Aalborg Universitet

Adapting Interaction Analysis to CSCL: a systematic review

Steier , Rolf ; Davidsen, Jacob Gorm

Published in:
14th International Conference on Computer-Supported Collaborative Learning (CSCL)

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
Steier , R., & Davidsen, J. G. (2021). Adapting Interaction Analysis to CSCL: a systematic review. In C. E. Hmelo-Silver, B. De Wever, & J. Oshima (Eds.), *14th International Conference on Computer-Supported Collaborative Learning (CSCL)* (pp. 157-160). International Society of the Learning Sciences (ISLS). https://drive.google.com/file/d/1cuUAcl3Nwxtqgcz7ufdvyOum8oWp8-wM/view

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: July 04, 2025

ISLS Annual Meeting 2021

Reflecting the Past and Embracing the Future

Bochum, Germany, June 8-11

Workshops: June 1-7

Ruhr University Bochum (Online Event)

14th International Conference on Computer-Supported Collaborative Learning (CSCL)

- Proceedings -

Edited by: Cindy E. Hmelo-Silver, Bram De Wever, & Jun Oshima

International Society of
the Learning Sciences

International Society of the Learning Sciences

ISLS Annual Meeting 2021
Reflecting the Past and Embracing the Future
Bochum, Germany, June 8-11
Workshops: June 1-7
Ruhr University Bochum (Online Event)

14th International Conference on Computer-Supported Collaborative Learning (CSCL) 2021

- CSCL Proceedings -

1st Annual Meeting
of the
International Society of the Learning Sciences (ISLS)

: Editors
Cindy E. Hmelo-Silver, Bram De Wever, & Jun Oshima

ISLS Annual Meeting 2021
Reflecting the Past and Embracing the Future
Bochum, Germany, June 8-11
Workshops: June 1-7
Ruhr University Bochum (Online Event)

14th International Conference on Computer-Supported Collaborative Learning (CSCL) 2021

© 2021 International Society of the Learning Sciences, Inc. [ISLS]. Rights reserved.

www.isls.org

ISBN: 978-1-7373306-2-2 (CSCL Proceedings, PDF Version)
ISSN: 1573-4552

Cite as: Hmelo-Silver, C. E., De Wever, B., & Oshima, J. (Eds.). (2021). *Proceedings of the 14th International Conference on Computer-Supported Collaborative Learning - CSCL 2021.* Bochum, Germany: International Society of the Learning Sciences.

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

Conference logo design by Cornelia Robrahn C74 Gestaltung & Design: https://c74.org/

International Society of the Learning Sciences

Conference Organizers and Committees

Conference Organizers
Astrid Wichmann*, Ruhr University Bochum, Germany*
Ulrich Hoppe, *University of Duisburg-Essen, Germany*
Nikol Rummel, *Ruhr University Bochum, Germany*

Program Committee Co-Chairs:
Cindy Hmelo-Silver, *Indiana University, USA*
Jun Oshima, *Shizuoka University, Japan*
Bram de Wever, *Ghent University, Belgium*

Technology Support Committee:
Oskar Lindwall, *University of Gothenburg, Sweden*
Meagan O'Hara, *University of Toronto, Canada*
Preeti Raman, *University of Toronto, Canada*
Jim Slotta, *University of Toronto, Canada*
Joel Wiebe, *University of Toronto, Canada*

Social Media Communications Co-Chairs:
Charleen Brand, *Ruhr University Bochum, Germany*
Jenny Olsen, *University of San Diego, USA*
Sebastian Strauß, *Ruhr University Bochum, Germany*

Online Conference Committee:
Irene-Angelica Chounta, *University of Duisburg-Essen, Germany*
Anouschka van Leeuwen, *Utrecht University, Netherlands*
Yotam Hod, *University of Haifa, Israel*
Joey Huang, *University of California, Irvine, USA*
Jun Oshima, *Shizuoka University, Japan*
Joseph Polman, *University of Colorado Boulder, USA*

Advisory Board:
Susan Goldman, *University of Illinois at Chicago, USA*
Melissa Gresalfi, *Vanderbilt University, USA*
Ilana Horn, *Vanderbilt University, USA*
Eleni Kyza, *Cyprus University of Technology, Cyprus*
Oskar Lindwall, *University of Gothenburg, Sweden*
Kris Lund, *University of Lyon, France*
Gerald Niccolai, *University of Lyon, France*
Jim Slotta, *University of Toronto, Canada*

Copy Editors
Liat Rahmian, *University of Haifa, Israel*
Shiri Kashi, *University of Haifa, Israel*

Senior Reviewers

Brigid Barron
Daniel Bodemer
Douglas Clark
Hanni Muukkonen
Heisawn Jeong
Irene Chounta
Jennifer Olsen
Jeremy Roschelle
Laura Bofferding
Marcelo Worsley
Mireille Betrancourt
Murat Cakir
Omid Noroozi
Päivi Häkkinen
Philippe Dessus
Pierre Tchounikine
Richard Alterman
Roberto Martinez-Maldonado
Sandra Y. Okita
Sanna Järvelä
Sébastien George
Stefan Trausan-Matu
Sten Ludvigsen
Toni Rogat
Ulrike Cress
Wouter van Joolingen
Yannis Dimitriadis

Reviewers

Alessia Ruf
Ali Raza
Alyssa Wise
Amanda Dickes
Anders Mørch
Andrea Gomoll
Andy Nguyen
Anna Shvarts
Annelies Raes
Anouschka van Leeuwen
Anthony Petrosino
Anuli Ndubuisi
Ayano Ohsaki
Blair Winograd
Bohdana Allman
Bram De Wever
Britte Cheng
Chao Zhang
Chih-Pu Dai
Christophe Reffay
Claudia Ramly
Crina Damsa
Cynthia D'Angelo
Daniel Spikol
Danielle Espino
Deanna Kuhn
Deborah Silvis
Dhvani Toprani
Dimitra Tsovaltzi
Edwin Chng
Ehsan Toofaninejad
Elena Boldyreva
Elise Lavoué
Elizabeth Charles
Emma Reimers
Etan Cohen
Feng Lin
François-Xavier Bernard
Gerry Stahl
Ha Nguyen
Hannah Smith
Hans Chr Arnseth
Heather Killen
Heather Leary
Hideki Kondo
Huiruo Chen
Iris Tabak
Ishari Amarasinghe
James Planey
Jamie Costley
Janet Shufor Bih Epse Fofang
Jennifer Kahn
Jesmine S. H. Tan
Ji-Eun Lee
Joel Wiebe
Joey Huang
Johanna Pöysä-Tarhonen
Jorge Garcia
Joshua Rosenberg
Kate Miller
Kshitij Sharma
Lauren Vogelstein
Lexie Zhao
Liat Rahmian
Louise Mifsud
Luis Perez Cortes
Man Su
Mario Niederhauser
Martin Greisel
Max Sherard
Michael Baker
Michelle Lui
Mik Fanguy
Mohsen Roshanian Ramin
Monaliza Chian
Nicole Hutchins
Nicole Wessman-Enzinger
Nikki Lobczowski
Pirkko Siklander
Pryce Davis
Ravi Sinha
Renato Carvalho
Robert Huang
Rolf Steier
Saadeddine Shehab
Saeed Latifi
Sai Raj Reddy
Sarah Bryans-Bongey
Sarah McGrew
Sari Widman
Shagun Singha
Shiri Kashi
Simona Pesaresi
Sreecharan Sankaranarayanan
Stephan Mende
Taehyun Kim
Therese Laferriere
Thomas Hillman
Thomas Lerche
Thomas Richman
Tugce Aldemir
Virginia J. Flood
Xiao Ge
Xiao Hu
Yana Landrieu
Yihong Cheng
Yingying Zhao
Yuchan (Blanche) Gao
Yugo Hayashi
Zarifa Zakaria

International Society of
the Learning Sciences

Preface

With its 14th edition, the International Conference on Computer-Supported Collaborative Learning (CSCL) has now moved to an annual event. Together with the International Conference of the Learning Sciences (ICLS), it is part of the 2021 Annual Meeting of the International Society of the Learning Sciences.

As the first of its kind, the ISLS Annual Meeting 2021 is a milestone in the evolution of the Learning Sciences and CSCL community. This meeting was envisaged to take place in Bochum, a location with a history that symbolizes the theme of the conference: "Reflecting the past and embracing the future." Bochum lies in the heart of Europe in a region that has been historically shaped by the heavy industries but also by the solidarity and conviviality of workers with a variety of cultural backgrounds. After the decay of the old industries, this spirit of solidarity is an important asset for embracing the present and future challenges. We hope this will inspire our growing international community, even though we do not have a chance to meet in place this time.

As part of the ISLS Annual Meeting 2021, CSCL 2021 invites academics, researchers, professionals, and educators to share and embrace their diverse views. This includes empirical, theoretical, conceptual, design-based work, and system development. The CSCL Proceedings feature long papers, short papers, posters and symposia, all subject to a rigorous double-blind peer review.

We had 104 submissions from 20 countries over Europe, Asia-Pacific and America, which covered a broad range of CSCL research and design. In total, 33% (17 out of 51) of long paper submissions, and 42% (14 out of 33) of short paper submissions were accepted in the category where they were submitted. In addition, a number of submissions were accepted in another category (short papers or posters). As a result, the CSCL Proceedings features 17 long papers, 24 short papers (work-in-progress), 32 posters, and one symposium. This year we have observed that the diversity of topics in CSCL research has been continuously expanding over the years. The program includes research on innovative technologies, learning analytics, instructional designs, equity and identity, and more. Given the challenges of this year of the COVID-19 pandemic, Computer-Supported Collaborative Learning was often a critical part of the educational infrastructure, which was reflected in the submissions to the CSCL program. The successful program would not be possible without authors, reviewers, and the local organizing team.

We would like to take this opportunity to offer a special thank you to the 110 reviewers and 27 senior reviewers who carried out over 300 reviews and meta-reviews and helped us in making the final decisions on each submission, as well as the numerous people around the process who have spent countless hours ensuring that the program is of high quality.

It is a great honor to edit this year's proceedings for the CSCL community in ISLS. We hope all of you enjoy your participation in the conference and social activities designed online.

Cindy Hmelo-Silver, Bram De Wever, and Jun Oshima
CSCL 2021 Conference Program Co-chairs

 © ISLS

International Society of
the Learning Sciences

Acknowledgments

ISLS Annual Meeting 2021 conference organizers would like to give special thanks to the following committees and people for their support:

- International Society of the Learning Sciences (ISLS)
- ISLS Board of Directors and Executive Committee
- Ruhr University Bochum, Germany
- Deutsche Forschungsgemeinschaft (DFG)
- Wallace Foundation
- National Science Foundation (NSF)
- ExOrdo Support Team
- Whova Support Team
- Gather/Meetingland Support Team
- Conference Organizing Committee including
    - LS and CSCL Program Committees
    - Social Media Communications Co-Chairs
    - Doctoral Consortium Co-Chairs
    - Early Career Workshop Co-Chairs
    - Mid-Career Workshop Co-Chairs
    - Interactive Tools and Demos Co-Chairs
    - Pre-Conference Workshops and Tutorials Co-Chairs
    - Technology Support Committee
    - Online Conference Committee
    - ISLS 2021 Advisory Board
- Reviewers and Sr. Reviewers
- Best Paper Award Committees
- Conference Proceeding Copy Editors

 © ISLS

International Society of the Learning Sciences

Table of Contents

CSCL Proceedings

Long Papers

International Society of the Learning Sciences

Short Papers

Symposia

Posters

Indexes

Long Papers

Impact of Learners’ Video Interactions on Learning Success and Cognitive Load

Alessia Ruf, University of Basel, alessia.ruf@fhnw.ch
David Leisner, Carmen Zahn
leisn_r@hotmail.com, carmen.zahn@fhnw.ch
University of Applied Sciences and Arts Northwestern Switzerland (FHNW)
Klaus Opwis, University of Basel, klaus.opwis@unibas.ch

Abstract: Enhanced video-based learning environments provide new tools (e.g., hyperlinks) – along with the well-known basic video control tools (e.g., play, pause, rewind) – that afford learners‘ enhanced interaction with videos. With these tools, learners can actively transform existing videos into their own hypervideo structures by adding hyperlinks and own materials. Unlike research on basic control tools that has revealed positive impacts on learning, research on enhanced tools is still rare and conflicting. It is thus open, whether the tools support generative interested learning or put too much extrinsic cognitive load onto learners. In the present study, we investigated the effects of video annotation and hyperlinking tools on learning success and cognitive load by analyzing tool-related interaction behavior data of 141 university students. Results indicated that the frequent use of enhanced video tools positively predicted learning success and a decrease in cognitive load. Implications of these results are discussed.

Keywords: interactive learning environments, video-based learning, interaction behavior.

Introduction and related work

Video is a popular, effective, and timely medium for supporting learning – and has been for a long time (for a review see Poquet, Lim, Mirriahi, & Dawson, 2018). Streaming media platforms such as YouTube contribute to a continuous increase of students’ access to digital video-based material (Poquet et al., 2018). Previous approaches emphasized that such dynamic audiovisual media support learning (both factual and procedural) when designed according to concrete guidelines (Mayer, 2005). Besides, the possibility to *interact* with such media plays a further decisive role in fostering learning processes: for example, Schwan and Riempp (2004) investigated the effects of basic control tools – such as play, pause and rewind – on learning nautical knots of varying difficulties and could show that learners successfully used these tools for strategic interactions. Further research suggested that the active use of basic control tools correlated significantly with knowledge acquisition (Zahn et al., 2004) as they allow learners to learn at their own pace which – in turn – minimizes the risk of cognitive overload (Cattaneo et al., 2015) – or as Schwan and Riempp (2004) put it: to adapt information flow to internal cognitive needs.

Today, *enhanced video-based environments* provide tools that additionally allow to annotate (e.g., with hyperlinks or annotations for self-written summaries), comment, discuss, and edit videos alone or in groups (e.g., Leisner, Zahn, Ruf, & Cattaneo, 2020; Sauli, Cattaneo, & van der Meij, 2018; Yousef, Chatti, Danoyan, Thüs, & Schroeder, 2015; Zahn, 2017). With such *enhanced interaction tools*, learners are able to actively transform existing video representations into their own enriched information structures (Schwartz & Hartman, 2007; Yousef et al., 2015) and, therefore, actively generate meaning (Wittrock, 1992) by designing their own learning content (e.g., Kafai & Resnick, 1996; Papert, 1994). Such an active participation of learners in constructing information is crucial for conceptual understanding and fosters deep processing and re-organization of concepts (Kafai & Resnick, 1996; Papert, 1994; Wittrock, 1992). Delen, Liew, and Willson (2014) provided evidence that using enhanced tools for generative note-taking was superior to working with basic control tools regarding learning success. Besides, Zahn, Pea, Hesse, and Rosen (2010) and Zahn, Krauskopf, Hesse, and Pea (2012) found that designing a hypervideo structure is suitable for successfully learning complex history topics. However, research on enhanced tools is conflicting (see Sauli, Cattaneo, & van der Meij, 2018): for instance, Merkt et al. (2011), who investigated the impact of a table of contents in videos, found no effects on learning success. Two possible explanations for these conflicting results are discussed as follows: first, learners may be overwhelmed by the complexity of enhanced tools (Krauskopf et al., 2014; Zahn et al., 2012), which may be manifested in an increase of extraneous cognitive load (Kirschner et al., 2018; Paas, 1992). Second, some previously investigated enhanced tools were more intended to be *optional supporters* for facilitating video interaction (e.g., table of contents, see Merkt et al., 2011), than tools that are *necessary to complete the learning task* (e.g., note-taking, see Delen et al., 2014). Learners seem to have a lack of strategies underlying the use of *optional* tools and therefore hardly use them (Merkt et al., 2011), which probably results in an increased extraneous cognitive load as learners need cognitive resources to process them as part of the learning environment but not necessarily need them to complete the task (Kirschner et al., 2018; Paas, 1992; Zahn et al., 2012). These issues could be solved when learners are provided with clear instructions about how to use enhanced tools efficiently and how to include them as part of a concrete learning task (Shin & Jung, 2020; Zahn et al., 2012). According to Sweller's (1999) Cognitive Load Theory, such *task-relevant* enhanced tools can reduce intrinsic cognitive load by helping learners to disaggregate the difficulty of the learning material by actively creating their own hypervideo structures (Kafai & Resnick, 1996; Papert, 1994; Wittrock, 1992; Yousef et al., 2015). This is also in line with constructivist approaches suggesting that learning is not a consequence of offering tools, but, after all, depends on internal processes associated with tool use – that is: concrete learning activities in a constructive learning process (Clark, 1994; Kozma, 1994).

