script, we design and develop all materials and scaffolding for activities at the individual, small group and whole class levels. For example, a student's personal tablet computer (e.g., iPad) could guide him or her to the appropriate location in the room, or to a particular small group, and then solicit, guide and collect particular observations or reflections. Internet-based software collects all student contributions, such that intelligent agents can perform real-time data mining to compile an aggregated dataset which can be presented on tablets or Smartboards, or both. Students respond to this emergent view, perhaps recognizing cells in a table where they disagree or need more data. Teachers can be prompted on their own personal tablets for a variety of interactions. Through such intricately designed and scaffolded activities, we seek to create a sense of autonomy, creativity and inquiry, without "overscripting."

References

- Dillenbourg, P. & Tchounikine, P. (2007). *Flexibility in macro-scripts for CSCL*. *Journal of Computer Assisted Learning*, 23(1), 1-13.
- Dillenbourg, P., & Hong, F. (2008). The mechanics of CSCL macro scripts. *International Journal of Computer-Supported Collaborative Learning*. 3 (1), pp. 5-23
- Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro CSCL scripts. *Journal of Computer Assisted Learning*. Vol. 23 (1), pp. 1-13.
- Dillenbourg, P., Jarvela, S. & Fischer, F (2009). The evolution of research on computer-supported collaborative learning: from design to orchestration. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder & S. Barnes (Eds) *Technology enhanced learning: Principles and Products*. (pp. 3-20). Springer.
- Dillenbourg, P. (to appear) Design for Classroom Orchestration . *Computers & Education*
- Fischer, F., Kollar, I., Stegmann, K. & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, 48(1), 56-66.
- Goodyear, P. (2001). Effective networked learning in higher education: notes and guidelines. Volume 3 of the Final Report to JCALT: Networked Learning in Higher Education Project. Retrieved from <http://csalt.lancs.ac.uk/jisc/> (December 2nd, 2006).
- Jermann, P., & Dillenbourg, P. (1999). An analysis of learner arguments in a collective learning environment. C. Hoadley et J. Roschelle (Eds.), *Proceedings of 3rd Computer-Supported Collaborative Learning Conference*, pp. 265-273, Stanford .
- Jones, C., Dirckinck-Holmfeld, L. & Lindström, B. (2006). A relational, indirect, meso-level approach to CSCL design in the next decade. *International Journal of Computer-Supported Collaborative Learning (1)*, 35-56.
- Hmelo, C.E., Holton, D.L., Kolodner, J.L. (2000). Designing to Learn about Complex Systems. *Journal of the Learning Sciences*, 9 (3), 47-298.
- Kaptelinin, V. & Nardi, B. (2006). *Acting with Technology: Activity Theory and Interaction Design*. Cambridge: MIT Press.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts - a conceptual analysis. *Educational Psychology Review*, 18(2), 159-185.
- Linn, M. C., & Eylon, B.-S. (2006). Science Education: Integrating Views of Learning and Instruction. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of Educational Psychology* (2nd ed., pp. 511-544). Mahwah, NJ: Lawrence Erlbaum Associates.

- Quintana, C., Reiser, B.J., Davis, E.A., 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.
- Roschelle, J. (1990) *Designing for Conversations*. Paper presented at the AAAI Symposium on Knowledge-Based Environments for Learning and Teaching, Stanford, CA, March 90.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York, NY: Cambridge University Press.
- Schank, R. C. (1999). *Dynamic Memory Revisited*. Cambridge, MA: Cambridge University Press.
- Schauble, L., Glaser, R., Duschl, R. A., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *Journal of the Learning Sciences*, 4, 131-166.
- Schellens, T., Van Keer, H., De Wever, B., & Valcke, M. (2007). Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups? *International Journal of Computer-Supported Collaborative Learning*, 2(2-3), 225-246.
- Schoonenboom, J. (2008). The effect of a script and a structured interface in grounding discussions. *International Journal of Computer-Supported Collaborative Learning*, 3, 327-341.
- Slotta, J. D., & Peters, V. L. (2008). A blended model for knowledge communities. Proceedings of the Eighth International Conference for the Learning Sciences – ICLS 2008 (pp. 343-350).
- Slotta, J.D. & Najafi, H. (2012). Supporting Collaborative Knowledge Construction with Web 2.0 Technologies. In *Emerging Technologies for the Classroom: A Learning Sciences Perspective* (N. Lavigne, Ed.).
- Sobreira, P. & Tchounikine, P. (2012) A model for flexibly editing CSCL scripts. *International Journal of Computer-Supported Collaborative Learning* 7 (4).
- Stegmann, K., Weinberger, A., & Fischer, F. (2011). Aktives Lernen durch Argumentieren: Evidenz für das Modell der argumentativen Wissenskonstruktion in Online-Diskussionen {Active Learning Through Argumentation: Evidence for a Model of Argumentative Knowledge Construction in Online Discussions}. *Unterrichtswissenschaft*, 39 (3), 231–244.
- Suthers, D., Weiner, A. Connelly J. & Paolucci, M. (1995) Belvedere: Engaging students in critical discussion of science and public policy issues. In J. Greer(Ed). Proceedings of the International Conference in Artificial Intelligence in Education, Washington, August 16-19, pp. 266-273.
- Tchounikine, P. (2011). *Computer Science and Educational Software Design. A resource for multidisciplinary work in Technology Enhanced Learning*. New York: Springer.
- Tchounikine, P. (2008). Operationalizing macro-scripts in CSCL technological settings. *International Journal of Computer-Supported Collaborative Learning* 3(2), 193-233.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological functions*. Cambridge: Harvard University Press.
- Wecker, C. & Fischer, F. (2011). From guided to self-regulated performance of domain-general skills: The role of peer monitoring during the fading of instructional scripts. *Learning and Instruction*, 21 (6), 746-756.
- Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior*, 26(4), 506-515.
- Weinberger, A., Fischer, F. & Mandl, H. (2002). Fostering computer supported collaborative learning with cooperation scripts and scaffolds. In G. Stahl (Ed.), *Computer support for collaborative learning: foundations for a CSCL community* (pp. 573-574). Proceedings of the Conference on Computer Support for Collaborative Learning (CSCL) 2002, Boulder, CO.

Are CSCL and Learning Sciences research relevant to large-scale educational reform?

Nancy Law, University of Hong Kong, Pokfulam Road, Hong Kong. nlaw@hku.hk

Naomi Miyake, University of Tokyo, Bunkyo, Tokyo, Japan, nmiyake@p.u-tokyo.ac.jp

Chee-Kit Looi, National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616, cheekit.looi@nie.edu.sg

Riina Vuorikari, European Schoolnet, vuorikari@gmail.com

Yves Punie, European Commission Institute for Prospective Technological Studies (IPTS), Inca Garcilaso, s/n 41092 Seville, SPAIN, Yves.PUNIE@ec.europa.eu

Discussant: Marcia Linn, Graduate School of Education, University of California, Berkeley, 4611 Tolman Hall, CA 94720-1670, USA, mclinn@berkeley.edu

Abstract: Many scholars have contributed efforts to improve education in schools. A major motivation for learning scientists to develop design research as a methodology is to contribute to theory and educational practice through rigorous research without avoiding the complexities and messiness in authentic educational settings. There are many examples of successful implementation of collaborative, knowledge-construction oriented pedagogies using socio-cognitive and socio-metacognitive tools in formal and informal educational settings as well as in teacher professional development. However, there are many challenges to scaling up such innovations beyond small-scale implementation, including that of developing into “fatal mutations” (Brown, 1992). This symposium provides an opportunity for discussion and reflection on the impact that CSCL and Learning Sciences researchers have made on large-scale education reform and what, if any, may be done to extend this impact by bringing together a set of papers describing some large-scale education innovation initiatives in Asia and Europe.