It becomes clear from the research described above that investigating learners' interaction with videos is promising to understand how video tools can successfully be used for learning. This potential has been addressed by the research field on *learning analytics*, which suggests to measure learning behavior using logged interaction data (e.g., Mirriahi & Vigentini, 2017). Accordingly, the use of basic control tools or enhanced tools can be measured using log files that provide logged users' (inter-)actions – such as pressing buttons – in form of tabular representations. Thereby, it is important to note that the use of basic control tools (e.g., pressing the pause button to pause the video) is often reflected in a single logged action (e.g., logged action: pause), whereas the use of enhanced tools is usually reflected in multiple logged actions related to it: for example, a hyperlink can be added to the video timeline of the video, or moved within the timeline, or deleted from the timeline. Previous approaches further suggested to distinguish between different *levels of interactivity* resulting from the use of basic control or enhanced tools (Delen et al., 2014; Merkt et al., 2011). Accordingly, the use of basic control tools (e.g., play, pause, rewind) can be subsumed under the term *micro-level interactivity* and the use of enhanced tools (e.g., table of contents, hyperlinks, annotations) under the term *macro-level interactivity* (see Delen et al., 2014; Merkt et al., 2011). In line with these approaches, in the present study we summarized single learners' actions resulting from the use of basic control tools under the term "micro-actions". In addition, and as stated above, we further classified enhanced tools as either *optional supporters* for facilitating learning with videos (e.g., table of contents, Merkt et al., 2011) or as important and *necessary parts of a concrete learning task* (e.g., note-taking, Delen et al., 2014). Consequently, we summarized learners' actions resulting from the use of *task-relevant* enhanced tools under the term "task-actions".

The present study aims to add new original findings to the corpus of existing research on the effects of enhanced tools in video-based environments on learning by pursuing the two following research objectives: first, in consideration of the previously described conflicting results concerning enhanced tools and learning success (Delen et al., 2014; Merkt et al., 2011; Sauli et al., 2018; Zahn et al., 2012), we investigate the effects of learners' performed micro- and task-actions (i.e., actions resulting from *task-relevant* enhanced tools: annotations and hyperlinks) on learning success using frequencies of learners' actions (cf. Hung & Zhang, 2008) and, second, in order to address possible overwhelming situations provoked by enhanced tools, we additionally consider cognitive load by investigating both mental load and mental effort (Kirschner et al., 2018; Paas, 1992; Zahn et al., 2012). The study is guided by the following hypotheses:

(1) *Learning success*: frequently performed (H1a) micro-actions and (H1b) task-actions are positively related to learning success (i.e., objective learning success and self-assessed knowledge gain).

(2) *Cognitive load*: frequently performed (H2a) micro-actions and (H2b) task-actions reduce cognitive load (i.e., mental load and mental effort).

In the next section, we give a description of the study context, the data set, and the measures used.

Method

Study context and description of the data set

To answer our hypotheses, we used a subsample (N = 141) from a data set consisting of 209 Swiss University students (75% female, $M = 24.30$ years, $SD = 6.70$) who learned a complex learning topic from natural sciences (i.e., synaptic plasticity) with an enhanced video-based environment (i.e., *FrameTrail*, see Figure 1). The ethical standards of the controlled laboratory experiment were set through the institutional ethical committee of our institution. Participants received course credits for participation and had no or marginal experience with the learning topic prior to participation. They were randomly assigned to the experimental conditions of a 3 x 2 study plan where the first factor concerned the video-related learning task (adding hyperlinks vs. adding annotations for self-written summaries vs. considerate-watching) and the second factor related to the learning setting (individual vs. dyadic collaborative learning). After instructions concerning the task and the usage of the tools, participants learned the topic individually or in groups of two by adding either (1) hyperlinks containing further thematic information from prepared written texts (see *Further information texts* in Figure 1) or (2) self-written annotations based on these texts directly into the video, or (3) they received further information texts but were not able to add them into the video (i.e., considerate-watching condition). Participants learned at their own pace, so that they had the chance (1) to fully understand the learning topic, (2) to complete the task, and (3) to compensate for possible effects of extraneous cognitive load triggered by the (initially unfamiliar) learning environment and tools.

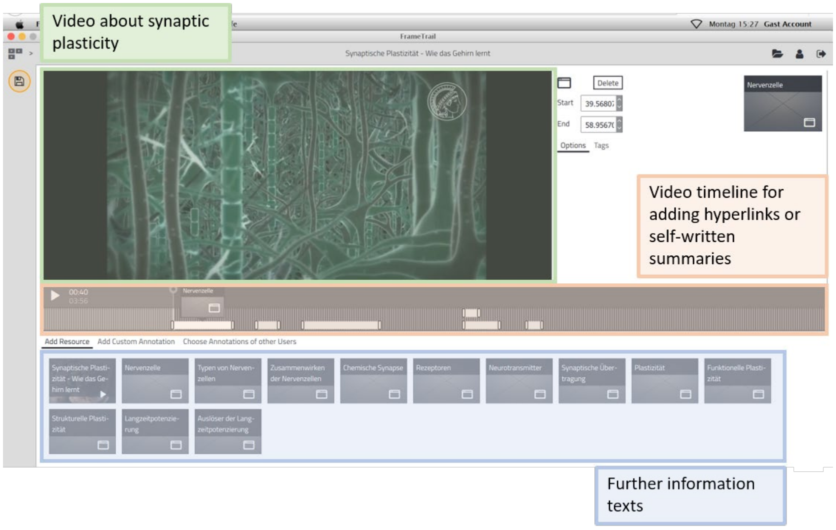


Figure 1. Enhanced video-based learning environment *FrameTrail* (see https://frametrail.org).

For the present study, only data from participants learning in the two "enhanced" learning task conditions (i.e., hyperlink and annotation) were considered, because only they had the possibility to perform task-actions with necessary tools according to their learning task. Thus, from the 209 datasets, 74 were excluded and the remaining data sample consisted of 141 participants (75% female, 79% psychology students, M = 24.27 years, SD = 6.70). Thereof, 53 participants learned in an individual learning setting and 88 learned collaboratively in 44 dyads. Furthermore, 71 used annotations and 70 used hyperlinks to complete the learning task. It is important to note that although only one set of interaction data was collected for dyads, because groups worked together on a shared desktop computer, dyad interaction data was used for the present study as *individual data* for the purpose of comparison between groups. With this, we refer to literature on *joint attention*, which revealed that interactions of collaborative dyads are closely coupled (Barron, 2003; Schneider & Pea, 2013). Moreover, two analysis of variance (ANOVAs) with micro- or task- actions as dependent variables and learning setting (individual vs. collaborative) as between-subject factor did not reach significant levels (p > .05). Thus, individuals and collaborative learners were comparable on these variables. This approach further allowed us to examine effects of video interaction on individual learning success and subjective perceptions of knowledge gain and cognitive load.

Measures

Learners' video interactions were collected with log files provided by the enhanced video-based environment used in the study (see https://frametrail.org). Table 1 lists the collected actions. As mentioned above, we summarized actions resulting from basic control tools into *micro-actions* and actions resulting from the use of enhanced tools (annotations and hyperlinks) into *task-actions*. The circumstance that participants learned at their own pace was reflected in a remarkable spread of variance for both *absolute learning time* (M = 42.17 min, SD = 15.22) and *absolute frequencies of performed actions* over all participants (micro-actions: M = 88.96, SD = 47.10; task-actions: M = 67.53, SD = 41.21). Therefore, we considered *relative values of actions* (division of absolute interaction frequencies of micro- and task-actions by learning time in minutes) to address individual learning pace (micro-actions per minute: M = 2.31, SD = 1.30; task-actions per minute: M = 1.58, SD = 0.65). Although we conducted analyses for both absolute and relative values of performed micro- and task-actions, for the purpose of this contribution as well as its substantive relevance (through the consideration of learning time), we only focused on relative values here.

Table 1: Collected micro-actions and task-actions

Micro-actions	Task-actions
Play	Adding hyperlink or annotation into video
Pause	Change annotation text
Jump backwards	Change displayed time of hyperlink or annotation on video timeline
Jump forward	Delete hyperlink or annotation from timeline

To measure *learning success*, participants were, first, asked to answer an objective knowledge test with 20 questions (post-experimental) developed with an expert of biopsychology at our institution. The questionnaire consisted of 15 multiple choice (four answer options with one correct answer) and five open short-answer questions. Out of these, eight questions addressed understanding of concepts (e.g., “What are vesicles?” referred to the understanding of the concept “vesicle”), eight related to understanding of concept interrelations (e.g., “What role do calcium ions play in synaptic transmission?” referred to the understanding of the concepts “calcium ion” and “synaptic transmission” and their interrelation), and four measured transfer knowledge (transferring learning information to other situations or circumstances, see Rebetez, Bétrancourt, Sangin, & Dillenbourg, 2010). The distinction between concept and concept interrelation was crucial with regard to our learning task conditions as the use of annotations is assumed to foster relations between prior knowledge and new information, and thus understanding of concepts (Zahn et al., 2010, 2012), and the use of hyperlinks is assumed to foster relations among concepts (Stahl, Finke, & Zahn, 2006). Participants received one point for a correct and zero points for an incorrect answer. Cronbach’s α for the final test was .76 (note: this analysis was conducted with the full data sample, N = 209). Second, we measured self-assessed knowledge gain (post-experimental) with a one-item scale (i.e., “How much do you think your knowledge in synaptic plasticity has improved?”) from 1 (not at all) to 5 (very much).

To measure *cognitive load*, we focused on Paas (1992) and De Jong (2010) and analyzed both concepts *mental load* (imposed by instructional parameters such as task structure) and *mental effort* (capacity assigned to instructional demands) separately to consider the large variety of definitions of the construct: first, mental load was measured according to Kalyuga, Chandler, and Sweller (1999). Participants rated the item “Please estimate how easy or how difficult you found the learning material.” from 1 (very easy) to 7 (very difficult). Second, in order to measure mental effort, we took a closer look into the subscale *effort/importance* of the short version of the Intrinsic Motivation Inventory (KIM, Wilde et al., 2009, see Table 2). Note that the items of this scale were originally validated by Wilde et al. (2009) in German language and were translated from German to English for the purpose of this contribution. The participants rated the subscale from 1 (totally disagree) to 5 (totally agree) on three items. A reliability analysis (conducted with the subsample, N = 141 of the present study) for this subscale revealed α = .457. However, when item 1 (i.e., “Editing the video in the learning environment was a considerable effort for me.”, see Table 2) was excluded from the scale, Cronbach’s α changed to .755. We consequently concluded that item 1 measured the actual “effort” while items 2 and 3 were more related to perceptions of “importance”. Item 1 was proximately used to measure mental effort and was extracted from the original subscale. Both scales were measured post-experimental.

Table 2: Subscale effort/importance of the short version of the Intrinsic Motivation Inventory (KIM)

Subscale	Cronbach’s α	Nr.	Item
Effort / Importance	.457	1	Editing the video in the learning environment was a considerable effort for me.
		2	I tried to do my best.
		3	It was my personal concern to perform well at editing the video in the learning environment

To answer the hypotheses described above, several multiple and multivariate multiple regression analyses were conducted with micro- and task-actions as predictors and measures regarding learning success and cognitive load as dependent variables.

Results

For data preparation, we first investigated the correlation of micro- and task-actions using Pearson correlations and found no significant result (p > .05). When conducting regression analyses, predictor variables are ideally independent to minimize the risk of suppressor effects (Bortz, 2005). Hence, we concluded that both predictor variables (micro- and task-actions) could be examined independently. Second, we calculated a Pearson correlation with the dependent variables mental effort and mental load and found no significant results ($p > .05$). Therefore, we examined these variables independently in two multiple regression analyses. For the statistical tests, an α-level of .05 was used.

Effects of interaction frequencies on learning success (H1)

To answer our hypotheses on learning success (H1), a multivariate multiple regression analysis with the three scores of objective learning success (understanding of concept, understanding of concept interrelations and transfer tasks) as dependent variables and micro- and task-actions as predictors was conducted (see Table 3). A significant regression equation was found for understanding of concepts for task-actions ($F(1,132) = 5.31$, $p = .023$). As expected, (Hb), this result indicates that the more task-actions were performed the higher were learning success outcomes in understanding of concepts. However, no other result reached a significant level ($p > .05$). Hence, we could not confirm a positive relation between micro-actions and objective learning success (H1a).

Table 3: Results on the impact of micro- and task-actions on objective learning success

	Concept				Concept interrelations				Transfer			
Predictors	β	SE β	R^2	ΔR^2	β	SE β	R^2	ΔR^2	β	SE β	R^2	ΔR^2
Micro-actions	.023	.107	.039	.024	-.088	.136	.003	-.012	.022	.062	.004	-.011
Task-actions	.493*	.214	.039	.024	-.044	.273	.003	-.012	.081	.125	.004	-.011

Moreover, a multiple regression analysis with self-assessed knowledge gain as dependent variable (see Table 4) yielded significance ($F(2,132) = 6.38$, $p = .002$). However, in contrast to our assumption (H1a), this result indicates that the more micro-actions were performed the lower was self-assessed knowledge gain ($\beta = -.148$, $p = .004$). Besides, a marginal significant effect was found for task-actions ($\beta = .193$, $p = .056$), indicating, according to expectations (H1b), that frequently performed task-action increased self-assessed knowledge gain.

Table 4: Results on the impact of micro- and task-actions on self-assessed knowledge gain

	Self-assessed knowledge gain			
Predictors	β	SE β	R^2	ΔR^2
Micro-actions	-.148*	.050	.088	.074
Task-actions	.193	.100	.088	.074

Effects of interaction frequencies on cognitive load (H2)

To answer the hypotheses on cognitive load (H2), two multiple regression analyses were conducted that addressed mental load and mental effort separately (see Table 5). First, we conducted an analysis with mental load as dependent variable and micro- and task-actions as predictors. The analysis showed a significant result ($F(2,131) = 3.94$, $p = .022$). A closer look at the predictors revealed that task-actions significantly predicted mental load ($\beta = -.435$, $p = .021$). As expected, (H2b), this result indicates that frequently performed task-actions reduce mental load. However, contrary to our expectations (H2a), no effects were found for micro-actions ($p > .05$).

Second, a similar analysis with mental effort as dependent variable did not reach significance ($p > .05$). Thus, we could not confirm our assumptions for mental effort (H2a, H2b).

Table 5: Results on the impact of micro- and task-actions on mental effort and mental load

	Mental load				Mental effort			
Predictors	β	SE β	R^2	ΔR^2	β	SE β	R^2	ΔR^2
Micro-actions	.134	.092	.057	.042	-.092	.068	.018	.003
Task-actions	-.435*	.186	.057	.042	-.113	.137	.018	.003

Discussion

The main goal of this study was to understand how learners' video interactions are related to learning success and cognitive load in natural science learning. To address this goal, we differentiated between micro- and task-actions and analyzed log file data of students' interactions with an enhanced video-based environment.

Our data indicates that frequently performed task-actions predict objective learning success. Following earlier considerations (Zahn et al., 2012), we therefore conclude that meaningful enhanced tools that are an integral and explicit part of the learning task can substantially foster learning processes. This conclusion is in line with related research suggesting that note-taking in enhanced videos is able to increase learning success (Delen et al., 2014) and that designing a hypervideo structure can foster learning of complex topics (Zahn et al., 2010, 2012; Zahn, 2017). The frequent use of enhanced video tools seems to help learners to design their own information structures (e.g., Clark, 1994; Kafai & Resnick, 1996; Yousef et al., 2015) and to actively generate meaning (Wittrock, 1992), which in turn is reflected in actual learning success. However, this could only be confirmed for understanding of concepts, whereas results on other measures (understanding of concept interrelations and transfer knowledge) did not yield significance. Thus, it is arguable that task-actions that are connected to annotations are more involved in fostering understanding of concepts than task-actions that are related to hyperlinks (see Stahl et al., 2006; Zahn et al., 2012, 2010). Future research should consider this by explicitly investigating differences between annotations and hyperlinks and their related actions.

Moreover, in contrast to earlier research (Zahn et al., 2004), we could not confirm a positive relation of frequently performed micro-actions with objective learning success and even found a negative relation with subjective knowledge gain. One possible explanation for these results may be that not the frequent but rather the target-oriented use of basic control tools (manifesting in performed micro-actions) is crucial when learning a complex learning topic (synaptic plasticity) with an enhanced video-based environment. For example, learners who first completely watch the video before starting with the task (interacting with enhanced tools) may need less micro-actions to complete the task than learners who directly start with the task and occasionally need to adapt initial decisions (e.g., skipping through the video to find appropriate places to add an annotation). Hence, frequently performed micro-actions in enhanced video-based environments might not necessarily be related to a deep engagement with the content of the learning material which, in turn, may be reflected in objective learning success and subjective perception of knowledge gain. This example further shows that micro- and task-actions are closely related – not in the sense of a correlation (see results above) – but rather in such a way that basic control tools are often used by learners to *meaningfully* use enhanced tools (e.g., rewind (= micro-action) the video to find an appropriate place to add a hyperlink (= task-action)). Thus, a learner's intention to use an enhanced tool not only includes task- but also micro-actions. Behavior sequence analyses could provide deeper insights into such intentions: learners' micro- and task-actions can be combined into meaningful sequences that can be associated with learning strategies. Such analyses have – although rarely – been considered in previous research on interactive videos (see Sinha, Jermann, Li, & Dillenbourg, 2014). Future research should increasingly exploit the potential of behavior sequence analyses for learning with interactive (video) environments.

Furthermore, we investigated cognitive load (by analyzing both mental load and mental effort) and found a negative relation of task-actions with mental load, indicating that students who frequently performed task-actions perceived the learning material as less difficult than students who made little use of them. In consideration of the above described research (Wittrock, 1992; Zahn et al., 2010, 2012; Zahn, 2017), we conclude that frequently performed task-actions can lead to a deeper understanding of concepts, which in turn can lead to a lower mental load. This could also explain our result suggesting that more performed task-actions increased self-assessed knowledge gain (marginal). However, it is important to note that these findings might also be interpreted in the other direction (e.g., learners who understand the topic more easily have more capacity available to use the enhanced tools meaningfully, which is reflected in a higher number of task-actions). The direction of causality should therefore be specifically considered in future work. Moreover, our results showed that enhanced video tools that are an important and necessary part of the learning task seem not to negatively impact mental effort (no relations found for micro- and task-actions with mental effort). However, these results should be interpreted with caution, as we used a not validated single-item scale. Subsequent studies should use standardized and validated scales specifically created to measure mental effort. Also, to get further insights into the effects of video interaction on cognitive load, future research should consider measurements for intrinsic, extraneous and germane load (with validated instruments, see for example Klepsch et al., 2017).

In sum, we conclude from our results that designs for enhanced video-based learning environments should include tools that are *task-relevant* and explicitly important to the learning task, instead of being optional. The frequent use of such tools seems not only to support learning, but also to reduce cognitive load.

Conclusion

The present study examined the impact of video interactions (micro- and task-actions) of learners who interacted with an enhanced video-based learning environment on learning success and cognitive load. Our results suggest that frequently performed task-actions – that are related to the use of *task-relevant* enhanced tools – not only positively impact learning success but also cognitive load. In summary, our study sheds light on the scientific knowledge about the effects of video tools on learning and leads to important practical implications for designing enhanced video-based learning environments concerning questions of how task instructions and video tools should be integrated. Future research should consider behavior sequence analyses of interaction behavior data for additional in-depth analyses of learning strategies by an equal investigation of micro- and task-actions and their impact on learning success and cognitive load. We hope this contribution will inspire future research in this important area.

References

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, *12*(3), 307–359. https://doi.org/10.1207/S15327809JLS1203_1

Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Springer. https://doi.org/10.1007/b137571

Cattaneo, A., Nguyen, A. T., Sauli, F., & Aprea, C. (2015). Scuolavisione: Teaching and learning with hypervideos in the swiss vocational system. *Journal of E-Learning and Knowledge Society*, *11*(2). https://doi.org/10.20368/1971-8829/1015

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, *42*(2), 21–29.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, *38*(2), 105–134. https://doi.org/10.1007/s11251-009-9110-0

Delen, E., Liew, J., & Willson, V. (2014). Effects of interactivity and instructional scaffolding on learning: Self-regulation in online video-based environments. *Computers and Education*, *78*, 312–320. https://doi.org/10.1016/j.compedu.2014.06.018

Hung, J., & Zhang, K. (2008). Revealing online learning behaviors and activity patterns and making predictions with data mining techniques in online teaching. *MERLOT Journal of Online Learning and Teaching*, *4*(4), 426–437.

Kafai, Y. B., & Resnick, M. (1996). Constructionism in practice: Designing, thinking, and learning in a digital world. In *Constructionism in Practice*. Lawrence Erlbaum Associates.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*(4), 351–371. https://doi.org/10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, J. R. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning*, *13*(2), 213–233. https://doi.org/10.1007/s11412-018-9277-y

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, *8*, 1997. https://doi.org/10.3389/fpsyg.2017.01997

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development*, *42*(2), 7–19. https://doi.org/10.1007/BF02299087

Krauskopf, K., Zahn, C., Hesse, F. W., & Pea, R. D. (2014). Understanding video tools for teaching: Mental models of technology affordances as inhibitors and facilitators of lesson planning in history and language arts. *Studies in Educational Evaluation*, *43*, 230–243. https://doi.org/https://doi.org/10.1016/j.stueduc.2014.05.002

Leisner, D., Zahn, C., Ruf, A., & Cattaneo, A. (2020). Different ways of interacting with videos during learning in secondary physics lessons. In C. Stephanidis & M. Antona (Eds.), *International Conference on Human-Computer Interaction* (pp. 284–291). Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-030-50729-9_40

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press.

Merkt, M., Weigand, S., Heier, A., & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction*, *21*(6), 687–704. https://doi.org/10.1016/j.learninstruc.2011.03.004

Mirriahi, N., & Vigentini, L. (2017). Analytics of learner video use. In *Handbook of learning analytics* (Vol. 1,

pp. 251–267).
Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429
Papert, S. (1994). *Revolution des Lernens : Kinder, Computer, Schule in einer digitalen Welt* (Issue Hannover). (Deutsche Ausgabe von: Brosche & Weigmann). Heidelberg, Heise.
Poquet, O., Lim, L., Mirriahi, N., & Dawson, S. (2018). Video and learning: A systematic review (2007-2017). *ACM International Conference Proceeding Series*, 151–160. https://doi.org/10.1145/3170358.3170376
Rebetez, C., Bétrancourt, M., Sangin, M., & Dillenbourg, P. (2010). Learning from animation enabled by collaboration. *Instructional Science*, *38*(5), 471–485. https://doi.org/10.1007/s11251-009-9117-6
Sauli, F., Cattaneo, A., & van der Meij, H. (2018). Hypervideo for educational purposes: a literature review on a multifaceted technological tool. *Technology, Pedagogy and Education*, *27*(1), 115–134. https://doi.org/https://doi.org/10.1080/1475939X.2017.1407357
Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning*, *8*(4), 375–397. https://doi.org/10.1007/s11412-013-9181-4
Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction*, *14*(3), 293–305. https://doi.org/10.1016/j.learninstruc.2004.06.005
Schwartz, D. L., & Hartman, K. (2007). It is not television anymore: Designing digital video for learning and assessment. In *Video research in the learning sciences* (pp. 335–348).
Shin, Y., & Jung, J. (2020). The effects of a visible-annotation tool for sequential knowledge construction on discourse patterns and collaborative outcomes. *Australasian Journal of Educational Technology*, *36*(4), 57–71. https://doi.org/https://doi.org/10.14742/ajet.4875
Sinha, T., Jermann, P., Li, N., & Dillenbourg, P. (2014). Your click decides your fate: Inferring Information Processing and Attrition Behavior from MOOC Video Clickstream Interactions. *Proceedings of the EMNLP'2014 Workshop*, 3–14.
Stahl, E., Finke, M., & Zahn, C. (2006). Knowledge acquisition by hypervideo design: An instructional program for university courses. *Journal of Educational Multimedia and Hypermedia*, *15*(3), 285–302.
Sweller, J. (1999). *Instructional design in technical areas*. ACER Press.
Wilde, M., Bätz, K., Kovaleva, A., & Urhahne, D. (2009). Testing a short scale of intrinsic motivation. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *15*, 31–45.
Wittrock, M. C. (1992). Generative learning processes of the brain. *Educational Psychologist*, *27*(4), 531–541. https://doi.org/10.1207/s15326985ep2704_8
Yousef, A. M. F., Chatti, M. A., Danoyan, N., Thüs, H., & Schroeder, U. (2015). Video-mapper: A video annotation tool to support collaborative learning in moocs. *Proceedings of the Third European MOOCs Stakeholders Summit EMOOCs*, 131–140.
Zahn, C. (2017). Digital design and learning: Cognitive-constructivist perspectives. In S. Schwan & U. Cress (Eds.), *The Psychology of Digital Learning: Constructing, Exchanging and Acquiring Knowledge with Digital Media.* (pp. 147–170). Springer International Publishing A. https://doi.org/https://doi.org/10.1007/978-3-319-49077-9
Zahn, C., Barquero, B., & Schwan, S. (2004). Learning with hyperlinked videos - Design criteria and efficient strategies for using audiovisual hypermedia. *Learning and Instruction*, *14*(3), 275–291. https://doi.org/10.1016/j.learninstruc.2004.06.004
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. https://doi.org/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. https://doi.org/10.1080/10508401003708399

Acknowledgments

This work was funded by the Swiss National Science Foundation (SNSF) under Grant 100014_176084. The authors acknowledge the financial support of the SNSF.

Automating Characterization of Peer Review Comments in Chemistry Courses

Blair A. Winograd, Amber J. Dood, Solaire A. Finkenstaedt-Quinn,
Anne R. Gere, Ginger V. Shultz
bawinogr@umich.edu, ajdood@umich.edu, quinnsa@umich.edu, argere@umich.edu, gshultz@umich.edu
University of Michigan

Abstract: While writing-to-learn (WTL) pedagogies are a promising way for students to construct knowledge, one limiting factor to implementation is time the instructor spends grading. We conducted two WTL assignments in two undergraduate general chemistry courses combined with collaborative peer review. We used a previously developed scheme to classify peer review comments generated by 1,732 students enrolled in two undergraduate chemistry courses as praise, problem/solution, and verification/summary. Problem/solution comments were further separated into greater-level, mid-level, and word-sentence descriptors. Using the SciBERT language model we then developed a classifier which accurately identifies comments where human coding was considered the ground truth. In the future, this model can provide an efficient way for instructors to monitor peer review collaborations and help instructors use peer-generated insights to guide their instruction.

Introduction

Writing-to-learn (WTL) activities are typically short, low-stakes writing tasks that help students think through key concepts or ideas presented in a course. WTL interventions have been shown to help elicit deep levels of reasoning necessary to make meaningful connections about concepts. WTL interventions, however, create a logistically challenging situation for instructors to provide timely and effective feedback on student writing. A collaborative approach to WTL introduces a social aspect that can help students connect concepts and guide each other's revisions (Reynolds et al., 2012). Peer review is one solution to solve the logistical challenge for instructors to provide feedback; its contents may also be a rich resource of information for instructors. Student peer review may provide a filtered perspective through which an instructor can gain insight about the learning experiences of their students and alternative conceptions their students are expressing through written work. Through peer review, students can be provided with feedback on their writing and have the opportunity to learn through reading and providing feedback on other students' writing (Finkenstaedt-Quinn et al., 2019). An efficient way to monitor the quality of student interactions could be useful for instructors; instructors could also use the content of these peer interactions to influence their instruction. Using natural language processing (NLP), peer review comments can be automatically classified to provide instructors with a timely, new viewpoint of their students' understanding. We use machine learning methods to develop a generalized model which can classify peer review comments in an undergraduate organic chemistry course and a general chemistry course without needing to be rebuilt and retrained for specific assignments or courses. The model has potential for expansion to include peer review assignments in other disciplines.

Background

WTL pedagogies have shown evidence of effectiveness as a learning tool across disciplines for personal construction of knowledge (Bangert-Drowns et al., 2004). The purpose of these assignments is to help students build connections between facts, promoting a deep conceptual understanding of topics (Reynolds et al., 2012). WTL has been used widely in STEM fields; examples of widely-implemented WTL programs include the Science Writing Heuristic (Hand et al., 1999; Keys et al., 1999) and Calibrated Peer Review (Russell et al., 2017; Walvoord et al., 2008). Recently, the effectiveness of WTL in chemistry courses has been studied with assignments about acid-base concepts, light-matter interactions, and polymer properties (Finkenstaedt-Quinn et al., 2017; Moon et al., 2018; Schmidt-McCormack et al., 2019). The results of these studies indicated that the WTL assignments helped students better understand and explain the concepts covered in the assignments. It has been found that the most effective WTL assignments include four main aspects: the assignment is a meaning-making activity, the writing process is interactive, the assignment has clear expectations, and involvement of aspects of metacognition (Anderson et al., 2015; Bangert-Drowns et al., 2004; Gere et al., 2018). While meaning-making, assignment expectations, and metacognition aspects are dependent upon the assignment design, interactivity depends on how the assignment is implemented. One way to make the assignment interactive is through peer review (Finkenstaedt-Quinn et al., 2019; Halim et al., 2018; Pelaez, 2002). Peer review can help students in courses with large enrollment receive feedback on their writing even though it is not feasible for the instructor to provide feedback to all students (Kulkarni et al., 2013). Additionally, studies have shown that students benefit even more from providing feedback to other students than they do from receiving feedback (Lundstrom et al., 2009). Though peer review ideally eliminates the need for individualized instructor feedback, it is still helpful for instructors to see what students are writing about in their peer review, both as a way to ensure their students are providing high quality feedback to each other and to identify common inaccuracies in student writing. Instructors can be provided with timely feedback about their students' peer review comments at the classroom level through automated formative feedback.

In most automated feedback schemes for writing, the goal is to use NLP to analyze text written in response to a specific prompt. Symbolic NLP has been widely researched in STEM disciplines to provide both formative and summative feedback on constructed response items for which models are specifically trained (Dood et al., 2018, 2020; Haudek et al., 2012; Knight et al., 2020; Moharreri et al., 2014; Noyes et al., 2020). Compared to constructed response items, WTL pedagogies ask students to provide lengthier and less focused responses to a prompt which produces highly unstructured textual responses. Developments in the field of NLP have led to improvements in the ability to analyze unstructured texts; yet, even the most sophisticated NLP methods are unlikely to provide effective feedback on this type of writing. One way to circumvent this could be to focus the NLP tasks on the slightly more structured peer review comments students give to each other.

Modern developments in the field of NLP include language models, called transformers, that pre-train the representation of language based on large amounts of text. On top of these pre-trained descriptors of transformers, bidirectional training can be employed, such as is used for Bidirectional Encoder of Representations and Transformers (BERT) (Devlin et al., 2018). This approach can combat one of the more prevalent issues with NLP: lack of sufficient training data to solve the problem at hand. The BERT language model is unique because it was trained on a very large dataset of approximately 3,200M words and takes into account the context of words within sentences rather than just looking at single words. The general BERT model was developed on a breadth of literature including novels, newspapers, and Wikipedia. Other models that are more specific to science writing have been developed using text from scientific journals, such as sciBERT (Beltagy et al., 2019). The sciBERT model was built from a corpus size of 1.14M papers and 3.1B tokens from full scientific texts found at Semantic Scholar (www.semanticscholar.org). Even with more advanced language models, providing feedback on standalone essays is far-fetched; however, the structured nature of peer review comments for WTL assignments provides an avenue to be explored.

A study by Dixon and Moxley (2013) looked at a large sample of instructor feedback on student writing and found that instructors were primarily focused on higher-order concerns with student writing (e.g., use of reasoning, accuracy of content) rather than lower-order concerns (e.g., grammatical errors). Yet, when instructors are dealing with a multitude of long writing assignments to score, they often find themselves only marking for low-order problems. Studies have also shown that computer support may help instructors focus on higher-order concerns. Online tools have been developed to facilitate peer review, such as Calibrated Peer Review (Russell et al., 2017; Walvoord et al., 2008) and MyReviewers (Branham et al., 2015), but these tools do not provide instructors with an understanding of conceptual problems prevalent in a class.