Introduction

Many of the CSCL and Learning Sciences researchers work in Faculties of Education, and they often conduct their research in naturalistic classroom settings rather than in special experimental arrangements. This is particularly the case for researchers engaged in design-based research. There have been many advances in learning theories and learning technologies in the past two decades resulting from such research (e.g. Sawyer, 2006), but are these developments making impact on educational practice at large? Collins and Halverson (2009) arrive at the conclusion that the transformative potential of ICT for education is not likely to have impact on publicly funded education because of the inherent conservative nature of these schools. Studies of large scale education reforms and scaling up of innovations have mainly been the concerns of researchers in the field of educational administration and management (e.g. Fullan, 2008, 2010; Hargreaves and Fink, 2012), and much of that literature does not pay specific attention to the learning theories or technologies underpinning the changes involved other than as a contextual variable. Arguably, it is this latter literature that has so far been much more successful in capturing the attention of school leaders and education policy makers, rather than the work of the learning scientists.

Immediately preceding the CSCL 2011 conference in Hong Kong, CITE (the Centre for Information Technology in Education at the University of Hong Kong) took advantage of this global gathering of to bring together top learning scientists and policy leaders in a forum on how to restore learning as the core of education policy concerns, and to make sure fore-running research results on learning will inform policy-making and impact education (<http://backtolearning.cite.hku.hk/>). There was agreement at this forum that while there is much that research on learning has to contribute to the focal concern of current education policy to nurture 21st century skills in learners, the impact of such research on the practice of education professionals or on the understanding of the wider community about education is still very limited. This symposium is organized to explore the following questions:

- (1) How relevant is CSCL and Learning Sciences research to large-scale education reform?
- (2) What unique contributions can research on learning make to the sustainability and scalability of ICT-supported learning innovations in schools?
- (3) Are there ways through which the CSCL and Learning Sciences community can increase their social and professional impact?

The symposium presenters have all engaged in studies on the implementation and scaling up of research-informed ICT-supported learning innovations in mainstream school education in Asia and Europe, while the discussant has similarly rich experience in the US. They will share with participants their insight on the above questions based on their work.

Paper 1: From e-Learning Pilot Scheme to Scalable e-Learning Innovations: Wishful thinking or reality?

Nancy Law and Yeung Lee

University of Hong Kong

nlaw@hku.hk, yeunglee@hku.hk

The Hong Kong Education Bureau launched a three-year e-Learning Pilot Scheme in September 2011 with the aim to identify good models of integrating ICT in the school curriculum to bring about effective interactive learning, self-directed learning and/or to cater for learner diversity, and to build models of change conducive to the sustainability and scalability of the innovations piloted. To this end, an evaluation project, both formative and summative in nature, was commissioned to start when the scheme was launched to identify if the intended goals were achieved and to summarize the lessons learnt. At the time when the CSCL 2013 Conference takes place, this scheme would have completed its second year of operation. This paper draws on data collected from the evaluation study to explore whether, and for what reasons, this e-Learning Pilot Scheme can be a successful mechanism for system-wide, scalable e-Learning innovations.

All publicly funded primary and secondary schools were invited in May 2010 to submit innovative e-Learning project proposals for funding. The goal was to make use of the pilot schools as test beds to help the Bureau to develop, try out and evaluate when and how e-Learning works best (i.e. using the scheme to build up knowledge about successful e-Learning pedagogical models) No specific learning theory or model of e-Learning was prioritized nor discouraged—as long as the project was learning focused and student-centered, and the call did not reference any local or international experiences.

Another objective of the pilot scheme was to build knowledge about sustaining ICT-enabled learning innovations. An underpinning project assumption is that engagement and support from the private sector (e.g. e-Learning resources providers, publishers, learning technology companies and Internet service providers) is of critical importance in order to evolve a viable business model for e-Learning. Hence all submitted proposals must include some partnership arrangements with the private sector, and schools were also encouraged to partner with other organizations such as tertiary institutions, NGOs, etc.

Altogether 21 pilot proposals were selected for funding, involving a total of 61 primary, secondary and special education needs schools (details from <http://edbsdited.fwg.hk/e-Learning/eng/index.php?id=3>). Because of the atheoretical stance taken by the scheme on learning, the pilot projects were selected to achieve maximum variations in school and curriculum contexts as well as in partnership arrangements. Some of the pilot projects simply focused on developing graded, self-accessed learning materials while others target the development of inquiry, collaborations and information literacy skills. Twelve of the projects involve collaboration among two or more schools while the other nine involve a single school each.

Due to the lack of a common pedagogical theory underpinning the different pilot projects and the large diversities in the innovation foci and school contexts, a generic, multilevel framework was developed by the project evaluation study to conceptualize how ICT-using pedagogical practices contribute to students' learning outcomes within the bigger context of overall pedagogical practices found in schools, which are in turn influenced by the teachers' characteristics as well as school and system level factors. Indicators for each of the identified contextual factors were developed to chart how these influence (1) the effectiveness of an e-Learning pilot project in enhancing students' information literacy and self-directed learning skills, and (2) the sustainability and scalability of the innovation. These indicators, both quantitative and qualitative, are derived from data collected at four levels: classroom, school, project and system levels. Quantitative data are collected through surveys to principals, ICT coordinators, teachers and students in the pilot schools. Qualitative data include interviews with different stakeholder groups and documentary records of the pilot projects such as the project proposals and project annual reports. The most important source of data to shed light on the ICT-using pedagogical practice and the associated students' learning outcomes at the classroom level was collected through two instruments, to be submitted by a nominated teacher from each project on one curriculum unit of their choice: (1) a curriculum design cover sheet to describe the targeted learning outcome(s), the curriculum activity(ies) designed, the role of technology in the process and how these are connected, and (2) samples of students' authentic work generated during the course of the curriculum unit that can demonstrate different levels of outcomes achieved (high, middle and low) in the areas of information literacy and self-directed learning. These two instruments were developed on the basis of similar instruments used in the Microsoft Innovative Schools Program, (Shear et al., 2009)).

End of year 1 evaluation results reveal that most of the curriculum examples submitted by teachers are largely very traditional and content focused, and the samples of students' work collected generally show little evidence of information literacy or self-directed learning skills being exercised. For the few cases where such outcomes were evidenced, some common characteristics the associated e-Learning pedagogical practices were

observed: the students had direct access to use ICT for tasks that had some levels of openness, and they had opportunities to observe the work of peers and to receive feedback.

While it is not possible yet to draw conclusions on the sustainability and scalability of the 21 e-Learning pilot projects being evaluated after just one year of operation, we find large diversities in the progress made in project implementation and the extent to which pedagogical changes were observed in the process. Only in a few of the pilot projects were rapid cycles of learning and advances in pedagogical designs and ICT use observed. In all these cases, the projects have built-in organizational infrastructures (e.g. co-planning teams, peer observations of teaching and debriefing sessions) to facilitate and scaffold interaction, communication and sharing of ideas among teachers and the leadership team. Changes in practice are most evident in those cases where there are mechanisms to make adjustments and changes to school and/or classroom routines such as timetabling, staffing or resource allocation priorities as discussed in Spillane, Parise and Sherer (2011). This presentation will reflect on the findings, particularly on the apparent lack of progress at the system/policy level in learning about what constitute the primary pedagogical characteristics of e-Learning practices that foster 21st century skills or what features of innovation implementation would be conducive to scale and sustainability, despite the many reform efforts implemented over the past 15 years.

Paper 2: Restoring “how people learn” as the core of educational reform in Japanese classrooms

Naomi Miyake

University of Tokyo

nmiyake@p.u-tokyo.ac.jp

In Japan, there have been a good number of educational reforms utilizing collaborative learning, based on many different “theories” of how people learn. Some have evolved from strong beliefs about learners’ self-construction of “hypotheses” about what they experience, imagine and think, and to discuss these with classmates through carefully ordered series of scientific problem solving, often by observing sequences of experiments (Itakura, 1971, Hatano and Inagaki, 1991). There is also a long history of creating learner-centered practices, some of which, like the “lesson study” movement (e.g. <http://www.wals2011.com/>) and the Japanese Association for the study of Cooperation in Education (<http://jasce.jp/indexe.html>), have attracted international attention. Yet these often lack direct conversation with policy makers, which hinder their expansion beyond certain points in their scale.