The goal of this project is to be able to automate detection of the higher-order concerns provided by student peer reviewers and provide instructors with insight into the type of feedback their students are providing to other students. We adopted a scheme by Finkenstaedt-Quinn et al. (2019) which hierarchically characterizes peer review comments in an undergraduate general chemistry course as verification, summary, praise, and problem/solution as well as delineates between higher and lower-order feedback. We applied a similar coding scheme to peer review comments on WTL assignments in an undergraduate general chemistry course and an undergraduate organic chemistry course. Using SciBERT, we explore automating the characterizations of peer review comments within the coding scheme.

Research questions

This work is guided by one question: To what extent can characterization of peer review comments (e.g., as verification, summary, praise, problem/solution and higher and lower-order) in different courses be automated?

Methods

Data acquisition

This study included analysis of peer review comments and student writing from large undergraduate courses in general and organic chemistry at a large, public research university. While in the same discipline, these two courses were chosen due to differences in the nature of their content. Specifically, the content in general chemistry courses is very quantitative, while organic chemistry courses focus on qualitative aspects of the domain.

In both courses, students were given a three-part WTL assignment which included an initial draft, a peer review portion, and a revised, final draft. Each student reviewed the assignments of three other students via an in-house peer review tool and received feedback from one to four other students. Students were then asked to revise their writing based on feedback and submit a revised draft. Students in the general chemistry course received completion credit for the writing assignment and peer review. Assignments in the organic chemistry course were graded with a low-stakes grading scheme; assignments in the general chemistry course were graded for completion. In the general chemistry course, students were given a WTL assignment that involved reading and summarizing a 1916 paper written by Gilbert Lewis proposing what was at the time a new method for modeling molecular structures (Lewis, 1916). In this activity, students were provided with a detailed assignment description requiring them to write about specific molecular concepts. Peer reviewers were guided by a rubric that prompted students via six questions, five of which directed students to provide feedback on specific content areas and one of which focused more on writing style and mechanics. A similar WTL structure was implemented in the organic chemistry course as the first of three WTL assignments. Students were provided with a base-free Wittig reaction from the research literature and asked to compare the reaction with the traditional Wittig reaction (Schirmer et al., 2015). Detailed requirements for the assignment asked students to describe the reaction mechanisms at a molecular level and explain why the base-free reaction works under one set of conditions but not another. Peer reviewers in the organic course were prompted by a rubric requiring them to respond to four questions about different areas of content. IRB approval was obtained to collect and use student data and every student enrolled in the course provided consent for our use of their responses.

Characterization of peer review

Peer review comments were first characterized in the general chemistry course. The process of characterization was completed as described by Finkenstaedt-Quinn et al. (2019). First, we considered the codes present at the top of the hierarchy of the scheme: praise, verification, summary and problem/solution. Due to the imbalanced and uncommon occurrence of student revision based on type of feedback, we chose to combine verification and summary into a single label. These two types of feedback could be considered "neutral" as they do not promote revision or provide the instructor with feedback about the written work. We also chose to study the distinction between peer reviews that suggest lower level versus higher level problems in the student writing. This distinction is described by the scope of problem and scope of solution codes given in the characterization scheme by Finkenstaedt-Quinn et al. Of these types of comments, whether the peer reviewer described only the problem, only the solution, or both, little distinction of the occurrence of revision was present; therefore, in this work we chose to combine code types scope of problem greater-level with scope of solution greater-level, scope of problem mid-level with scope of solution mid-level, and so on. A summary of the scheme applied in this work is given in Figure 1.

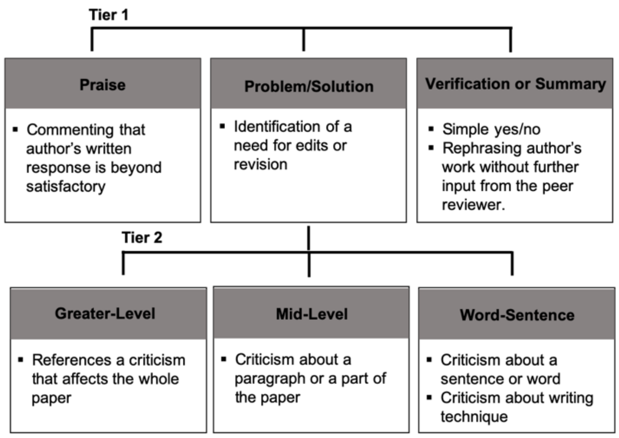


Figure 1. Hierarchical structure and short definition of coding scheme used. All peer review comments are filtered into three labels: praise, problem/solution, and verification or summary. All problem solution comments are further filtered into three more labels: higher-level, mid-level, and word-sentence.

Overall, 1,132 comments from the general chemistry course were characterized based on the Figure 1 scheme.

To explore how well the existing coding scheme can apply to other coursework as well as how well the algorithm transfers to a unique set of data, we also coded a new set of 600 peer review comments from the organic chemistry course. These 600 comments were chosen at random out of 9,145 total peer review comments in the WTL intervention; however, the comments in this set always contained the full collection of reviews one peer reviewer gave to an author. Comments were also coded for the labels listed above. Two raters coded all 600 peer reviews and Krippendorff's alpha was calculated to be 0.90 (Krippendorff, 2011).

Computational Approach

Because the characterization scheme is hierarchical, allowing for only single labels, this provides a clear environment to use a multi-classification labelling method. Due to the lack of structure and variety in comments, we chose to use fine-tuning with the SciBERT model which encodes language bi-directionally on top of the pre-trained SciBERT transformer model with a linear layer for text sequence classification on top. The input peer review comments, for which one response to one criterion was considered one input, were first tokenized with the BERT basic tokenizer to perform punctuation splitting, lowercasing and invalid character removal. The longest comment in the training set contained 168 tokens. The maximum sequence length was defined as 175 tokens to ensure that no textual data would be lost. Shorter sequences were padded with zeros and longer sequences were truncated to the maximum sequence length.

Results

First, the 1,132 peer review comments from the general chemistry course were partitioned into a training set and testing set. We chose an approximate split of 80% training set (900 comments) and 20% testing set (232 comments). The training set was further split into a 90% training and 10% validation set which is a common practice when training neural networks which work in epochs. The training set was fed into the network. Three epochs were chosen as this was the smallest gap achieved where validation loss overcame the training loss. Then, the test data were fed into the network in batches of 32 in the spirit of k-fold resampling to allow us to observe any unexpected results within the data. An overall Matthew's Correlation Coefficient (MCC) score of 0.906 was achieved with the individual MCC scores of each batch provided in Figure 2 (the MCC score can be read comparably to Cohen's Kappa but is more trustworthy to describe imbalanced data sets). The confusion matrix of the results is provided in Table 1. The overall accuracy for each type of comment is as follows: summary/verification = 0.96 praise = 0.95, and problem/solution = 0.97.

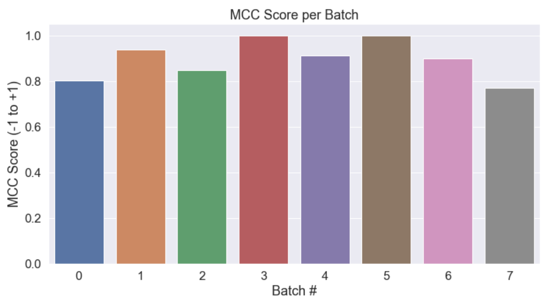


Figure 2. Batched view of MCC scores for the types of comments in the general chemistry course.

Table 1. Confusion matrix for classification of types of comments in the general chemistry course (testing set).

	Summary/verification	**Praise**	**Problem or solution**
Summary/verification	**33**	**3**	**0**
Praise	**4**	**64**	**1**
Problem/solution	**2**	**3**	**122**

Based on the accuracy of these results, we then considered further filtering the problem/solution comments based on the scope label. Of the 1,132 peer review comments, 545 of the comments were problem/solution type. Each of these were coded by hand and identified either greater-level, mid-level, or word/sentence level scope. In this work, the data shows great imbalance in the types of codes where there were many instances of praise and problem/solution, but fewer summary and verification. For the purpose of training, imbalance is best avoided. Given previous work that suggests instructors struggle to give formative feedback far beyond the word sentence level (Dixon & Moxley, 2013), we can assume that just separating word/sentence feedback compared to mid- and greater-level feedback would be beneficial. We combine the mid-level and greater-level feedback into one code for this work.

Similar to what was done with the full set, this subset of 545 comments were broken down into an approximate 80% training and 20% testing split. Again, the training set was further split into a 90% training and 10% validation set. The training set was fed into the algorithm. Three epochs were chosen as this was the smallest gap achieved where validation loss overcame the training loss. Then the test data were fed into the network in batches of 32 in the spirit of k-fold resampling to allow us to observe any unexpected results within the data. An overall MCC score of 0.814 was achieved with the individual MCC scores of each batch given in Figure 3. The confusion matrix of the results is given in Table 2.

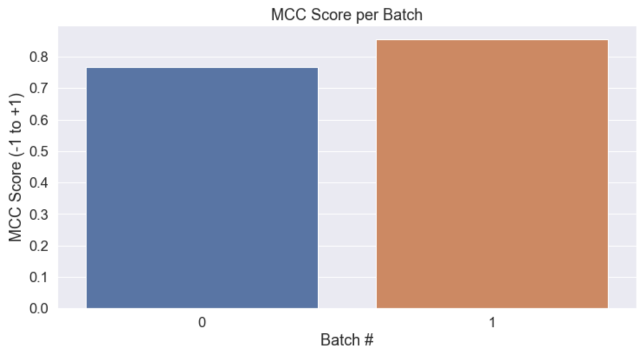


Figure 3. A batched view of the MCC scores for the sample of problem/solution type comments from the general chemistry course.

Table 2. Confusion matrix for the problem/solution type comments from the general chemistry course.

	Greater and mid- level	**Word-sentence**
Greater and mid- level	**45**	**0**
Word-sentence	**5**	**14**

Finally, to view how well the algorithm transfers to a completely new set of data, we considered the 600 comments coded from the organic chemistry course. In this case, the algorithm was not trained on data from the same peer review rubric. Instead, the weights determined from training to identify problem/solution, verification/summary and praise comments in the general chemistry course were applied to the comments from the organic chemistry course. In the organic chemistry course, the rubric guiding students' peer review was more scaffolded, resulting in many instances of peer review beyond 175 tokens. When creating the embeddings for these sentences, we chose a max length of 300 tokens, even though a small number of peer review comments exceeded this. As the number of tokens increases, the time for training also increases. Cutting off at 300 tokens provided a balance between an efficient testing period and inclusion of data in longer peer review comments.

The test data were fed into the network in batches of 32 via k-fold resampling to allow us to observe any unexpected results within the data. An overall MCC score of 0.602 was achieved with the individual MCC scores of each batch, given in Figure 4. The confusion matrix of the results is in Table 3.

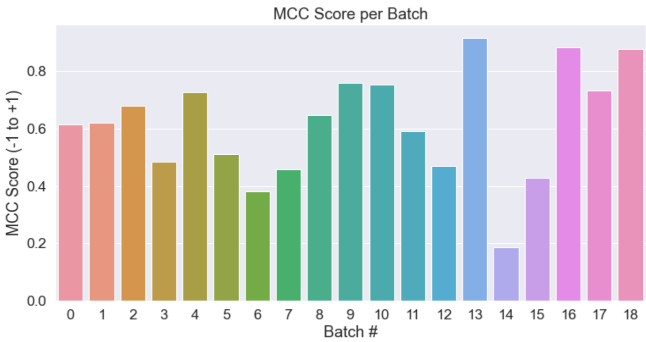


Figure 4. Batched view of MCC scores for the types of peer review comments in the organic chemistry course.

Table 3. Confusion matrix for characterization of types of comments in the organic chemistry course.

	Summary/Verification	**Praise**	**Problem or Solution**
Summary/Verification	**28**	**1**	**11**
Praise	**26**	**50**	**40**
Problem/Solution	**4**	**12**	**428**

Conclusions

Overall, the results of this work support a positive outcome of the research question: To what extent can this characterization scheme for peer review comments in different courses be automated? In the general chemistry course, identification by the computer of the first-tier type of comments was achieved to a very high, almost perfect level (MCC = 0.906). The work supports previous work (Finkenstaedt-Quinn et al., 2019) in that it confirms that the coding scheme suggested can depict peer review in two unique chemistry writing contexts even when the scientific content present in the essays is substantially different. Without training a new set, we transferred the weights from the general chemistry course training set (with characterizations described by tier 1 of Figure 1) and achieved a substantial level of accuracy (MCC = 0.602) to identify peer review types in the organic chemistry course. This introduces promising support that the coding scheme and model may also be applicable to peer review in other natural science, engineering and math courses. We could further identify peer review comments described by the second-tier of Figure 1 in the general chemistry course to an almost perfect level of accuracy (MCC = 0.814).

Considering our effort is not to grade high-stakes assignments on the individual level but rather to provide formative feedback to instructors about student learning in the class, the accuracy of this model is more than satisfactory. This is also promising considering we were able to use a relatively small amount of training data (i.e., 900 comments). Using most standard machine learning models with non-bidirectional encodings of language, typical algorithms require tens of thousands of data points in a training set in order to produce a suitable model. This work supports other findings showing that BERT is able to make meaningful predictions even with smaller data sets. While BERT may seem like a computationally difficult and expensive approach for identifying problems, it is useful here due to the relatively small sets of data being used. Additionally, when we begin to consider more features prevalent in useful peer review, the BERT approach will likely prove useful for the subtleties that will present themselves.

Currently, our model has only been tested in two different contexts. Though the scope is currently limited, the success of the model when evaluating peer review in a second context without additional training on the new set is promising for the broader applicability of this model. Further work should be done to expand the training set to contain instances of human-labeled peer review from several different courses.

Automatic characterization of peer review can be beneficial in many ways. This benefit can be honed on two sides of the classroom: for students and for instructors. For students, automated characterization of the content within a single peer review comment can provide support to peer reviewers about whether they are providing high quality and supportive peer review to the author. Immediate feedback can serve as a coach to support students as they learn to provide quality reviews, a skill that is typically not formally taught in undergraduate or graduate curricula. Additionally, students with problems in their writing may need to be reminded to revisit and revise their work. If we can automate detection of occurrences of problematic peer review, an automatic reminder system could be developed to encourage students to edit their work.

Instructors may be limited by class size and unable to read the content of all peer reviews. This could inhibit their ability to make inferences about what their students know and do not know. Based on our work, we envision a type of automated dashboard or roadmap of the peer review process. Instructors could be given all the peer review feedback in a type of filtration system where they can choose to view specific filters of the types of peer review. A tool like this could help an instructor focus on greater-level problems pointed out in student writing, allowing them to intervene during lectures to redeliver or revisit material that is consistently incorrect in the class. Furthermore, the instructor may be given flags on students who are doing exceptionally well, students who are struggling with alternative conceptions, students who are struggling with writing, or even students who are going above and beyond in the peer review process. As course sizes continue to increase, instructors will need more tools to identify and address problems in a class. Furthermore, many classes within the United States in 2020 have gone virtual; in this case, instructors may have even less feedback from their students. In person, instructors have the ability to catch a glimpse of confusion or distress when teaching concepts. Virtually, more tools are required to catch these instances. Being able to efficiently view and understand what is happening in students' writing and collaborations is an important and efficient way to provide instructors with opportunities to deliver quality instruction.

References

Anderson, P., Anson, C. M., Gonyea, R. M., & Paine, C. (2015). The Contributions of Writing to Learning and Development: Results from a Large-Scale Multi-institutional Study. *Research in the Teaching of English*, *50*(2), 199–235.

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The Effects of School-Based Writing-to-Learn Interventions on Academic Achievement: A Meta-Analysis, *Review of Educational Research*, *74*(1), 29–58. https://doi.org/10.3102/00346543074001029

Beltagy, I., Lo, K., & Cohan, A. (2019). SciBERT: A Pretrained Language Model for Scientific Text. *ArXiv:1903.10676 [Cs]*. http://arxiv.org/abs/1903.10676

Branham, C., Moxley, J., & Ross, V. (2015). My reviewers: Participatory design & crowd-sourced usability processes. *Proceedings of the 33rd Annual International Conference on the Design of Communication - SIGDOC '15*, 1–6. https://doi.org/10.1145/2775441.2775482

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *CoRR*, *abs/1810.04805*. http://arxiv.org/abs/1810.04805

Dixon, Z., & Moxley, J. (2013). Everything is illuminated: What big data can tell us about teacher commentary. *Assessing Writing*, *18*(4), 241–256. https://doi.org/10.1016/j.asw.2013.08.002

Dood, A. J., Dood, J. C., Cruz-Ramírez de Arellano, D., Fields, K. B., & Raker, J. R. (2020). Analyzing explanations of substitution reactions using lexical analysis and logistic regression techniques. *Chemistry Education Research and Practice*, *21*(1), 267–286. https://doi.org/10.1039/C9RP00148D

Dood, A. J., Fields, K. B., & Raker, J. R. (2018). Using Lexical Analysis To Predict Lewis Acid–Base Model Use in Responses to an Acid–Base Proton-Transfer Reaction. *Journal of Chemical Education*, *95*(8), 1267–1275. https://doi.org/10.1021/acs.jchemed.8b00177

Finkenstaedt-Quinn, S. A., Snyder-White, E. P., Connor, M. C., Gere, A. R., & Shultz, G. V. (2019). Characterizing Peer Review Comments and Revision from a Writing-to-Learn Assignment Focused on Lewis Structures. *Journal of Chemical Education*, *96*(2), 227–237. https://doi.org/10.1021/-acs.jchemed.8b00711

Finkenstaedt-Quinn, Solaire A., Halim, A. S., Chambers, T. G., Moon, A., Goldman, R. S., Gere, A. R., & Shultz, G. V. (2017). Investigation of the Influence of a Writing-to-Learn Assignment on Student Understanding of Polymer Properties. *Journal of Chemical Education*, *94*(11), 1610–1617. https://doi.org/10.1021/-acs.jchemed.7b00363

Gere, A., Knutson, A., Limlamai, N., McCarty, R., & Wilson, E. (2018). A Tale of Two Prompts: New Perspectives on Writing-to-Learn Assignments. *ETSU Faculty Works*, *29*, 147–188.

Halim, A. S., Finkenstaedt-Quinn, S. A., Olsen, L. J., Gere, A. R., & Shultz, G. V. (2018). Identifying and Remediating Student Misconceptions in Introductory Biology via Writing-to-Learn Assignments and Peer Review. *CBE Life Sciences Education*, *17*(2), ar28. https://doi.org/10.1187/cbe.17-10-0212

Hand, B., Lawrence, C., & Yore, L. D. (1999). A writing in science framework designed to enhance science literacy. *International Journal of Science Education*, *21*(10), 1021–1035. https://doi.org/10.1080/095006999290165

Haudek, K. C., Prevost, L. B., Moscarella, R. A., Merrill, J., & Urban-Lurain, M. (2012). What are they thinking?

Automated analysis of student writing about acid-base chemistry in introductory biology. *CBE Life Sciences Education*, *11*(3), 283–293. https://doi.org/10.1187/cbe.11-08-0084

Keys, C. W., Hand, B., Prain, V., & Collins, S. (1999). Using the Science Writing Heuristic as a Tool for Learning from Laboratory Investigations in Secondary Science. *Journal of Research in Science Teaching*, *36*(10), 1065–1084. https://doi.org/10.1002/(SICI)1098-2736(199912)36:10<1065::AID-TEA2>3.0.CO;2-I

Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., Wight, R., Lucas, C., Sandor, A., Kitto, K., Liu, M., Mogarkar, R. V., & Buckingham Shum, S. (2020). AcaWriter: A learning analytics tool for formative feedback on academic writing. *Journal of Writing Research*, *12*(1), 141–186.

Krippendorff, K. (2011). Computing Krippendorff's Alpha-Reliability. *Departmental Papers (ASC)*. https://repository.upenn.edu/asc_papers/43

Kulkarni, C., Wei, K. P., Le, H., Chia, D., Papadopoulos, K., Cheng, J., Koller, D., & Klemmer, S. R. (2013). Peer and self assessment in massive online classes. *ACM Transactions on Computer-Human Interaction*, *20*(6), 33:1–33:31. https://doi.org/10.1145/2505057

Lewis, G. N. (1916). The Atom and the Molecule. *Journal of the American Chemical Society*, *38*(4), 762–785. https://doi.org/10.1021/ja02261a002

Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18(1), 30-43.

Moharreri, K., Ha, M., & Nehm, R. H. (2014). EvoGrader: An online formative assessment tool for automatically evaluating written evolutionary explanations. *Evolution: Education and Outreach*, *7*(1), 15. https://doi.org/10.1186/s12052-014-0015-2

Moon, A., Zotos, E., Finkenstaedt-Quinn, S., Gere, A. R., & Shultz, G. (2018). Investigation of the role of writing-to-learn in promoting student understanding of light–matter interactions. *Chemistry Education Research and Practice*. https://doi.org/10.1039/C8RP00090E

Noyes, K., McKay, R. L., Neumann, M., Haudek, K. C., & Cooper, M. M. (2020). Developing Computer Resources to Automate Analysis of Students' Explanations of London Dispersion Forces. *Journal of Chemical Education*. https://doi.org/10.1021/acs.jchemed.0c00445

Pelaez, N. J. (2002). Problem-based writing with peer review improves academic performance in physiology. *Advances in Physiology Education*, *26*(3), 174–184. https://doi.org/10.1152/advan.00041.2001

Reynolds, J. A., Thaiss, C., Katkin, W., Thompson, R. J., & Wright, R. L. (2012). Writing-to-Learn in Undergraduate Science Education: A Community-Based, Conceptually Driven Approach. *CBE—Life Sciences Education*, *11*(1), 17–25. https://doi.org/10.1187/cbe.11-08-0064

Russell, J., Horne, S. V., Ward, A. S., Bettis, E. A., & Gikonyo, J. (2017). Variability in students' evaluating processes in peer assessment with calibrated peer review. *Journal of Computer Assisted Learning*, *33*(2), 178–190. https://doi.org/10.1111/jcal.12176

Schirmer, M.-L., Adomeit, S., & Werner, T. (2015). First Base-Free Catalytic Wittig Reaction. *Organic Letters*, *17*(12), 3078–3081. https://doi.org/10.1021/acs.orglett.5b01352

Schmidt-McCormack, J. A., Judge, J. A., Spahr, K., Yang, E., Pugh, R., Karlin, A., Sattar, A., Thompson, B. C., Gere, A. R., & Shultz, G. V. (2019). Analysis of the role of a writing-to-learn assignment in student understanding of organic acid–base concepts. *Chemistry Education Research and Practice*. https://doi.org/10.1039/C8RP00260F

Walvoord, M. E., Hoefnagels, M. H., Gaffin, D. D., Chumchal, M. M., & Long, D. A. (2008). An Analysis of Calibrated Peer Review (CPR) in a Science Lecture Classroom. *Journal of College Science Teaching*, *37*(4), 66–73.

Acknowledgements

We would like to acknowledge the undergraduate students who helped in the IRR process of coding the data and the undergraduate students who participated in the WTL assignments and peer review.

Problematic Interaction Patterns During Online-Collaboration. A Library and a Survey

Sebastian Strauß, Nikol Rummel
sebastian.strauss@rub.de, nikol.rummel@rub.de
Ruhr-Universität Bochum

Abstract: Implementing collaborative learning into online courses can help mitigate central challenges for this learning setting which often stem from a lack of interaction in the course. During online-collaboration, however, learners can experience a number of interaction patterns that not only reduce the effectiveness of the collaboration but also lead to frustration. If learners are frustrated with the collaboration they are prone to reducing their participation or even exiting the learning setting. As collaborative learning can only be effective if all group members interact with each other, frustrating interaction patterns pose a challenge to the effectiveness of collaborative learning. We compiled a library consisting of 14 potentially frustrating interaction patterns through a literature-review and an analysis of collaboration data (n = 10 groups). Conducting an online survey among university students (n = 100) revealed that social loafing, deadline rush, ineffective communication, and unequal participation are the most severe interaction patterns.

Introduction

Distant online education has received increasing attention over the recent years. Not only the development of Massive Open Online Courses (MOOCs, Baturay, 2015; Pappano, 2012) in the 2010s but especially the recent COVID-19 pandemic has reinvigorated the discussion about effective instruction in online settings (e.g., Reynolds & Chu, 2020). Notably, online courses face a number of challenges such as feelings of isolation that result from a lack of social interaction (Khalil & Ebner, 2014). Implementing collaborative learning activities is one way to mitigate these challenges (Rosé, Goldman, Zoltners Sherer, & Resnick, 2015). In collaborative learning, interaction between learners is a conditio sine qua non (cf. "interaction paradigm", Dillenbourg, Baker, Blaye, & O'Malley, 1996) as the majority of processes during collaboration and especially those processes that are conducive to learning are based on interaction between the group members. Besides interaction that is required to develop as a group (Tuckman & Jensen, 1977), coordinate (e.g., Wittenbaum, Vaughan, & Strasser, 2002) or distribute roles (Strijbos & Weinberger, 2010), group members also need to achieve a shared understanding regarding concepts or the task (Clark & Brennan, 1991; Hadwin, Bakhtiar, & Miller, 2018), develop a transactive memory system (Wegner, 1995), and group awareness (Bodemer & Dehler, 2011), or resolve conflicts (Darnon, Doll, & Butera, 2007). Further, students have to pool information to make informed decisions (Stasser & Titus, 1985), co-construct new knowledge (Weinberger & Fischer, 2006), and engage in interactions that benefit learning (e.g., explaining or cognitive modelling, cf. King, 2007). During collaboration, a group also monitors and regulates the interaction processes, and the motivation and emotions of the group members (Järvelä et al., 2016). All these processes are challenging, especially for unexperienced learners. Unsurprisingly, productive interactions in groups do not automatically occur simply because students are assigned to a team (Kreijns, Kirschner, & Jochems, 2003). Further challenges result from interactions that are not only unproductive but can also cause frustration. Frustration can be defined as a negative affective state that results if a person is unable to achieve their goal or if expectations are not met (Capdeferro & Romero, 2012; Ortony, Clore, & Foss, 1987). The impact of interaction patterns that cause frustration should not be overlooked because dissatisfied learners may reduce their participation or even drop out of the learning setting (Khalil & Ebner, 2014; Levy, 2007). Consequently, there would be less opportunities for those interactions that are necessary for successful collaboration.

Despite their pivotal role during collaboration, affective variables such as frustration have received less attention in CSCL research in comparison to cognitive variables (Jeong, Hmelo-Silver, & Jo, 2019). Only few studies consider satisfaction and frustration during collaborative learning and often neglect to elaborate on the role of these variables during collaboration. Against this background, we aim at shedding light on interaction patterns that are unproductive and frustrating for students because this knowledge empowers researchers, practitioners, and educators to consider these unfavorable interaction patterns and design online learning activities to mitigate the challenges that arise from these patterns or develop collaboration support to promote students' collaboration competence (i.e., internal collaboration scripts, cf. Fischer, Kollar, Stegmann, & Wecker, 2013) so they can engage in fruitful collaboration. The present study has two goals. First, we provide an overview of potentially frustrating interaction patterns that can occur during online collaboration in small groups. In this regard, we extend the preliminary findings of a prior study that used a similar procedure (Strauß, Rummel, Stoyanova, & Krämer, 2018). Secondly, we report on the results of a survey that sought to determine which of these interaction patterns are the most severe and thus may be considered as primary targets for instructional design and support.

A library of problematic interaction patterns in online-collaboration

To compile a list of potentially frustrating interaction patterns, we conducted a literature review on Google Scholar and Web of Science using the search string '*(collaborative learning OR Collaboration OR group) AND (satisf* OR frustr*) AND interaction AND (online OR e-learning)*'. We scanned the titles and abstracts if the publication included student-student interactions and students' affects (e.g., frustration, satisfaction, positive group climate) during collaboration (work or learning). During this process we were intentionally over-inclusive when selecting article because affective variables are rarely assessed in conjunction with interaction patterns. We did not consider publications that only reported on the relationships between interaction and variables such as learning gains, and publications that dealt with the relationship between affects and other forms of interaction (e.g., teacher-learner or learner-content interaction).

After selecting publications, we read them closely and discarded articles that did not provide enough information to derive specific interaction patterns. While reading the articles, we identified additional references and scanned them using the procedure described above. If a publication described one or more interaction patterns that are associated with frustration, we added them to our list. If a publication described interactions related to satisfaction, the opposite or absence of this particular interaction was added to the library since the absence of beneficial interactions also constitutes a challenge.

In addition to the literature review, we analyzed collaboration data to potentially identify additional challenges that were not yet covered by the literature review. Therefore we analyzed a small sample of log files from an online course (reported as Course 2 in Erdmann et al., 2017). During this online course, students worked on collaborative assignments using a group forum for communication and a shared text-editor to construct a joint answer text on Moodle. We randomly selected ten out of 55 groups and analyzed their log files (written text, time stamps and number of words). Two raters independently analyzed the logs for all ten groups and listed situations where students expressed frustration or contentment with the collaboration process. For example, one student wrote in the group forum that they found it difficult to know if a group member had already completed their current task. Afterwards, the raters compared their lists. We only added interactions to the library that both raters had identified. To finalize this step, we performed a literature search to theoretically ground the situations identified in the collaboration data.

Problematic interaction patterns

This procedure resulted in a list of fourteen situations that we assumed to affect students' satisfaction with the collaboration (cf. Table 1). In the following we present these interaction patterns.

Communication

Six interaction patterns concern communication during collaboration. The first potential source of frustration is *impolite communication* (Park, 2008). In her study, Park (2008) reports a positive correlation between polite communication and satisfaction. Consequently, we listed the opposite, that is, impolite communication. Following Kellermann and Park (2001), a person acts politely if they behave "mannerly, courteous, and respectful" (Kellermann & Park, 2001, p. 4). From this perspective, impolite communication refers to neglecting socially constructed expectations of what constitutes manners, being cold, unfriendly, ignorant, and disrespectful during collaboration. Similarly repeated 'nagging' or '*nudging*' fellow group members to make them contribute their fair share to the task can also be expected to have a negative effect on the group climate and satisfaction (Walther, 1996).

Another potentially frustrating interaction pattern is *inefficient communication*. We derived this interaction pattern from a cluster analysis reported by Kwon and colleagues (2014) who identified three types of groups, namely 'Late Collaborator', 'Passive Task-oriented Collaborator' and 'Early Active Collaborator'. In comparison, groups classified as Early Active Collaborator exhibited a high number of activities that targeted group-regulation and social-emotional behavior. Besides a higher sense of community, groups in this cluster also perceived a more positive group climate than the groups in the other clusters. One communication-related aspect of the groups in the 'Early Active Collaborator' cluster was efficient communication. Thus, we added the opposite, inefficient communication, to our library. Inefficient communication can be defined as communication that does not focus on the current task (Park, 2008). Following Park's (2008) argumentation, inefficient communication violates Grice's (1975) four conversational maxims, quantity, quality, relation and manner. From this perspective, the communication in a group would be considered *inefficient* if (1) group members sent messages that contained too little or too much information for the current purpose of the conversation (quantity), (2) group members sent messages that were mere assumptions that lack evidence, false messages or messages that did not reflect the sender's beliefs (quality), (3) group members sent messages that were irrelevant for the current topic (relation) or (4) if group members sent obscure, ambiguous, wordy and disorganized messages (manner).

Another communication-related interaction pattern concerns the flow of communication. While timely responses help build a sense of community (Sung & Mayer, 2012), our analysis of interaction as well as existing studies revealed that *long waiting times for replies* were perceived as frustrating by the students (Draskovic, Holdrinet, Bulte, Bolhus, & van Leeuwe, 2004; Goold, Craig, & Coldwell, 2008).

Another interaction pattern that we identified in the collaboration data was that students *neglect to communicate when they will be able to work on the task*. This creates obscurity which violates Grice's maxim of manner (Grice, 1975). This interaction often occurred at the beginning of the group task. If students are not aware of when group members plan to start working on the task and how much time they will be able to invest, the other group members may mistake a group member's intended absence (e.g., due to a competing deadline) as lack of engagement (social loafing, Aggarwal & O'Brien, 2008). Also, coordinating the collaboration becomes more difficult if not all members make their individual time constraints transparent.

The last potentially frustrating interaction pattern in this category are *phases of no interaction or communication.* We derived this interaction pattern from the cluster analysis reported by Kwon and colleagues (2014) who found that group members in the 'Early Active Collaborator' cluster continuously interacted with the other members of their group, thus, we added the inverse to our list of potentially frustrating interaction patterns.