The University of Tokyo launched in 2009 an initiative strongly grounded in the learning sciences to contribute to this movement of renovating Japanese education with two important strategic orientations. One is to bring university research closer to policy makers at the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and in boards of education throughout Japan, so that what has been researched and developed in universities would have direct influence on schools. The other is to base the renovation on the conversation between universities and business sectors, so that the needed educational reform could be supported by the entire society. To implement this, the Consortium for Renovating Education of the Future (henceforth CoREF; <http://coref.u-tokyo.ac.jp/>) was founded at the University of Tokyo, in conjunction with one city-level board of education and four other leading universities. This project is one of the rare cases where learning science research is adopted seriously to guide the renovation in classroom practices using a concrete framework (Miyake, in press) and with the joint efforts of regional boards of education. This initiative is also unique in spanning all subject areas taught at all school levels, including vocationally oriented tertiary education (high school level). As the administrative leaders at the school and board of education levels as well as the participating teachers are learning the sciences of how people learn, not in abstract forms but through implementing them in classrooms, they are also building up the capacity to scale up the initiative by themselves. Promising outcomes have been observed in the better designed classes in developing students’ 21st century skills: higher levels of learning gains, higher motivation to extend learning outside of school into homes, and a stronger sense of “learning” among the students.

Relevance of CSCL and learning sciences

CoREF has started to work with prefectural and city/town boards of education, to develop learner-centric teaching curricula using a concrete form of collaborative learning named Constructive Jigsaw (Miyake, in press), based on research findings on how people change their concepts (c.f. Vosniadou, 2008; Sinatra & Pintrich, 2003). The key focus of the renovation is to change teacher practices from being teacher-centric to learner-centric, by working collaboratively with teachers to create a new set of curricular that allows learners to learn in collaborative, knowledge-constructive ways. Another focus is to work with members of the boards of education from the start to research on ways to support the renovation, and to share experiences with other boards of education through networked support systems for both novice and experienced teachers. A further focus is to enhance the natural adoption of ICT. The teachers involved in this project are learning quickly to

take advantage of being networked, and are encouraged to identify uses of ICT natural to the learners that would enhance the quality of learning in the new curricular.

This reform started in 2000, involving 3 prefectural boards of education covering some 300 high schools and 18 city/town education boards covering some 80 elementary and middle schools. More than 600 teachers have developed some curricula, with some also having changed the main part of their practices into the new pedagogical form. The developed curricula cover almost all subject areas in all school types, including language art, math, science, humanities, English as second language, art, music, as well as some part of special education. Encouragement through the project has prompted participating teachers to actively exchange the teaching plans and teaching materials they developed, localizing them to fit each class context, and to co-examine the class activities and outcomes. Because these collaborative efforts take place across different boards of education, Internet connectivity plays a critical role.

Assessment and Outcomes

The key learning outcomes targeted in this renovation are assessed using three criteria: outcome portability, dependability and sustainability (Miyake, et al., 2007). Concretely, each individual student is assessed on whether s/he is able to (1) express their understanding of the subject matter through “justification with acceptance,” (2) demonstrate confidence in collaborative knowledge construction, and (3) show increase in motivation to continue and expand what they have learned, both at the end of each class as well as through longer periods of half-a-year to two or three years. Preliminary analyses at the end of the second year show a very favorable pattern. Of the 527 elementary and 461 middle school students surveyed, more than 85% considered the reformed classes they attended as “enjoyable because we understand”, and more than 78% preferred their next class to be taught in the reformed fashion. For the 1556 high school students surveyed, the corresponding percentages are 73% and 54%. The learning outcomes achieved showed a greater variability, but their sustainability, that is, how well students can remember and reconstruct what they have learnt, hovered around 80% in cases where this information was made available to us. The teachers’ reflective comments are also in favor of this new renovation. The project has also been identified as an implementable model for senior professionals to work together with teachers to improve the curricular quality.

Unique Contributions of ICT for Scaling-up and Enhancing Social/Professional Impact

The plan for scaling-up is to form a large community to work with MEXT by networking small networks that are currently emerging among teachers, schools, and boards of education. Each network may consist of about 5 to 10 active members who share common interests and goals. A teacher or an education leader could participate in several networks, according to their motivation, needs, energy and time. Some networks may enjoy the power derived from the participation of senior professionals from industries and governments. CoREF has just started to investigate the possibility of connecting remote learners in very small schools scattered in many parts of Japan, through remotely operable robots acting as learning classmates and mediators to cloud resources. It is hoped that this network will be research-oriented, with strong learning sciences underpinning, working directly with practitioners and education policy makers.

Paper 3: Scaling up rapid collaborative practices in Singapore schools

Chee-Kit Looi

National Institute of Education, Nanyang Technological University

cheekit.looi@nie.edu.sg

One characteristic of education research in Singapore is the close partnerships between researchers and practitioner communities. The Singapore government has funded educational research at the National Institute of Education (NIE) of the Nanyang Technological University (NTU) over the past decade, including the setting up of the Learning Sciences Lab in 2004. This support is intended to not only advance the body of knowledge about designing, implementing and supporting educational innovations, but also to help inform educational policy and practices.

Since 2007, researchers from NIE have carried out a programme of research introducing rapid collaborative knowledge improvement (RCKI) practices using a technology called GroupScribbles (GS) to many schools in Singapore (Looi, So, Toh and Chen, 2011). The first school we worked with was a primary school. In the second year, the research was extended to two secondary schools. Because of the success from these research interventions, the Ministry of Education (MOE) has worked with another six schools to bring in GS as a “deep” intervention. Separately, other schools have adopted and adapted these innovations on their own.

Through the research work with schools, 109 GS lesson plans have been created in various subjects (Math, Science, English and Chinese language), and 146 GS lessons have been enacted, observed and studied. The research team has conducted numerous professional development sessions for teachers.

Two workshops were held during Jan-Feb 2010, reaching out to 50 teachers from more than 20 schools. In NIE itself, which trains pre-service teachers, the GS pedagogy was introduced to many cohorts of in-service teachers doing the leadership programmes, and to pre-service teachers taking the core Instructional Technology modules over the years.

Much CSCL and Learning Sciences research have focused on understanding or deriving design principles for learning, while others are intervention studies. There is a gap between such contextualized studies and the imperatives of large scale education reform. One approach to reducing the research-practice gap is to do implementation studies that systematically study an intervention to understand the conditions for successful implementation. Penuel, Fishman and Cheng (2011) put forward a compelling argument for a new form of implementation research termed as "Design-based Implementation Research (DBIR)", which comprises four elements: (a) a focus on persistent problems of practice from multiple stakeholders' perspectives, (b) a commitment to iterative, collaborative design, (c) a concern with developing theory related to both classroom learning and implementation through systematic inquiry, and (d) a concern with developing capacity for sustaining change at the system level. In the sister disciplines of medicine and public health, DBIR has a robust infrastructure and a clear focus on the interdisciplinary challenge of bringing about large-scale improvements to complex systems (Fixsen et al., 2005).

In our reflective analysis of the GS intervention study in Singapore schools, we have incorporated elements of DBIR, by working closely with schools to bring about the routine adoption of PCKI learning practices in the classroom. In our role as academics, we seek to identify and refine design principles and our theoretical understanding through our research. What have we learned about design principles through our iterative RCKI work in schools using a DBIR approach? Face-to-face classroom situations can host a broad variety of pedagogical patterns involving student-student and student-teacher interactions that go beyond IRE. However, the prevailing modularity of class periods, in chunks of between 40 and 90 minutes, constraints the adoption of pedagogical approaches whose characteristic timescales are measured in days or months or even years. Of particular interest to teachers and school administrators are pedagogical patterns that would carry the burden of scaffolding students to learn content (e.g. science, mathematics, language learning) as well as enhance participating students' development of so-called 21st century skills, such as communication, collaboration and critical thinking skills. RCKI refers to a collection of evidenced-based pedagogical patterns in which the learners brainstorm and contribute ideas, and build on each other's ideas to derive better ideas. It is a set of design principles that seeks to harness the collective intelligence of groups to learn collaboratively in a dynamic live setting (Looi, Chen & Patton, 2010; Wen, Looi & Chen, 2012).