Information processing and decision making

The following four interaction patterns relate to information processing and decision making. Probably the most widely reported challenge and source of frustration for groups is *social loafing*, that is, the tendency for individuals to exert less effort in a group as compared to when working alone (Aggarwal & O'Brien, 2008; Capdeferro & Romero, 2012; Karau & Williams, 1993). Although these group members contribute less to the joint work, they often still receive the same rewards as the other group members ('free-riding') (Aggarwal & O'Brien, 2008). The bottom-up analysis of our sample of collaboration data revealed that social loafing in the form of a large difference in participation (measured as the number of words) was not as frequent as a somewhat uneven distribution of participation. Thus, in the present study we distinguished between *unequal participation*, that is, group members contribute varying amounts to the joint task, and *social loafing* that is, individual group members exert only very little or no effort at all.

Undemocratic decisions are another potential source for frustration. In their meta-analysis, Foels and colleagues (2000) showed that group members are more satisfied with a democratic leadership compared to autocratic leadership. Also, satisfaction was higher when group members could participate in decisions of the group.

The final interaction pattern in this category was derived from a study by Draskovitc and colleagues (2004). One aspect of their investigation concerned the relationship between dysfunctional behavior and satisfaction. They found that *interactions that are dominated by a single group member* are also associated with low levels of satisfaction (Draskovic et al., 2004). According to their conceptualization, "dominant personality" (p. 453) refers to a group member repeatedly providing (un)solicited explanations, which hinders collaborative exchange of ideas.

Coordination

The final set of interaction patterns concerns coordination during collaboration. All four interaction patterns in this category were derived from Kwon and colleagues' (2014) cluster analysis (see above) by reversing the interaction patterns found in successful groups (i.e., 'Early Active Collaborator'). Based on the characteristics of these groups, we assume that *low intensity interactions during early phases of the collaboration* and *neglecting to distribute tasks clearly* may be associated with a bad group climate or dissatisfaction.

Group members in this cluster further regularly reported their progress on the task to their group. Communicating one's process on the task helps group members develop group awareness which is necessary for coordination (Bodemer & Dehler, 2011). Our analysis of collaboration data revealed the opposite interaction pattern, that is, students felt confused if it was not clear to them if and to what degree their group members had already worked on the joint task. Therefore, we assume that *not displaying individual progress towards the joint task* may be a source of frustration.

We term the final potentially frustrating interaction pattern *deadline rush.* Again, we derived this interaction pattern from the cluster analysis by Kwon and colleagues (2014). Some students in their study complained that they had to rush a large amount of work shortly before the deadline. Arguably, this phenomenon could be the result of a number of the interaction patterns described above.

Survey to identify the most severe interaction patterns

So far, we revealed over a dozen interaction patterns that can occur during online-collaboration and could potentially lead to frustration. However, our analyses do not indicate how frequently these interaction patterns occur or which are perceived as particularly frustrating by the learners. To answer these questions, we conducted an online survey at the end of the summer term 2020 which was the first term where European Universities had to resort to remote online education due to the COVID-19 pandemic. Thus, we expected that many students would have enrolled in online courses that employed collaborative activities.

The online questionnaire presented the participants with the fourteen interaction patterns that we presented above. Each interaction pattern was presented on an individual questionnaire page. The top of the page presented participants with a brief description of the interaction pattern, followed by the items. For example, the description for *social loafing* read 'Individual group members contribute little or nothing to the joint work'. Since we expected that not all respondents would complete the questionnaire, we randomized the sequence of the interaction patterns for each respondent so we would achieve an even number of responses for each interaction pattern.

Sampling process and retained sample

We drew a convenience sample by contacting the teaching staff at our university and asking them to encourage their students to participate in the survey. We pursued two strategies to contact teachers. First, we contacted the teaching staff at our own department via email. Second, we contacted teachers from other departments by posting a blog article with the survey link on the university's teaching blog. Data collection took place from mid-July until the end of August 2020 (six weeks). In total, 128 students answered the questionnaire. However, 28 students did not continue the questionnaire beyond the socio-demographic data and were hence removed from the sample. To retain as much data for the individual interaction patterns, we included data from all participants who filled in the questions for at least one interaction pattern. The retained sample consisted of 100 participants (age: m = 23.15, SD = 3.63; 68% female students). The majority of the participants (65%) studied two social sciences in parallel (predominantly educational research plus a language) while 22% were enrolled in a STEM subject (e.g., civil engineering or applied computer science). The remaining participants studied either one social science or one social science plus a STEM subject. In summary, students reported that 52.3% of their courses included collaborative learning activities. On average, students worked in 6.03 (SD = 4.69) different groups that consisted of 4.23 (SD = 2.51) students and worked together for 23.13 (SD = 24.32) days.

Measures

After entering demographic information (age, gender) and information about their courses (*total number of courses taken, number of courses that included online collaborative learning, size of the online groups, duration of the collaborative tasks),* students replied to the items for the interaction patterns. For each interaction pattern participants indicated if the interaction pattern had *occurred* in at least one of their online groups ('yes' or 'no'). This item served as a filter for the remaining items. If a participant indicated that they had experienced this interaction pattern, they received one item asking them to rate *how frequently* this interaction pattern occurred in their groups and one item that asked them to rate the *degree of dissatisfaction* that was caused by the interaction pattern. Both items used a five-point Likert scale that ranged from 1 ('very rarely' or 'very little', respectively) to 5 ('very frequently' or 'very much', respectively). At the end of the questionnaire, participants could suggest *new interaction patterns that caused frustration* through an open-ended item.

To determine which interaction patterns were both, frequent and frustrating, we determined the *severity of the interaction patterns* by calculating a rank-sum that incorporated the frequency of the interaction pattern and the degree of dissatisfaction associated with the interaction pattern. Specifically, we calculated this variable for each interaction pattern as follows. First, we created two lists that contained all interaction patterns and ranked the interaction patterns, in the first list by the frequency of occurrence in the second list by the degree of dissatisfaction. Raking refers to ordering positions and assigning a rank number to each position in the list. In the resulting lists a position at the top of the list (i.e., indicated by a small number such as '1') indicated that the interaction pattern was the most frequent, or caused the most dissatisfaction, respectively. Afterwards we summed the two ranks for each interaction pattern. This resulted in a score that could range from '2' (i.e., the interaction pattern was on the top position in both lists) to 28 (i.e., the interaction pattern was on the bottom position in both lists). The more frequent an interaction pattern occurred and the more students perceived it as frustrating, the lower the resulting sum of ranks.

Results

Table 1 provides an overview about the interaction patterns, how frequently they occurred, how frustrating the interaction pattern was perceived by the students, and the severity of the interaction patterns. The table is sorted by severity, that is, interaction patterns at the top of the list occurred more frequently and at the same time were perceived as most frustrating. Note that the number of cases (n) reported in the table varies because not all participants filled out all items. Further, participants only answered the items for an interaction pattern if they had experienced it in at least one of their groups. Thus, only these participants were included in the analysis.

Table 1: Descriptive statistics and severity of each interaction pattern

#	Interaction pattern	n	Experienced by	Frequency M (SD)	Frustration M (SD)	Rank sum
1	Unequal participation	74	78.72%	3.72 (0.93)	3.62 (1.04)	9
2	Deadline rush	61	65.96%	3.79 (0.82)	3.46 (1.18)	10
3	Ineffective communication	60	65.93%	3.30 (0.96)	4.07 (0.78)	10
4	Social loafing	69	75.82%	3.51 (0.99)	3.97 (0.97)	10
5	Lack of communication about working times	54	67.50%	3.19 (0.97)	4.06 (0.98)	13
6	Undemocratic decisions	28	31.11%	3.54 (0.88)	3.39 (1.34)	14
7	Low intensity work during early phases	71	75.53%	3.85 (0.82)	2.80 (1.10)	14
8	Impolite communication	16	17.39%	3.06 (1.00)	4.13 (0.96)	15
9	Dominant group member	39	41.49%	3.49 (0.97)	3.21 (1.10)	17
10	Long waiting times for replies	49	51.58%	3.08 (1.08)	4.00 (0.82)	17
11	No clear distribution of tasks	33	35.48%	3.18 (0.95)	3.61 (1.20)	18
12	Nudging	42	46.67%	3.48 (1.04)	3.17 (1.51)	19
13	Not displaying individual working progress	31	32.29%	3.10 (1.04)	3.42 (1.18)	21
14	Phases of no interaction and communication	76	82.61%	3.29 (0.91)	2.78 (1.29)	23

Note. Interaction patterns are sorted by severity (rank-sum). Positions at the top of the table (lowest rank sum) indicate higher severity.

Our results showed that all of the interaction patterns occurred during online-collaboration. Mean values of above the scale's midpoint suggest that as soon as a group experienced an interaction pattern, it occurred somewhat frequently in the group. Students reported that a *slow start of the collaboration* (M = 3.85; SD = 0.82), *deadline rush* (M = 3.79; SD = 0.82), *unequal participation* (M = 3.72; SD = 0.93), *undemocratic decisions* (M = 3.54; SD = 3.39), and *social loafing* (M = 3.51; SD = 0.99) occurred most frequently in their groups. In terms of the degree of frustration, students perceived all interaction patterns as rather frustrating as indicated by mean values above the scale's midpoint. Altogether, students perceived five situations as especially frustrating, that is, the mean values are close to 4.0. These situations are *impolite communication* (M = 4.13; SD = 0.96), *ineffective communication* (M = 4.07; SD = 0.78), *lack of communication about working times* (M = 4.06.; SD = 0.98), *long waiting times for replies* (M = 4.00; SD = 0.82), and *social loafing* (M = 3.97; SD = 0.97). Interestingly, interaction patterns that occurred frequently were not necessarily perceived as frustrating and vice versa (e.g., impolite communication or low intensity work during early phases). Thus, in a next step, we combined the frequency and the degree of frustration to investigate which interaction patterns are the most severe. Based on the rank sums we conclude that *unequal participation*, *deadline rush*, *ineffective communication* and *social loafing* were both frequent and perceived as frustrating (cf. Table 1). Eventually, participants were also able to add interaction patterns that were not covered in the questionnaire. However, students did not mention specific interaction patterns, but instead reported frustrating situations that resulted from technical difficulties, a lack of competence in using communication technologies, unclear course requirements, and that the collaborative tasks were not suited for collaboration.

Discussion

The key to collaborative learning lies in productive interactions between the learners of a group. Thus, implementing collaborative activities in online courses can help mitigate negative effects that result from a lack of social interaction. At the same time, interaction between learners also represents a key challenge for effective collaboration. In this paper, we focused on interactions that are associated with frustration because frustration can be expected to lead to motivation loss and can thus reduce the participation of students. This is especially important because active participation and interaction among students are crucial for learning during and about collaboration. We developed a library with potentially frustrating interaction patterns that can occur during online-collaboration and used a survey to specify which interaction patterns might require special attention. In particular, our analyses revealed four interaction patterns that rank closely in terms of severity. 1) Unequal participation, 2) deadline rush, 3) ineffective communication and 4) social loafing occur frequently and are also perceived as frustrating by learners. Notably, the interaction patterns that we identified revolve around communication and participation. Our results suggest that learners tolerate phases with no interaction, however, if interaction happens to unequal parts or in an ineffective manner, they tend to become frustrated. We hypothesize that perceptions of fairness (e.g., collaboration norms, Karau & Williams, 1993) or the prospect to miss the joint goal may play a role for the development of dissatisfaction. The interaction patterns include two phenomena that result from a lack of interaction with certain team members and a perception of unfairness. Unequal participation appears to be a central challenge as it not only reduces motivation to participate but also decreases the performance of the group (Harding, 2018). However, there are many operationalizations of participation (cf. Hrastinski, 2008) and so far is has not been investigated which indicators students use to form their perception of unequal participation. However, studies on social loafing usually use self-reports to assess unequal participation (e.g., Peñarroja, Orengo, & Zornoza, 2017). While the results of this survey suggest that unequal participation and social loafing are indeed sources of frustration, Strauß and Rummel (2021) discuss after their experiment that there may be additional factors that affect the relationship between unequal participation and satisfaction with the collaboration.

The present study is not without limitations. First, the sample size of the survey was relatively small and predominantly contained students from the social sciences. As all survey, our study was prone to self-selection as the teachers were free to forward the questionnaire and the students participated voluntarily. Consequently, the generalizability of our results may be limited and future studies should aim for a larger, more diverse and randomized sample. Further, using a questionnaire at the end of the teaching semester reduced the reliability of the measurement because participants had to rely on their long-term memory to answer the questions instead of relying on memories about recent collaborative activities. This increases the potential influence of participants' implicit theories about collaborative learning in online-groups on our results. Against this background, we welcome studies that collect data shortly after a collaborative activity, or which employ behavioral data to capture interaction patterns. Finally, we acknowledge that focusing on frustration as an outcome is a selective approach. Therefore, we would like to encourage researchers and practitioners to add interaction patterns that affect other important variables such as learning outcomes or motivation.

In conclusion, instead of presenting the usual model of good collaboration, we accumulated suboptimal interaction patterns and identified interaction patterns that call for particular attention by researchers, practitioners and educators. To corroborate and expand our findings, future studies should employ research designs that allow to infer causality and analyze collaboration processes in greater detail. For example, experimental studies with confederate group members can create the interaction patterns in controlled settings and thus can help investigate the relationship between interaction patterns and variables such as satisfaction or motivation. Our paper highlights the role of dissatisfaction and frustration during collaboration and would like to encourage researchers to further investigate the role of these and similar affective variables. We are confident that the present study can also inform educational practice, especially for settings that rely on remote online learning. We propose that educational design should not only aim for increasing the opportunities for productive interactions but also should consider interactions that may cause frustration. Focusing on challenges for collaboration in online-collaboration offers a valuable alternative perspective because instruction or collaboration support can be directly targeted towards unfavorable processes. In this regard, educators can take precautions to mitigate interaction patterns and prepare students for web-based collaboration. This concerns not only the design of collaborative assignments or the configuration of the learning setting but also instruction, training or collaboration support for the learners. To address the four most severe interaction patterns, forming small groups and developing tasks that create social interdependence among learners can help reduce the danger of unequal participation or even social loafing (Johnson & Johnson, 2009). Further, students may benefit from support or training that helps them communicate effectively, especially in computer-mediated settings (e.g., in terms of grounding and media constraints, Clark & Brennan, 1991). Additionally, the learning environment, dedicated instructions or collaboration support could afford or facilitate monitoring the collaboration and early coordination and thus mitigate a deadline rush towards the end of the collaboration (cf. Jeong & Hmelo-Silver, 2016).

Acknowledgements
This study research was funded in parts by the German Federal Ministry of Education and Research (grant number: 16DHL1012) and in parts by the Stifterverband Digitales Lehrfellowship NRW. We want to acknowledge Nicole Krämer and Filipa Stoyanova for their help with the prior version of this study (cf. Strauß et al., 2018).

References
Aggarwal, P., & O'Brien, C. L. (2008). Social loafing on group projects: Structural antecedents and effects on student satisfaction. *Journal of Marketing Education*, *30*(3), 255–264.

Baturay, M. H. (2015). An overview of the world of MOOCs. *Procedia - Social and Behavioral Sciences*, *174*, 427–433.

Bodemer, D., & Dehler, J. (2011). Group awareness in CSCL environments. *Computers in Human Behavior*, *27*(3), 1043–1045.

Capdeferro, N., & Romero, M. (2012). Are online learners frustrated with collaborative learning experiences? *The International Review of Research in Open and Distributed Learning*, *13*(2), 26.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on socially shared cognition*. (13), 127–149.

Darnon, C., Doll, S., & Butera, F. (2007). Dealing with a disagreeing partner: Relational and epistemic conflict elaboration. *European Journal of Psychology of Education*, *22*(3), 227–242.

Dillenbourg, P., Baker, M. J., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds.), *Learning in Humans and Machine: Towards an interdisciplinary learning science* (pp. 189–211). Oxford: Elsevier.

Draskovic, I., Holdrinet, R., Bulte, J., Bolhus, S., & van Leeuwe, J. (2004). Modeling Small Group Learning. *Instructional Science*, *32*(6), 447–473.

Erdmann, J., Rummel, N., Christmann, N., Elson, M., Hecking, T., Herrmann, T., . . . Wichmann, A. (2017). Challenges in implementing small group collaboration in large online courses. In B. K. Smith, M. Borge, E. Mercier, & K. Y. Lim (Chairs), *12th International Conference on Computer Supported Collaborative Learning (CSCL).* Symposium conducted at the meeting of International Society of the Learning Sciences.

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.

Foels, R., Driskell, J. E., Mullen, B., & Salas, E. (2000). The Effects of Democratic Leadership on Group Member Satisfaction. *Small Group Research*, *31*(6), 676–701.

Goold, A., Craig, A., & Coldwell, J. (2008). The student experience of working in teams online. In *Hello! Where are you in the landscape of educational technology? Proceedings ascilite,* Melbourne.

Grice, H. P. (1975). Logic and conversation. In P. Cole (Ed.), *Speech acts* (pp. 41–58). New York: Academic Press.

Hadwin, A. F., Bakhtiar, A., & Miller, M. (2018). Challenges in online collaboration: Effects of scripting shared task perceptions. *International Journal of Computer-Supported Collaborative Learning*, *13*(3), 301–329.

Harding, L. M. (2018). Students of a feather "flocked" together: A group assignment method for reducing free-riding and improving group and individual learning outcomes. *Journal of Marketing Education*, *40*(2), 117–127.

Hrastinski, S. (2008). What is online learner participation?: A literature review. *Computers & Education*, *51*(4), 1755–1765.

Järvelä, S., Kirschner, P. A., Hadwin, A. F., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual- and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning*, *11*(3), 263–280.

Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported collaborative learning: How to Support collaborative learning? How can technologies help? *Educational Psychologist*, *51*(2), 247–265.

Jeong, H., Hmelo-Silver, C. E., & Jo, K. (2019). Ten years of computer-supported collaborative learning: A meta-analysis of CSCL in STEM education during 2005–2014. *Educational Research Review*, *28*, 100284.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, *38*(5), 365–379.

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, *65*(4), 681–706.

Kellermann, K., & Park, H. S. (2001). Situational Urgency and Conversational Retreat: When Politeness and Efficiency Matter. *Communication Research*, *28*(1), 3–47.

Khalil, H., & Ebner, M. (2014). MOOCs completion rates and possible methods to improve retention - A literature review. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 1236–1244.

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). Boston, MA: Springer US.

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.

Kwon, K., Liu, Y.-H., & Johnson, L. P. (2014). Group regulation and social-emotional interactions observed in computer supported collaborative learning: Comparison between good vs. poor collaborators. *Computers & Education*, *78*, 185–200.

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & Education*, *48*(2), 185–204.

Ortony, A., Clore, G. L., & Foss, M. A. (1987). The referential structure of the affective lexicon. *Cognitive Science*, *11*(3), 341–364.

Pappano, L. (2012, November 2). The Year of the MOOC. *The New York Times*. Retrieved from https://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html

Park, H. S. (2008). The Effects of Shared Cognition on Group Satisfaction and Performance. *Communication Research*, *35*(1), 88–108.

Peñarroja, V., Orengo, V., & Zornoza, A. (2017). Reducing perceived social loafing in virtual teams: The effect of team feedback with guided reflexivity. *Journal of Applied Social Psychology*, *47*(8), 424–435.

Reynolds, R., & Chu, S. K. (2020). Guest editorial: Introduction to the Special Issue on Emergency Remote Teaching (ERT) under COVID-19. *Information and Learning Sciences*, *121*(5/6), 233–239.

Rosé, C. P., Goldman, P., Zoltners Sherer, J., & Resnick, L. B. (2015). Supportive technologies for group discussion in MOOCs. *Current Issues in Emerging eLearning*, *2*(Issue 1, Article 5). Retrieved from http://scholarworks.umb.edu/ciee/vol2/iss1/5

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*(6), 1467–1478.

Strauß, S., & Rummel, N. (2021). Promoting regulation of equal participation in online collaboration by combining a group awareness tool and adaptive prompts. But does it even matter? *International Journal of Computer-Supported Collaborative Learning.* https://doi.org/10.1007/s11412-021-09340-y

Strauß, S., Rummel, N., Stoyanova, F., & Krämer, N. (2018). Developing a library of typical problems for collaborative learning in online courses. In J. Kay & R. Luckin (Eds.), *Rethinking Learning in the Digital Age: Making the Learning Sciences Count, 13th International Conference of the Learning Sciences (ICLS) 2018* (pp. 1045–1048). London, UK: International Society of the Learning Sciences.

Strijbos, J.-W., & Weinberger, A. (2010). Emerging and scripted roles in computer-supported collaborative learning. *Computers in Human Behavior*, *26*(4), 491–494.

Sung, E., & Mayer, R. E. (2012). Five facets of social presence in online distance education. *Computers in Human Behavior*, *28*(5), 1738–1747.

Tuckman, B. W., & Jensen, M. A. C. (1977). Stages of Small-Group Development Revisited. *Group & Organization Studies*, *2*(4), 419–427.

Walther, J. B. (1996). Computer-Mediated Communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, *55*(4), 828–846.

Wegner, D. M. (1995). A computer network model of human transactive memory. *Social Cognition*, *13*(3), 319–339.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, *46*(1), 71–95.

Wittenbaum, G. M., Vaughan, S. I., & Strasser, G. (2002). Coordination in task-performing groups. In R. S. Tindale, L. Heath, J. Edwards, E. J. Posavac, F. B. Bryant, Y. Suarez-Balcazar, . . . J. Myers (Eds.), *Theory and research on small groups* (Vol. 4, pp. 177–204). Boston, MA: Springer US.

International Society of
the Learning Sciences

Using Minecraft to Reconstruct and Roleplay Local History: Intersubjectivity, Temporality, and Tension

Anders I. Mørch, Carl Sebastian Eielsen,
anders.morch@iped.uio.no, carlse@student.uv.uio.no
Dept. of Education, University of Oslo, Norway

Louise Mifsud, Dept. of Primary and Secondary Teacher Education, Oslo Metropolitan University
louise.mifsud@oslomet.no

Abstract: This paper presents a design-based study of pupils' use of Minecraft in a whole-day school project in social studies involving three seventh-grade classes, student teachers and amateur historians. We used qualitative methods for data collection and analysis. We followed three groups through the following activities: 1) searching for historical information (introduction), 2) building in Minecraft and creating roleplay scripts (reconstruction), and 3) acting out the scripts and making videos for a class presentation (transformation). The activities combined generic and domain-specific skills practices in different ways. We analyze how these two modes intertwine and argue that the teaching model we used can bridge the gap between learning in and out school. Key concepts used in the analysis are intersubjectivity, tension, and temporality. Our findings indicate that through Minecraft pupils, teachers, and amateur historians contribute to intersubjectivity toward shared knowledge by setting and releasing tensions between generic and domain specific knowledge.

Introduction

An overall aim of our research is to bridge learning in and out of school, i.e. theoretical and practical knowledge. Resnick (1987) uses examples from mathematics to show this educational gap and argues that schools need to focus more on cross-cutting themes such as thinking and learning abilities to motivate children for school learning. Today these general thinking and learning abilities are referred to as generic, soft, or 21st-century skills, and include among others collaboration, problem solving, and creativity (Resnick, 2017; Trilling & Fadel, 2009). The study presented here addresses the educational gap by employing a popular three-dimensional (3D) virtual world in the teaching of history and a teaching model where generic skills are intertwined with subject matter knowledge. Three seventh-grade classes in social studies recreated part of a 19th-century industrial community in Minecraft (Saw Valley River) with its industrial buildings, which were workplaces for those who lived in the municipality of the school during the Industrial Revolution in the time period 1840-90.

Computer-supported collaborative learning (CSCL) shares several characteristics with generic skills. Key features of CSCL are information sharing, interaction between learners, joint meaning-making based on negotiation in the group, and developing common artifacts (Engen et al., 2018; Stahl et al., 2006). Furthermore, Stahl et al. (2006) suggest that the problem of intersubjectivity is of particular relevance for understanding how learning is produced through interaction, advocating for more in-depth interdisciplinary research and arguing that this issue has implications for research methods and for CSCL system design. In the study presented here, collaborative learning is a dynamic process of combining generic and domain-specific skills aimed at developing common artifacts from the perspective of intersubjectivity. Minecraft Education Edition (MEE) is used as a CSCL system in two respects: 1) a design environment for reconstructing historical buildings and 2) roleplaying historical events in the buildings to learn social-studies concepts pertaining to a particular period in time.

The block-building and sandbox game Minecraft serves as a domain-oriented design environment. The users interact by placing and breaking 3D building blocks. Actions in Minecraft (building and destroying) have a persistent effect, keeping areas in the state that the user leaves them and enabling the continuous development of digital artifacts. The notion of block building or a "sandbox game" is analogous to a child playing in a sandbox; the sand's affordances for design are virtually unlimited and have no instructions or objectives, but constraints can be imposed by tools, artifacts, and knowledge-based activities (Mørch & Thomassen, 2016). Furthermore, making and destroying are legitimate actions toward artifacts. The challenges and opportunities of integrating this type of learning environment in three middle-school classrooms are the focus of the paper.

Several studies of Minecraft in education have highlighted the potential of Minecraft to support creativity (Karsenti & Bugmann, 2018; Lorence, 2015). However, studies have indicated several challenges in using MEE in learning environments such as teachers' reluctance to using Minecraft due to the gap between students' and teachers' game knowledge (Kuhn & Stevens, 2017) and lack of focused learning objectives, inflexible curriculum, and no previous gaming skills (Baek et al., 2020). Callaghan (2016) argues that the pedagogical use of Minecraft promotes conditions that are favorable for learning, not only in relation to creativity but also for collaboration. Cipollone et al. (2015), for example, show that Minecraft gives players an opportunity to be creative in virtual environments that would otherwise be difficult to recreate in the real world. In studying how Minecraft might be integrated into the curricula, Baek et al. (2020) argue that by using Minecraft, students are interested and enthusiastic while acquiring curricular knowledge and skills in subjects such as science, math, social-science, and language-arts and composition classes. Detailed depictions of history in a game that models real-life historical and present conditions are an appealing alternative to static pictures and descriptions used in traditional materials. Students can navigate through the virtual game space and observe the scenes that simulate real-life situations, promoting student interest and engagement (Baek et al., 2020). Spikol and Milrad's (2008) study using mobile technologies for learning local history indicated that giving pupils the possibility to involve themselves in authentic historical settings in which to collaborate with peers gives rise to meaningful learning.

Therefore, we address the following research question:

- How are generic and domain-specific skills intertwined in pupils' use of Minecraft in a seventh-grade social-study (local-history) project?

Integrating intersubjectivity and domain-oriented design environments

We adopt a theory of experience based on temporality and emergence referred to as social consciousness (Mead, 1910). Mead argues "there is a continuity of experience, which is a continuity of presents" (Mead, 1910, p. 1). Mead's interest was in understanding the past in the present as an emergent phenomenon of social reality, during which reconstruction is a central component (Mead, 1929). The past arises in memories and is represented in visual images (Mead, 1929, p. 235). The past is not stable or fixed, according to him, because "[t]he past consists of the relations of the earlier world to an emergent affair – relations which have therefore emerged with the affair. . . . The past thus belongs to a generalized form of experience" (Mead, 1929, p. 5).

This theory inspired the model we use for teaching history with Minecraft, both in terms of levels of temporality and the use of a design environment for reconstruction. Furthermore, we conjecture that when reconstructing the past in the present, tensions (not only relations) in temporality emerge, which we use as analytic concept in the analysis of intersubjectivity. By tension we mean a conflict between two elements that must be resolved to advance development. In our case it is used to align elements of domain specific and generic skills practices and past, present, and future events. We draw on Ludvigsen et al.'s (2010) characterization of horizontal temporality (levels of change according to time scale) and vertical temporality (in-depth discursive analysis on a specific level). In our study, we include dynamic visual artifacts as a context for analyzing discursive practices.

Intersubjectivity is a type of social consciousness, which in the work of Rommetveit (1976) is depicted as an expansive process of communication in a spatial–temporal–interpersonal space. According to Rommetveit, intersubjectivity is a temporarily sustained and partially shared social world that depends on access to historical information (common pre-understanding), which is projected forward by anticipatory cues (shared prolepsis). Participants in conversation collaboratively construct knowledge by expanding intersubjectivity toward the future, the past, social relationships, and specific localities (Rommetveit, 1976). Researchers in computer-supported cooperative work (CSCW) and CSCL have adapted the framework for analyzing technology-mediated communication in distributed work (Fugelli et al., 2013; Stahl, 2016) and collaborative construction of knowledge (Stahl et al., 2006). Technology can support or hinder intersubjectivity, and Suthers (2006) suggests that CSCL systems should be designed to support communication and constrain the activities toward learning trajectories.

Domain-oriented design environments (Fischer, 1994) are digital tools to mediate two interdependent design activities, constructive design, and argumentative design. Constructive design is mainly a visual activity of combining building blocks into functional designs, whereas argumentative design is mainly a verbal activity, including the discussion of desired relations among the design units (Fischer, 1994). The two activities of domain-oriented design environments inspired the design of complementary modes of activity for the teacher and pupils to shift their focus as they engage in different learning activities by toggling between generic and domain-specific skills practices. Mørch, Mifsud & Eie (2019) have developed a teaching model to support this process. The teachers, in collaboration with the researchers, used this model to organize the classroom activities (see Table 1).

Table 1 provides steps for developing intersubjectivity in phases: from a vague object of shared understanding to one that is more complete (ending with a roleplay video). Tensions are inherent in temporal orientations (past, present, and future), in the difference of visual and verbal activities (Fischer, 1994), and in discursive practices (Ludvigsen et al., 2010). From a temporal perspective on social consciousness (Mead, 1929) and intersubjectivity (Rommetveit, 1976) set in a contemporary digital context of sandbox video games, the aim of reconstruction in our research is to use a domain-oriented design environment (Minecraft) to create the historical context for developing intersubjectivity toward shared knowledge and memorable shared experience.

Table 1: A model for teaching with Minecraft in social-studies classrooms (Mørch, Mifsud & Eie, 2019)

Phase (temporality) (*time scale: frame*)	**Skills-practice intertwining** (foreground vs. background)	**Example of tensions** (and techniques for resolving them)
Introduction (oriented toward the past) *(slow: 50 years)*	Domain-specific vs. generic (teacher-centered activity leading to an incomplete object of shared knowledge)	Historical buildings and events vs. searching for relevant information (resolved by amateur historians, online searches, and site visits)
Reconstruction (oriented toward the present) *(inter-mediate: day-hours)*	Generic vs. domain-specific (learner-centered activity leading to a fragmented object)	Minecraft building blocks vs. building architecture pictures (resolved by teacher's scaffolding and pupils' creativity)
Transformation (oriented toward the future/out of school) *(fast: minutes)*	Domain-specific vs. generic (learner-centered activity leading to a focused object; varying degrees of quality)	Enacting social concepts in roleplay vs. Minecraft stage props (resolved by personalization and humor)

Research design and methods

The pupils used Minecraft as an educational game to learn about a social-studies topic, 19th-century forestry industrialization and timber trade in eastern Norway. The topic was adapted for a one-day school project, where pre-service teachers from a nearby teacher-education college participated in the activity. Three senior citizens with in-depth knowledge of local history (industrial, architectural, and labor history) were invited by the school to give an introductory presentation on the topic, and we refer to them as amateur historians. We used aspects of design-based research (DBR) to organize the activity (Brown, 1992; Hoadley, 2002). Our intervention is based on three previous iterations in a teacher-education program using the same model to prepare the student teachers to teach seventh-grade pupils social-studies topics using a virtual world that builds on pupils' prior (out-of-school) experiences (Mørch, Mifsud & Eie, 2019), and adapted in this iteration by a new location (school rather than university), theory refinement, and scaffolding by amateur historians. The pupils' work was not assessed by grades, but was discussed in the classroom by student teachers, amateur historians, and researchers.

We collected data from three seventh-grade classes (N=80) using field notes, video observations (three groups of four pupils each), and audio-recorded interviews (12 pupils). After transcribing the data, six researchers participated in a data-analysis workshop to code the material. We used a version of thematic analysis based on abductive classification to organize the textual data (Reichertz, 2014). The model (Table 1) provides three overarching themes (introduction, reconstruction, transformation) and our conceptual framework provides additional analytic concepts (intersubjectivity, temporality, tension). Several themes emerged during data categorization and we profile the following: scaffolding, cooperation, collaboration, problem solving, creativity, humor, domain knowledge. The transcript notation we used includes these symbols: (..) short pause, ((text)) comment by researcher, [..] excluded (non-audible) speech, and :: abruption of talk. The nine extracts presented below are chosen to illustrate the different phases as well as to highlight the intertwining of generic and domain-specific practices during collaborative learning as it developed over time. The names of participants are fictitious.