We make the case that some research in the CSCL and learning sciences community must address issues of intervention and implementation study to understand the conditions for adaptations and sustainability of innovations in different contexts and settings – towards contributing to reducing the research-practice gap conundrum in education. We consider the complex interplay of multiple dimensions of education reforms, and approach our programme of research from a systemic change perspective that recognises the micro, meso, and macro levels of educational systems (Looi, 2011; Looi, So, Toh, & Chen, 2011).

Singapore's Third Masterplan of ICT in Education (mp3) seeks to enrich and transform the learning environments of students and equip them with the critical competencies and dispositions to success in a knowledge economy (mp3, 2009). While mp3 has an explicit focus on helping students develop competencies for self-directed and collaborative learning through the effective use of ICT as well as become discerning and responsible ICT users, the policies are couched from the perspective of policy makers. The challenge for school principals and teachers is how to implement technology-enabled pedagogies that foster self-directed and collaborative learning. This is where university researchers come in as meso-level actors who work with school leaders and teachers to interpret collaborative learning outcomes and processes in the context of the needs of a particular school. This re-contextualization of the pedagogic discourse is a "meso-level" mechanism. By approaching this pedagogy-driven reform at the macro, meso and micro levels, we seek the alignment of systemic forces at work to provide a buttress for sustainability. Thus we, as researchers working as the meso-level actors, help the school practitioners understand and interpret policy imperatives and translate them into classroom teaching and learning practices in ways that are informed by research and learning theories.

In many countries and regions, education authorities are keen for their reform initiatives to be well received by various stakeholders, namely: district-level leaders, school leaders, teachers, students and their parents. Typically, these stakeholders have different notions of scaling from researchers, such as holding a more top-down view of scaling and a linear progression model of research interventions. The learning sciences community can be part of this conversation towards articulating different models of

evidence-based scaling that work in different socio-political-cultural contexts and contributing towards creating more existential examples of sustainable and scalable innovations.

Paper 4: eTwinning: a European Network Community for Teachers to support cross-border school collaboration

Riina Vuorikari
European Schoolnet
vuorikari@gmail.com

This contribution presents a study into the process of scaling up eTwinning, the community for schools in Europe (www.etwinning.net). By early 2013, the eTwinning community has attracted more than 100,000 schools and close to 200,000 teachers from 33 European countries to participate in international school collaboration through the use of Information and Communication Technologies (ICT). The initiative is one of the most successful actions under the European Union's Lifelong Learning Programme in the school sector. The participation amounts to more than 3% of all primary and secondary teachers who are eligible for participation.

Core to eTwinning is an online platform that offers participating teachers three types of activities:

1. Partner finding activities to create cross-border school collaboration projects using ICT;
2. Various continuing professional development (CPD) activities such as online Learning Events (distance courses) with formal certification and some more informal ones such as Online Interest Groups and Teachers' Rooms on various topics;
3. Social networking using tools that include profile pages with personal and professional information.

In the beginning, eTwinning was supported through the Pedagogical Advisory Group (PAG) comprising experts from teacher training, school inspection and pedagogical research. Its role was to analyse, reflect and comment on the eTwinning activities, and to develop a theoretical framework to ensure the lasting pedagogic value of the eTwinning activity in schools. Now in its 9th year of operation, eTwinning has evolved from simple school collaboration projects into more complex ones that impact not only on the level of innovative pedagogical practices in the classroom (Galvin, 2009) and students' involvement (Wastiau et al., 2011), but also impinge on institutional factors and the organisational climate of the whole school (Vuorikari, 2013). eTwinning also provide many CPD opportunities to teachers through formal and informal upskilling activities, and through teacher participation in networks (Vuorikari et al., 2011, Vuorikari et al., 2012). Similar results were reported in a large external study on the impact of eTwinning (European Commission, 2013).

To complement the above-mentioned qualitative studies on eTwinning practices and their impact, a number of longitudinal studies using data extracted from the platform have also been conducted. The eTwinning Analytics framework was created on the basis of OECD's indices for teachers' co-operation (OECD, 2009) to identify emerging behaviours and patterns within eTwinning. It operationalizes various activities for measurement and monitoring purposes. For example, the activities carried out while conducting school collaboration belong to the category of *teachers' professional development*, which can be seen as enhancing teachers' professionalism and self-efficacy.

To experience a full range of professional development activities in eTwinning, and therefore to take full advantage of it, a substantial time investment is needed. From our studies, it is clear that eTwinning has a potential to engage its users over a long period of time. Evidence was found that one eTwinning member in six, who registered on the platform between 2005 and 2006, still returns to it.

Paper 5: Mainstreaming ICT-enabled innovations in Education and Training in Europe: Challenges and Opportunities

Yves Punie and Panagiotis Kampylis
European Commission, Joint Research Centre, Institute for Prospective Technological Studies (IPTS)
Yves.PUNIE@ec.europa.eu

This paper presents results from a European-wide research project, running from December 2011 to June 2013, on up-scaling ICT-enabled innovation in Education and Training (E&T), titled Up-Scaling Creative Classrooms in Europe (SCALE CCR). The aim of this project is to establish a sound understanding and evidence-base on ICT-enabled innovations *for* learning which has significant scale and/or impact at system level and to identify policy recommendations for the further mainstreaming of ICT in E&T in Europe. In addition to an in-depth literature review, a number of case studies are being undertaken (1:1 learning initiatives, Boccini et al. (2013)), eTwinning, Hellerup School and Asia/Europe exchange), which will have produced final results before the CSCL 2013 Conference. Also, consultations with a wide spectrum of experts will have been undertaken with regard to the policy recommendations.

In order to investigate the sustainability and scalability of Creative Classrooms (CCR), we need to capture the complexity and richness of these learning ecosystems (Law et al., 2011). A multi-dimensional concept for CCR comprising eight encompassing and interconnected dimensions is proposed (Bocconi, et al., 2012) to capture the essential nature of these learning ecosystems: *Content and Curricula, Assessment, Learning Practices, Teaching Practices, Organization, Leadership and Values, Connectedness, and Infrastructure*. A set of 28 reference parameters have also been developed for policymakers, researchers and practitioners, which depict the *systemic approach* needed for the sustainable implementation and progressive up-scaling of Creative Classrooms across Europe. This holistic framework (see Figure 1) takes into account the key characteristics of innovative pedagogical practices at organizational, curricular, and assessment levels, and articulates with the systemic capability involving practices at classroom, school and whole community levels (i.e. at micro, meso and macro levels).

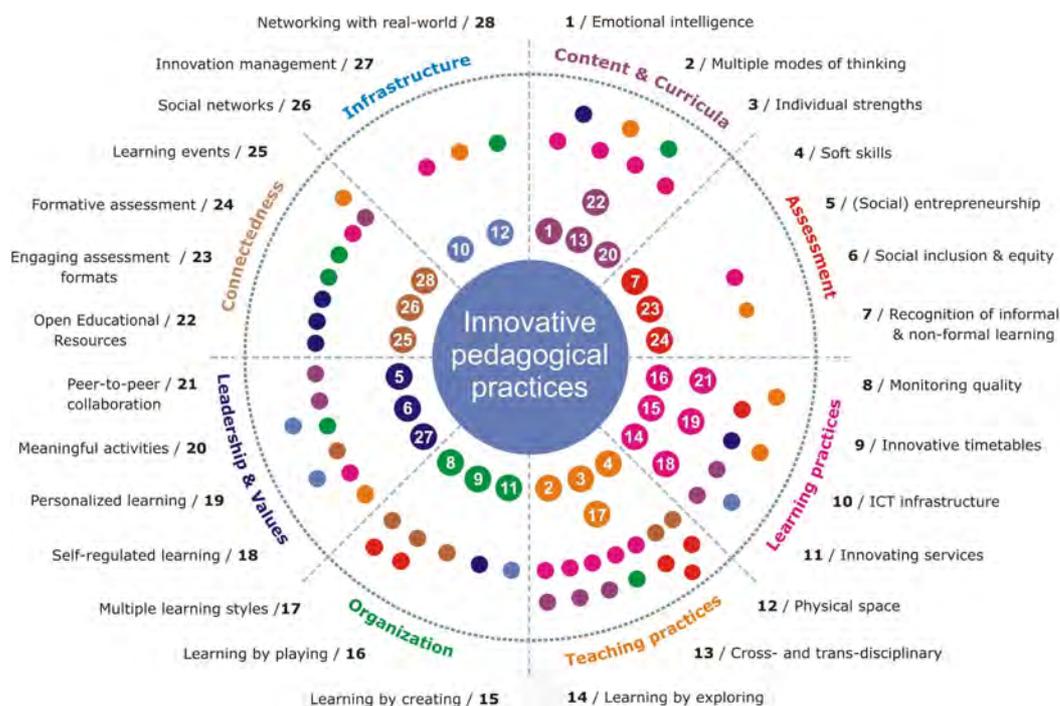


Figure 1. Bocconi, S., Kamylyis, P. & Punie, Y. (2012) Innovating Teaching and Learning Practices: Key Elements for Developing Creative Classrooms in Europe, *eLearning Papers*, No. 30, September, http://www.elearningeuropa.info/sites/default/files/asset/In-depth_30_1.pdf, p. 3.

Figure 1. Key dimensions and building blocks of Creative Classrooms

This paper will draw on the key findings from the SCALE-CCR project to address the three focal questions for this symposium, highlighting the multi-dimensional and holistic nature of ICT-enabled innovations in learning.

References

- Bocconi, S., Kamylyis, P. & Punie, Y. (2013). Framing ICT-enabled Innovation for Learning: the case of one-to-one learning initiatives in Europe. *European Journal of Education*. 48(1), 113-
- Bocconi, S., Kamylyis, P. G., & Punie, Y. (2012). *Innovating Learning: Key Elements for Developing Creative Classrooms in Europe*. Retrieved from <http://ftp.jrc.es/EURdoc/JRC72278.pdf>
- 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.
- Collins, A., & Halverson, R. (2009). *Rethinking education in the age of technology: The digital revolution and schooling in America*. New York: Teachers College Press.
- European Commission. (2013). Study of the impact of eTwinning on participating pupils, teachers and schools. Retrieved from http://ec.europa.eu/education/more-information/reports-and-studies_en.htm
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

- Fullan, M. (2008). *The Six Secrets of Change: What the Best Leaders Do to Help Their Organizations Survive*. San Francisco, CA: Jossey-Bass.
- Fullan, M. (2010). *Change leader: Learning to do what matters most*. San Francisco, CA: Jossey-Bass.
- Galvin, C. (2009). *eTwinning in the classroom - A showcase of good practice*. European Schoolnet. Retrieved from http://www.etwinning.net/en/pub/news/press_corner/statistics.cfm
- Hargreaves, A., & Fink, D. (2012). *Sustainable leadership*. San Francisco, CA: Jossey-Bass.
- Hatano, G., & Inagaki, K. (1991). Sharing cognition through collective comprehension activity. In B. Resnick, Levine, J.M. & Teasley, S.D.(Eds.), *Perspectives on socially shared cognition*, Washington DC: American Psychological Association, 331-348.
- Itakura, K. (1971). *Kagaku to kasetsu-Kasetsu-Jikken-Jugyo heno michi*. [in Japanese] (In English: *Science and hypothesis: The road to the Hypothesis-Experiment-Instruction*), Nobi Shobo.
- Law, N., Yuen, A., & Fox, B. (2011). *Educational Innovations Beyond Technology: nurturing leadership and establishing learning organizations*. New York: Springer.
- Looi, C.K., Chen, W.L. & Patton, C. (2010), Principles and enactment of rapid collaborative knowledge building in classrooms, *Educational Technology Magazine*, September-October, 2010.
- Looi, C.K., So, H-J., Toh, Y. & Chen W. (2011), The Singapore experience: Synergy of national policy, classroom practice and design research, *ijCSCL*, 6(1), 9-37.
- Looi, C.-K. (2011). *Sustained Innovation in Classroom Practices: The Role of Educational Researchers as Meso-Level Actors*. Paper presented at the Beijing Forum 2011, Beijing, China.
- Miyake, N., Pea, R., Barron, B., Schwartz, D., & Hall, R. (2007). Redefining Learning Goals of Very Long-Term Learning Across Many Different Fields of Activity. *Proceedings of the Computer Supported Collaborative Learning (CSCL) Conference 2007, Part I* [CD-ROM], USA, 26.
- Miyake, N. (in press). "Conceptual change through collaboration", In S. Vosniadou (ed.), *Handbook of research on conceptual change, 2nd Edition*, Taylor & Francis, London: U.K.
- mp3 (2009). Singapore's Third Masterplan in ICT in Education (mp3) (2009-2014). Available at: (<http://ictconnection.edumall.sg/cos/o.x?c=/ictconnection/pagetree&func=view&rid=665>)
- OECD (2009). *Creating effective teaching and learning environments: First results from TALIS*. Paris: Author.
- Penuel, W. R., Fishman, B. J., & Cheng, B. (2011). *Developing the area of design-based implementation research*. Menlo Park, CA: SRI International.
- Sawyer, R. K. (2006). *The Cambridge handbook of the learning sciences*. Cambridge, MA: Cambridge University Press.
- Shear, L., Means, B., Gallagher, L., House, A., & Langworthy, M. (2009). *ITL research design* (pp. 35). Menlo Park, CA: SRI International.
- Sinatra, G. M., & Pintrich, P. R. (2003). *Intentional conceptual change*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Spillane, J.P., Parise, L.M., & Sherer, J.Z. (2011). Organizational routines as coupling mechanisms for policy, school administration, and the technical core. *American Educational Research Journal*, 48(3), 586-619.
- Vosniadou, S. (2008). *Handbook of research on conceptual change*. Taylor & Francis, London: U.K.
- Vuorikari, R. (2013). *School Teams: Case studies on teacher collaboration through eTwinning*. European Schoolnet. Retrieved from http://www.etwinning.net/en/pub/news/press_corner/statistics.cfm
- Vuorikari, R., Berlanga, A. J., Cachia, R., Cao, Y., Fetter, S., et al. (2011). ICT-Based School Collaboration, Teachers' Networks and their Opportunities for Teachers' Professional Development - A Case Study on eTwinning. In Leung, H., Popescu, E., Cao, Y., Lau, R.W.H. & Nejdil, W. (Eds.), *Advances in Web-based Learning-ICWL 2011* (LNCS Vol. 7048, pp. 112–121). Berlin: Springer.
- Vuorikari, R., Garoia, V., Punie, Y., Cachia, R., Redecker, C., Cao, Y., ... Sloep, P. B. (2012). *Teacher Networks - Today's and tomorrow's challenges and opportunities for the teaching profession*. European Schoolnet.
- Wastiau, P., Crawley, C., & Gilleran, A. (2011). *Pupils in eTwinning- Case studies on pupil participation*. Central Support Service for eTwinning (CSS), European Schoolnet. Retrieved from http://resources.eun.org/etwinning/25/EN_eTwinning_165x230_Report.pdf
- Wen, Y., Looi, C. K., & Chen, W. (2012), Supporting Teachers in Designing CSCL Activities: A Case Study of Principle-based Pedagogical Patterns in Networked Second Language Classrooms. *Educational Technology and Society*, 15 (2), 138–153.

Designing to Improve Biology Understanding Through Complex Systems in High School Classrooms: No Simple Matter!

Susan Yoon, Joyce Wang, University of Pennsylvania, 3700 Walnut Street Philadelphia, PA 19104

Email: yoonsa@gse.upenn.edu, joycew@gse.upenn.edu

Eric Klopfer, Josh Sheldon, Daniel Wendel, Ilana Schoenfeld, Massachusetts Institute of Technology, 20 Ames Street, 02139

Email: klopfer@mit.edu, jsheldon@mit.edu, djwendel@mit.edu, ilanasch@mit.edu

Hal Scheintaub, Governor's Academy, 1 Elm Street Byfield, MA 01922

Email: HScheintaub@govsacademy.org

David Reider, Education Design INC, 7 Gibson Road, Newtonville, MA 02460

Email: david@educationdesign.biz

Abstract: This symposium seeks to illustrate and discuss the salient elements of project design and design decisions of a two-year implementation effort in high school science classrooms focused on improving knowledge and skills in biology content, complex systems, and computational thinking using StarLogo TNG simulations. We present design challenges that emerged in the areas of 1) Professional Development and Workshop Design; 2) Designing for Computational Thinking Through Computational Modeling; 3) Curriculum Design and Development; 4) Issues in the Design of Learning Progressions; and 5) Designing Assessments for Larger Scale. Through interactive discussion with the CSCL audience, we hope to share experiences that will enable similarly oriented design researchers to build successful programs in real-world educational systems.

Symposium Focus

A core activity of learning science researchers is to design interventions to improve learning and implement those interventions in real-world educational environments. As Kolodner (2004) writes, “If we want to understand how learning happens in complex situations, then we should study learning as it is occurring in those environments—with all the messiness of the real world and requiring methodologies that can nonetheless extract trends and descriptions” (p. 6). To do this, we turn to design research, which has gained momentum in the learning sciences for several reasons: its emphasis on accommodating a wide range of variables that may have contextual importance (Confrey, 2006); the dynamic process of assessment and evaluation of system states toward the goal of higher levels of educational improvement (Reimann, 2010); and the grounded-in-practice nature of the research, which increases the likelihood that interventions will be successful and sustained in the real world (Bielaczyc & Collins, 2007). Although design research is popular as a methodology for iterative improvement, few studies have discussed how design trajectories unfold. As Puntambekar and Sandoval (2009) write, “The learning sciences could benefit from clear examples of research trajectories that explicate how microcycles of analysis inform macrocycles and how iterated macrocycles build new knowledge” (p. 325). In other words, we need better descriptions and theories about how design teams measure, make decisions about and redesign system variables to produce the desired system-wide outcomes.

This symposium seeks to illustrate and discuss the salient elements of project design and design decisions as they emerged and changed over a two-year time frame. We report on implementation activities of a large-scale US National Science Foundation project called *BioGraph: Graphical Programming for Constructing Complex Systems Understanding in Biology*. This project is relevant to the CSCL community in that our goal is to develop curricular and instructional strategies to help teachers and students improve knowledge of biology content through computational modeling tools, emphasis on complex systems concepts, and a hypothesized curricular learning progression. Below, we provide a brief review of educational literature that highlights the need for this research followed by the curriculum and instruction conceptual framework that underpins the project design and activities.

Misconceptions in the Biological Sciences

Recent advances in the biological sciences that include the mapping of the human genome and the ability to manipulate atoms and molecules at the nanoscale have yielded unprecedented opportunities for humans to fashion our own evolutionary pathway. As Venter and Cohen (2004) write, “If the 20th century was the century of physics, the 21st century will be the century of biology” (p. 73). Yet, despite the enormous contemporary saliency, studies in science education have revealed robust misconceptions about concepts and processes in high school biology that directly impact students’ abilities to understand these recent advances. For example, student misunderstandings have been found across the scale of atoms (Taber & Garcia-Franco, 2010), to cells and genes (Garvin-Doxas & Klymkowsky, 2008), to organisms and ecology (Gotwals & Songer, 2010), as well as in the

relationships between these various scales (Sewell, 2002). Education researchers have speculated that these problems exist due to a lack of understanding of the complex systems realms in which these entities exist and interact (Chi, 2005). Thus, educational agencies in the US have urged science curriculum and instruction to emphasize systems content (AAAS, 2009; NRC, 2011).

Learning about Complex Systems and Learning Progressions

For more than a decade, knowledge of how students develop an understanding of complex systems has been an important theme in learning sciences research (Hmelo et al., 2000; Jacobson & Wilensky, 2006; Wilensky & Reisman, 2006; Yoon, 2008; 2011). Complex systems provide a framework through which one can explain and understand how patterns emerge across scales, that is, macro scale phenomena emerge from micro scale individual interactions. Complex systems scientists and educational researchers speculate that students have a hard time understanding the mechanisms that drive the emergence of large scale global phenomena from smaller scales of interacting agents (Chi, 2000). Explanations for how patterns emerge require integrating and matching explanations across scales. For example, while local environmental conditions can impose hard limits on where species can live and thereby impose large distribution patterns, the interactions of individuals within and between species contributes substantially to pattern development, influencing biodiversity and even evolution (Levin, 1999). While a coherent understanding of complex systems presently eludes most students (Jacobson, 2001), a biology sequence that is grounded in concrete examples as a starting place can tap into student's intuition about such systems and help build a deep understanding about complex systems as applied to biology, and even more generally. Recent learning progressions research offers a systematic approach in structuring such learning sequences (Alonso & Steedle, 2006; CPRE, 2009). For example, Mohan and colleagues (2009) identified levels of increasing sophistication in students' perception of carbon-transforming events (e.g., combustion, respiration) in complex socio-ecological systems. These ordered descriptions represent a research-informed framework for structuring the learning of core scientific ideas (NRC, 2007; Songer et al., 2009).

Computational Modeling

In addition to identifying a learning progression, educational computational modeling software and associated curricula including StarLogo, NetLogo, Biologica, and handheld Participatory Simulations (Colella et al., 2001; Gobert, 2005; Klopfer et al., 2005; Stieff & Wilensky, 2003; Wilensky & Reisman, 2006) have been created for school age students to learn about and visualize systems. Agent-based programs like StarLogo and NetLogo reveal how simple rules for interaction ascribed to individual agents with varying traits can produce emergent population scale patterns such as flocking behavior in birds, slime mold aggregation, or ant colony organization. Despite the promise of and need for these computational tools, widespread adoption in classrooms has not happened, and there are few studies that provide conclusive objective evidence of their benefit on learning. With respect to adoption by teachers, we know that the incorporation of technologically advanced curricular material into classrooms is met with many well documented challenges, including teacher time constraints, teachers' understanding of technology, teacher confidence levels in terms of computer programming, access to technology, and the lack of supporting curricular materials (Fishman et al., 2004; Yoon & Klopfer, 2006).

Project Activities and the Curriculum and Instruction Framework

The project entails building a curricular and instructional sequence in four high school biology units – Chemistry of Life, Population Ecology, Community Ecology, and Evolution that promote student learning of complex systems for implementation in high school biology classrooms. The project was funded in September of 2010. We report on project activities that ensued for two years until November 2012, which included: building biology simulations through the StarLogo TNG software that combines a graphical blocks-based programming with a 3-D game-like experience; development of classroom curricular materials, e.g., student lessons and teacher guides; construction of a summer teacher professional development workshop; and project implementation in classrooms. We worked with our first pilot cohort of four teachers between August 2011 and June 2012. Work with our second full cohort of 10 teachers began in August 2012 with the summer PD workshop. In order to address the aforementioned educational needs we constructed a curriculum and instruction (C and I) framework to inform the design and implementation of project activities. The C and I framework emerged from a legacy of design work (cf. Klopfer & Begel, 2005), classroom testing (Yoon & Klopfer, 2006), and educational learning research (Klopfer, 2008; Klopfer & Yoon, 2005) that spans more than a decade. As seen in Figure 1, the C and I framework has four main components: 1) Curricular relevance to ensure that project materials will be implemented and useful for students and in the classroom; 2) Cognitively-rich pedagogies to build on relevant understanding of best practices in learning theories; 3) Tools for learning and teaching to scaffold computational and curricular experiences; and 4) Learning progressions to structure sequences that will enable optimal understanding of project goals in computational thinking, biology content and complex systems concepts. In the remaining sections, we describe the design and development of five major project activities that will be discussed in more detail in the symposium: 1) Professional Development and Workshop Design; 2) Designing for Computational Thinking Through Computational Modeling; 3) Curriculum

Design and Development; 4) Issues in the Design of Learning Progressions; and 5) Designing Assessments for Larger Scale.

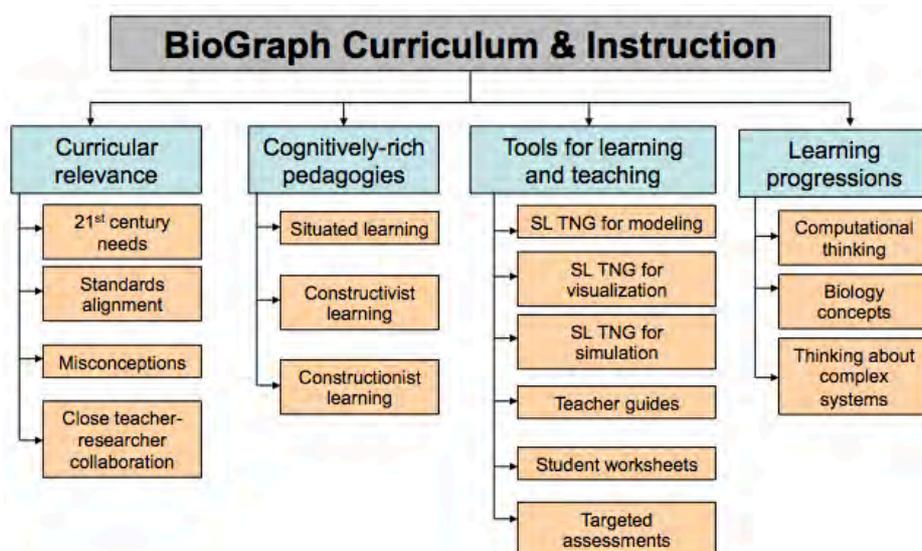


Figure 1. BioGraph Curriculum and Instruction Framework.

Since we are interested in describing the design parameters and trajectories, each description includes:

1. Initial ideas about the design of the activity as they are related to the curriculum and instruction framework and why and how they factored into the design.
2. How the activities as they unfolded did or did not address what we had envisioned to support the curriculum and instruction framework.
3. The rationale, decisions made, and steps taken in the redesign of the activity.

Professional Development Workshop Design

In this talk we will focus on the major design elements of the professional development activities that were constructed for our teacher participants. The four teachers in the pilot year came from three Cambridge and Boston area schools with an additional ten teachers in the second cohort coming from seven schools spanning eastern Massachusetts. Professional development activities for the pilot year teachers consisted of a one-week summer workshop followed by two half-day school year sessions.

Initially, the summer workshop was designed with the following goals: to introduce the topic of complex systems to teachers; to involve teachers in the co-design of curriculum activities (which were still in the process of being written and refined); and to introduce and begin helping teachers to become comfortable teaching computational modeling as part of their BioGraph instruction. The scope and sequence of activities was focused on gaining teachers' interest and engagement in adopting the C and I framework particularly in the way that computational modeling as a tool for learning could be used to encourage student inquiry and discovery in biology and related complex systems topics. In alignment with professional development research and the components needed for high-quality participation and enactment (e.g., Garet et al., 2001), we wanted to provide hands-on activities for teachers to learn about complex systems through participating in off-computer and on-computer simulations as their students would in inquiry-based activities. Additionally, we assumed from past experience that interacting with computer models and programming with teachers would be time-consuming. Thus, we made key design decisions to accommodate the finite amount of workshop hours which included decreasing the amount of time focused on learning about complex systems theory, fewer moments for reflection, and relatively little focus on classroom management and potential instructional issues in the implementation of project activities. Initial findings from the pilot teachers in workshop reviews, classroom observations, and follow-up interviews indicated a need to revise the professional development activities.

The pilot teachers indicated that they did not fully understand which core complex systems concepts were being illustrated in the StarLogo models and did not feel confident in teaching those to their students. Although they appreciated the active engagement and exploration of the models, they felt they were left to infer the meanings of terms. In classroom observations, teachers rarely made connections between the modeling activities and the higher-level complex systems concept being modeled e.g., self-organization, randomness, or decentralization. In the subsequent workshop in 2012, greater emphasis was placed on acquiring a more robust theoretical understanding of complexity. Teachers were provided several readings in advance of the workshop, were shown a number of videos and illustrations of different models and asked to reflect on the core similarities,

and importantly in order to make time for this theoretical investigation, two more professional development days were added prior to the summer workshop.

Despite the great emphasis in the workshop in computer programming, during our observations none of the teachers worked extensively with their students on learning how to build different models. Due to known challenges in time constraints, and teachers' learning curves, none of the teachers went beyond interacting with pre-existing models. Realizing the time-intensive nature of computer programming and the relatively little impact on classroom practice, we decided de-emphasize its focus in the subsequent summer PD workshop, instead opting for smaller chunks of programming interspersed through the week. However, this design decision showed interesting results with the second group of teachers. It allowed more time to focus on pedagogy and complex systems and having understood those variables better, provided some cognitive space to focus attention on programming. During the August 2012 workshop, several of the teachers engaged in a fruitful discussion on their own about how to change parameters within StarLogo to support or refute hypotheses about the biological systems being studied and initial observations in at least two classrooms showed teachers and students studying the programming blocks and making changes to the code to glean deeper biological understanding (as discussed in the following talk).

Another major design decision in the first PD workshop was to focus less on classroom management and pedagogical issues. Instead, we attempted to scaffold cognitively-rich interactions into the curriculum worksheets that students would work with while using the StarLogo models. Examples of scaffolds (to be discussed in more detail in the curriculum section of the proposal) included instructions to work with a partner and to fill in the worksheet together. However, in classroom observations, we found little collaboration among students. Instead, their participation in project activities looked much more like traditional plug and chug behavior as students (although interested and engaged) answered the questions one-by-one individually. In the subsequent workshop, we made a decision to work with teachers on how to include collaboration and argumentation (McNeill & Martin, 2011) in their pedagogy in conjunction with providing more directed opportunities in the worksheets for students to collaborate and argue their findings. Initial observations with the second cohort of teachers indicate greater amounts of peer-to-peer interaction and we hope to be able to report more detailed findings on this curriculum and instruction design variable in the symposium.

Designing for Computational Thinking Through Complex Systems Modeling

This talk focuses on the design decisions that influenced the construction of the simulations in the StarLogo TNG modeling tool and the computational thinking aspects implicit in the models that enable better understanding of biology content. We discuss the challenges of one of BioGraph's central goals, which is to bring computational thinking into science classrooms. Computational Thinking (CT) is increasingly understood to be a critical component of a 21st century education (NRC, 2010). However, CT is often relegated to technology classes, rarely bridging the gap to other subjects, despite its real-world prevalence in those fields.

CT is a large knowledge domain with many interpretations (NRC, 2010), but a few skills and concepts stand out as particularly relevant in a life sciences context. These are: the ability to interrogate and understand the underlying assumptions of models and simulations, and a basic paradigm for understanding and creating agent-based models. Two more general CT ideas underlie these life-science-specific skills. First is the idea that computers follow instructions literally—they do not add knowledge or interpretation beyond what is provided by the programmer. Second, combinations of these simple instructions form algorithms, which in an agent-based context lead to agent behaviors. Stated as the converse: modeling behavioral attributes requires reducing these behaviors to combinations of simple commands. In BioGraph, these CT ideas are critical for developing student understanding of complex systems, which is the cross-cutting theme throughout the curriculum, and for enabling the constructionist pedagogy of the C and I framework. For example, one cannot understand emergence without first understanding the basic rules of the individual agents. Similarly, the concepts of decentralized control and self-organization cannot simply be demonstrated by an animation, because students continue to imagine a large organizing force at work until they are able to construct a model without one.

However, there are many challenges inherent in bringing CT to the biology classroom (as discussed in the section on *Professional Development Workshop Design*). The primary challenge is the time required to integrate what is often taught as its own course into another course that is already saddled with more topics than can reasonably fit into a school year. Perhaps equally challenging is biology teachers' comfort level with teaching CT and programming. Our design efforts have focused on mitigating these challenges through intentional efforts in improving accessibility through the technology platform itself, how we embed programming activities in the curriculum, and low learning threshold activities in professional development.

StarLogo TNG, BioGraph's simulation platform, brings several technological solutions to these problems. The blocks-based visual programming interface eliminates the programming language syntax learning curve inherent to most programming languages, which can take weeks to cover in traditional computer science classes. Additionally, the 3D visualization engine lets students immediately see the effects of programming instructions, and changes to them, within an agent-based modeling paradigm. These features

VOLUME 1 AUTHOR INDEX

- Abdu, Rotem, 2
 Acosta, Alisa, 391
 Adams-Wiggins, Karlyn R., 18
 Adamson, David, 10, 105
 Ahn, June, 113, 549
 Alterman, Richard, 26
 Andrade, Luis, 34
 Arastoopour, Golnaz, 42
 Arvaja, Maarit, 49
 Asensio-Pérez, Juan I., 383
 Ashe, Colin, 10
 Ashe, David, 478, 494
 Barrat, Alain, 169
 Ben-Zvi, Dani, 240
 Biemans, Harm J.A., 375
 Blanton, Amos, 399
 Block, Florian, 153
 Bodemer, Daniel, 65
 Bonsignore, Elizabeth, 113, 549
 Borchers, Moritz, 57
 Bozelle, Christelle, 336
 Brassil, Chad, 176
 Brekelmans, Mieke, 518
 Buder, Jürgen, 65
 Bétrancourt, Mireille, 336
 Chanel, Guillaume, 336
 Chang, Chih-Hsuan, 73
 Chen, Bodong, 391
 Chen, Fei-Ching, 73
 Chen, Gaowei, 105
 Chen, Mei-Hwa, 81
 Chieu, Vu Minh, 89
 Ching, Dixie, 264
 Chinn, Clark, 564
 Chiru, Costin-Gabriel, 97
 Clarke, Sherice N., 105
 Clegg, Tamara, 113, 549
 Cober, Rebecca, 121
 Cress, Ulrike, 129, 557
 Cuendet, Sébastien, 137
 Danish, Joshua, 34, 192
 Dascalu, Mihai, 145
 Davis, Pryce, 153
 de la Torre, Luis, 161
 DeLiema, David, 192
 Dessus, Philippe, 145
 Diamond, Judy, 153
 Dillenbourg, Pierre, 137, 430, 564
 Dimitriadis, Yannis, 320, 383
 Dormido, Sebastian, 161
 Dowell, John, 502
 Druin, Allison, 549
 Dyke, Gregory, 105
 Eberle, Julia, 169
 Eck, Adam, 176
 Edbauer, Marie-Theresa, 320
 Edelmann, Jörg, 57
 Engelmann, Tanja, 184
 Enyedy, Noel, 192
 Erkens, Gijsbert, 518
 Evans, E. Margaret, 153
 Fields, Deborah, 200
 Filsecker, Michael, 208
 Fischer, Frank, 169, 256, 526, 564
 Georgiou, Yiannis, 272
 Giang, Michael, 200
 Greeno, James G., 105
 Gruson, Brigitte, 216
 Gubbels, Michael, 113
 Guha, Mona Leigh, 549
 Gunnarsson, Bjorn Levi, 26
 Hadjichambi, Demetra, 272
 Hadjichambis, Andreas, 272
 Halverson, Richard, 360
 Hashida, Tomoko, 224
 Hausknecht, Simone Nicole, 534
 Hayashi, Yugo, 232
 Heradio, Ruben, 161
 Herbst, Patricio, 89
 Hesse, Friedrich W., 57
 Hickey, Daniel Thomas, 208
 Hod, Yotam, 240
 Horn, Michael, 153
 Hourcade, Juan Pablo, 549
 Howley, Iris, 105
 Huang, Kun, 542
 Iida, Makoto, 224
 Jang, Hyeju, 10
 Janssen, Jeroen, 518
 Jara, Carlos, 161
 Jeong, Allan, 248
 Jermann, Patrick, 430
 Johri, Aditya, 470
 Jorrín-Abellán, Iván M., 383
 Ju, Wendy, 288
 Järvelä, Sanna, 280
 Kafai, Yasmin, 200
 Katz, Sandra, 105
 Kay, Judy, 320
 Kelly, Nick, 486
 Kennedy-Clark, Shannon, 486
 Klopfer, Eric, 580
 Kolikant, Yifat Ben-David, 414
 Kollar, Ingo, 256, 526, 564
 Kolodziej, Richard, 184
 Kozlov, Michail, 184
 Kwah, Helen, 264
 Kyza, Eleni, 272
 Langer, Sybille, 256
 Laru, Jari, 280
 Law, Nancy, 572
 Lee, Jiyeon, 81
 Lee, Woon Jee, 248
 Lewis, Sarah, 288
 Lewittes, Becky, 113
 Linn, Marcia, 572
 Loibl, Katharina, 296
 Looi, Chee-Kit, 462, 572
 Lui, Michelle, 304
 Lund, Kris, 169
 Lyons, Leilah, 264
 Magee, Rachel M., 312
 Martinez-Maldonado, Roberto, 320
 Mascaro, Christopher M., 312
 McCann, Colin, 121
 Medina, Richard, 328
 Mishenkina, Maria, 414
 Miyake, Naomi, 438, 572
 Mock, Philipp, 57

- Moher, Tom, 121
 Molinari, Gaëlle, 336
 Moreno, Yanín, 34
 Mulder, Martin, 375
 Naemura, Takeshi, 224
 Namdar, Bahadir, 344
 Nash, Pádraig, 352
 Nomura, Koki, 224
 Nüssli, Marc-Antoine, 430
 Öner, Diler, 454
 Ossorio, Pilar N., 422
 Owen, V. Elizabeth, 360
 Oztok, Murat, 368
 Parisio, Martin, 478, 494
 Pea, Roy, 406
 Pérez, Lenin, 34
 Phillips, Brenda, 153
 Popov, Vitaliy, 375
 Prieto, Luis P., 383
 Pun, Thierry, 336
 Punie, Yves, 572
 Rebedea, Traian, 97
 Reichersdorfer, Elisabeth, 526
 Reider, David, 580
 Reiss, Kristina, 526
 Resendes, Monica, 391
 Resnick, Lauren B., 105
 Rhodes, Emily, 113
 Rogat, Toni Kempler, 18
 Roque, Ricarose, 399
 Rosé, Carolyn Penstein, 10, 105
 Rosé, Carolyn P., 10
 Rudat, Anja, 65
 Rummel, Nikol, 296
 Rusk, Natalie, 399
 Sailer, Michael, 169
 Scardamalia, Marlene, 391
 Scheintaub, Hal, 580
 Schneider, Bertrand, 406
 Schoenfeld, Ilana, 580
 Schrementi, Laurel, 153
 Schwarz, Baruch B., 414
 Schwind, Christina, 65
 Sensevy, Gérard, 216
 Shaffer, David Williamson, 42, 352
 Shapiro, R. Benjamin, 360, 422
 Sharma, Kshitij, 430
 Sheldon, Josh, 580
 Shen, Chia, 153
 Shen, Ji, 344
 Shirouzu, Hajime, 438
 Siqin, Tuya, 446
 Slotta, James D., 121, 304, 564
 Soh, Leen-Kiat, 176
 Stahl, Gerry, 312, 454
 Stainton, Catherine, 105
 Stegmann, Karsten, 169, 564
 Sun, Daner, 462
 Suthers, Daniel, 328
 Tchounikine, Pierre, 564
 Teo, Evelyn, 462
 Teo, Hon Jie, 470
 Thompson, Kate, 478, 486, 494
 Trausan-Matu, Stefan, 97, 145
 Tscholl, Michael, 502
 Ufer, Stefan, 526
 van Aalst, Jan, 446
 van de Sande, Carla, 510
 van Leeuwen, Anouschka, 518
 Villagrà-Sobrino, Sara, 383
 Vogel, Freydis, 526
 Vuorikari, Riina, 572
 Wah Chu, Samuel Kai, 446
 Wang, Joyce, 580
 Wardak, Dewa, 478
 Wecker, Christof, 256, 564
 Wendel, Daniel, 580
 Wheeler, Penny, 486
 Wise, Alyssa Friend, 534
 Xie, Kui, 542
 Yacef, Kalina, 320
 Yang, Cheng-Yu, 73
 Yaron, David, 10
 Yeoman, Pippa, 478, 494
 Yip, Jason C., 113, 549
 Yoon, Susan, 580
 Zahn, Carmen, 57
 Zhang, Jianwei, 81
 Zhao, Yuting, 534