Data and empirical analysis

In this section we show a series of data extracts, organized in three subsections according to the three phases, and illustrating similarities and differences of three groups' collaborative learning. We focus in-depth on one group (Group 2) in the second subsection to show how the group worked and shifted focus as the work changed over time.

Introduction

In the beginning of the assignment, the pupils were engaged in information seeking and knowledge acquisition. We present this theme (introduction) from three different perspectives in order to foreground multiple methods for information seeking. The data below (see Table 2) are from two interviews (Groups 1 & 3) and from video observation (Group 2).

Table 2: Extract 1 (Wood Factory), Extract 2 (Sawmill Factory), and Extract 3 (Steel Factory)

Group 1 (00:28:57) Interview	**Group 2** (00:10:33) Video	**Group 3** (00:00:13) Interview
1. Interviewer: How did you know it started to burn in the Wood Factory? **2. Daniel**: It said so on a website. I think it was Wikipedia. It said it started to burn, and people came here to talk about it. **3. Interviewer**: Who were those guys? **4. Anna**: It was ((researchers)). We got pictures and ((building)) measurements from them. **5. Daniel**: And we used information about what we thought had happened in the factory. **6. Anna**: And when we got the measurements, we copied them from a sheet of paper [..].	**1. Student teacher:** Do you know where the Sawmill Factory is located? Is it here in ((City A))? **2. Lisa**: It's in ((Town B)). **3. Student teacher:** In ((Town B))? **4. Lisa**: Yes. **5. Student teacher:** Because then you can say that you are in ((Town B)) and that the timber comes from the nearby forest. It is useful to include some of the historical information, such as where the logs come from and the process of cutting them into planks.	**1. Kris**: We asked the pro how it was. **2. Interviewer**: The pensioners you mean? Those who came here? You asked him questions and then you created something afterward? **3. Kris**: Yes. **4. Thea**: And then we made something up. We got this ((sheet of paper with the building's measurements)), but before we could start, we went through how it ((working life)) was at the time. It was someone who explained all these things for us ((about buildings and working life)).

Group 1 sought information about the Wood Factory and cited multiple sources, including Wikipedia and building measurements they received from one of the researchers (Extract 1). They mentioned later in the interview that building the factory in Minecraft required more domain knowledge than scripting the roleplay, which could be a reason why they gathered information from multiple sources. Group 2 could not find any historical information about their building, the Sawmill. In this group, the student teacher played a central role by suggesting they look for information about the industrialization process of making planks from logs felled in the nearby forest (Extract 2). Group 3 was the most positive toward the information provided by the pensioners (amateur historians), asking one of them questions and using the information combined with their own ideas (Extract 3). Analyzing and comparing these extracts, we see that the three groups were able to find domain knowledge and start the process of developing intersubjectivity. Group 1 focused on a fire that broke out in the Wood Factory, Group 2 on the log-cutting process, and Group 3 on the working conditions in the Steel Factory.

Reconstruction

After having acquired knowledge of the buildings and important events, the next step for the three groups was to reconstruct the buildings in Minecraft and write a script for the roleplay. In this section, we focus on the building process and compare three extracts of the same group (see Table 3), showing how the pupils gradually learned to work together with the help of the student teacher. We illustrate how Group 2 developed their MEE building (Sawmill Factory) in parallel with building their understanding and how they incorporated historical information through negotiation supported by scaffolding, cooperation, and collaboration:

Table 3: Extract 4 (Scaffolding), Extract 5 (Cooperation), and Extract 6 (Collaboration)

Group 2 (00:18:16) Scaffolding	**Group 2** (00:22:22) Cooperation	**Group 2** (00:23:22) Collab.
Jon: Should we have red on the top? **Geir**: Should we use red terracotta ((searching MEE inventory, writes "red" in the search bar for options))? **Jon**: Or red wood? **Geir**: Or red concrete? **Student teacher**: Remember how it ((the building)) looked like then	**Lisa**: But we were going to have a wooden floor, didn't we? **Gro**: Yes, I'll find it. **Lisa**: What type of wood? **Gro**: Oak wood! **Student teacher**: What if you two ((Lisa and Gro)) start to build the wooden floor, given you aren't as experienced as the other guys?	**Geir**: Maybe we should choose a red block (..)? **Jon:** I think you can do it as it is (..). If you take the windows in the middle (..) assuming we have six spaces in between. **Geir:** Six spaces? **Jon:** It's not that many windows in the picture

(..). Was it made of wood or concrete? **Geir**: I think it was made of wood. Do we have any red wood ((writes “wood” in the MEE search bar))? **Student teacher**: You can also use red wool as well if you want red (..) but does everybody ((referring to the whole group)) agree that the building should be red? **Jon**: Yes, it was ((red)) on the picture ((Geir places “red wool” in the search bar)).	Then you ((Jon and Geir)) can start on the roof? **Jon**: Where is the floor on the picture ((Jon opens OneNote and looks at photo in Fig. 1))? **Student teacher**: ((Repeats his former comment)) If you ((Jon and Geir)) could start with the roof and the windows, then the rest of the group ((Lisa and Gro)) can start with the floor, given that you guys are more experienced.	((looks at the picture of the building, Fig. 1 middle))? **Geir:** Should I take the other side ((of the building))? **Jon:** Let us see how it will look first. **Geir:** We are going for six spaces? **Jon:** Or five? **Geir:** Is it on the fifth or the sixth ((asking for a confirmation of options))?

In Extract 4, Jon and Geir discussed the options for the red building blocks to match the building (see Figure 1). The student teacher reminded them that they should refer to the material the real building consisted of. Geir believed it was made of painted red wood and asked if there were any red wooden blocks in the MEE inventory. The student teacher also suggested that they could use red wool, a versatile MEE building block, if everybody in the group agreed, implying the block’s color was more important than its functionality, to which Jon answered yes, referring to a picture they had received earlier (see Figure 1, middle). In Extract 5 (see Table 3, middle), the student teacher advised the group to split the work into subtasks, and suggested the boys take the roof and the girls the floor. The student teacher assumed that the girls had less experience playing Minecraft and that the roof was more difficult to construct. In Extract 6, Jon and Geir discussed the distance between the building’s windows to determine how many they could fit on one of the walls. The photo that the pupils used as a reference shows four windows on one side (see Figure 1, middle), but the pupils created seven (see Figure 1, right). The historical photos they received from the amateur historian did not cover the entire building, leaving the rest to the pupil's imagination, own interpretation, or their searching for additional information on their own (e.g., as a comparison, Group 1 used Google Maps to look for a current picture to see more detail).

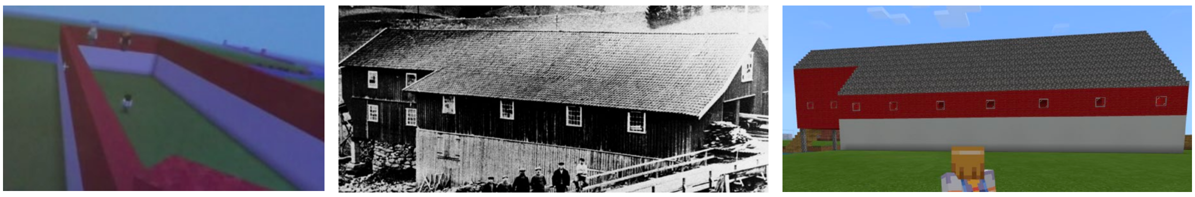


Figure 1. Two screen snapshots in the development of Group 2’s MEE building (left and right) and a photograph of the historical building that served as model (middle). They were told the color was red.

In analyzing the data material from the three groups’ reconstruction activities, we see that intersubjectivity is now knowledge-based, but remains fragmented. For example, domain-specific knowledge and scaffolding played a central role to help the pupils to cope with the challenges of relying solely on generic skills and help the group work closer together (e.g., problem solving, division of labor, collaboration). The student teacher suggested in two rounds that Group 2 should divide their work based on perceived gameplay experience. All the groups tried to create buildings that resembled the pictures they received from the amateur historians and from other sources when this was insufficient. All the groups met challenges when trying to recreate certain parts of the buildings, sometimes leading to creative workarounds, such as Group 1 creating a restaurant-like seating area outside their building, or Group 2 increasing the number of windows along a wall (Extract 6).

Transformation

In the last phase of the assignment, transformation, the pupils roleplayed historical events set to the scenery of the 19th-century industrial architecture they had created in Minecraft. We will present this theme from three different perspectives in order to foreground the degree of seriousness and domain-knowledge accuracy (or alternatively the lack of it and the inclusion of humor and entertainment) in the roleplay, as we see in Extracts 7–9 in Table 4:

Table 4: Extract 7 (Building on fire), Extract 8 (Work accident), and Extract 9 (Exploitation)

Group 1 (00:25:50–6)	**Group 2** (00:36:32–3)	**Group 3** (00:15:21–6)
Iris: What are we going to produce in the factory today? **Anna**: Probably a door. **Daniel**: OK, let's start! **Nils**: I have to use the restroom (..). I have to use the restroom (..). I have to use the restroom (..). **Anna**: Look, a fire has started to burn! **Daniel**: Let's get out of here ((everybody leaves the building))! **Anna**: We have to get out of here now! **Nils**: I have to use the restroom (..)! I have to use the restroom!	**Student teacher**: Have you started planning the roleplay yet? What are you going to say? **Gro**: Yes, we have started, but we are not yet finished. **Lisa**: ((Laughs)) It ends with Kaare dying (..). ((Starts reading the script from OneNote document.)) It starts when he arrives at work for the first time. The day after, he comes to work as usual and does not sense any danger. He goes to the cutting machine to start his work (..). Ouch (..)! What happened (..)? I cut off my hand (..) **Geir**: The story is not very long ((Lisa opens Minecraft)). **Lisa**: No, we don't have enough!	**Kris**: Albert ((manager)), we haven't received our wages in over three months ((walks straight across the MEE screen toward a table higher up)). **Mikkel**: Too bad for you guys. **Kris**: Why are you so cranky all the time? **Mikkel**: Why do you ask so many questions? **Kris**: Because I'm curious. **Mikkel**: Please leave, now ((narrator voice says, "One minute later in the roleplay"))! **Thea**: Arne ((worker)) contacted me and said you don't give these folks their wages, but you have to do that, otherwise you will be arrested.

Comparing the three roleplays, it seems the groups had different approaches and expectations with regards to domain knowledge. Group 1 used domain knowledge as a basis for their roleplay (Extract 7) with a fire as the second act of their roleplay, following a door-production scene. They also included a humorous element that was important for two of the group members. Group 2 could not find any historical information on their building, the Sawmill Factory, but they included elements of the industrialization process that started in a nearby forest. In the final act (Extract 8), one of the characters was injured in a log-cutting machine. The lack of a safe work environment was commented on by the amateur historians as an important element of that time, after the group's presentation. The third group (Extract 9) brought up a societal issue (exploitation of workers by factory leaders) illustrated by a dialog between a worker and his manager. The group was praised for its use of domain knowledge. The roleplays were created for video and the intersubjectivity was focused, but the quality of shared knowledge varied. Comparing the three groups along a scale of foregrounding entertainment and domain knowledge, Group 1 chose entertainment, Group 3 aimed for domain-knowledge accuracy, and Group 2 was in between the other two.

Discussion

In this section, we address the research question, *how are generic and domain-specific skills intertwined in pupils' use of Minecraft in a seventh-grade social-studies (local-history) project?* We discuss this from the two perspectives identified through the empirical analysis and informed by our conceptual framework: 1) Tensions in temporality and contextual reconstruction, and 2) setting and releasing tension.

Tensions in temporality and contextual reconstruction

The pupils relied on different techniques of information seeking to find information about the historical buildings, which was an essential generic skill in the beginning of the project. When the pupils discussed features of the buildings (see for example Table 1 and Extracts 2 & 3 in Table 2), we see that they try to connect the historical information of the buildings that are still in use, to the historical buildings they learned about and reconstructed in Minecraft, despite some of the buildings having a new function today (e.g. the place they know as Mall used to be Steel Factory). This approach to understanding history is in line with Mead's theory of temporality (Mead, 1929), which suggests that people reconstruct the past in the present aided by visual imagery and the images serve as a context for their understanding. In reconstructing the buildings in Minecraft, the pupils created "history in the present" (Ludvigsen et al., 2010, p. 109) and they relied on pictures from the past and present buildings.

In designing the buildings, the pupils spent considerable time finding the right building blocks (generic practice in foreground) and creating buildings that could be used in the knowledge-based roleplay (demonstrating domain knowledge). This discrepancy between generic and domain specific practices became a challenge for all the groups, which we refer to as a tension of context and understanding, or contextual reconstruction. The groups differed in how they emphasized context vs. understanding along a scale from picking good building blocks and scripts to demonstrating recognizable historical events and proper use of domain-specific concepts. For example, Group 1 demonstrates in Table 4 the concept of *door-production* (Extract 7), Group 2, *safety at work* (Extract 8), and Group 3, *exploitation* and *class struggle* (Extract 9). After each roleplay, this became a topic of discussion, when classmates and amateur historians praised each group but also gave them a constructive critique for what they could have done better (e.g., using concepts associated with more significant historic events and less humor). The amateur historians made important contributions to the activity of Group 3. Kris referred to them as 'pros' (see Extract 3 in Table 2), and Mikkel said later in the same interview "The pro taught us more about local history than the teacher." This indicates a tension between different epistemic positions, paralleling studies in community planning involving longtime residents and professionals (city planners). Resolving the tension by mutual adaptation led to new knowledge (Taylor, 2020). The lessons learned from these studies were used to design experimental teaching lessons for public schools in the same neighborhoods to teach about local history.

Setting and releasing tension

Group 2 did not find information about the Sawmill Factory online and was told by the student teacher to seek information about the wood industrialization process (see Extract 2 in Table 2). The two boys in the group had more experience in Minecraft than the girls, but both pairs needed frequent scaffolding. For example, before the two boys could decide on the better building block to use (wood or wool), the student teacher reminded them about the material the physical building was made of (see Figure 1, middle). This became a constraint for the pupils because their design would be compared and measured against this building (*setting tension*). Furthermore, the student teacher wanted to include the entire group in the deliberation process, and by intervening, he opened up a space for the two other members to join the discussion (*releasing tension*). By releasing tension, we mean that domain knowledge is put in the background and inclusion of all (a generic skill) is prioritized. Therefore, by setting and releasing tensions at appropriate times, the pupils were guided by the student teacher to resolve their discrepancies and move forward in their collaborative learning process.

In Extract 5, Group 2's division of labor separated the activities of two subgroups, partly as a result of scaffolding by the student teacher but later resolved when they completed the roleplay. The roleplay emphasized an important concern for workers, preventing accidents (Extract 7). However, as this topic did not specifically refer to a known event, it can be understood as a situation of releasing tension too soon (e.g., not spending enough time searching for reliable knowledge), which would be more appropriately handled by a knowledgeable person.

Humor was another way of releasing tension for the groups. Despite the effort of some members of Group 1 to use domain knowledge in their roleplay, the inclusion of humor in the script appeared to be important for Daniel and Nils as a form of entertainment for the whole class to enjoy (see Extract 7 in Table 4). Even though they knew humor might jeopardize the knowledge basis of their roleplay, they found it acceptable in their current setting. The roleplay created by Group 3 was dominated by domain knowledge (Extract 9, Table 4). The characters showcased a societal issue of exploitation and class struggle that may indicate a tension between the past and the future in terms of labor rights; later on (in 1920s), the workers in the Steel Factory formed a labor union.

In summary, we have used temporality and tension as analytic concepts to understand the development of intersubjectivity in three classrooms and to identify focus shifts in the pupils' collaborative learning with Minecraft in their efforts to learn about their own local history. Tensions are relations between one or more elements that can be classified as typical gaming activities and one or more elements that belong to school activities. We used these concepts to identify when the pupils switch from generic skills practice (gaming activities in the foreground) to domain knowledge practice (school activities in the foreground). Tensions are released when domain knowledge are put in the background. Frequent focus shifts stimulated the collaborative learning process, which we also supported by a teaching model. Shared knowledge was the result of the process for 2/3 of groups.

Conclusions, limitations, and directions for further work

Our research aims to contribute to bridging the educational gap of practical and theoretical knowledge in a school setting and we have used a popular digital game and a teaching model towards that end. While building in Minecraft does not help one become a better carpenter or mason, it can help one learn digital skills, history and landscape, and it can lower the threshold to theoretical knowledge. We consider our design experiment to be moderately successful based on motivation of pupils and teachers and the feedback by amateur historians.

Limitations: 1) Our qualitative approach draws on a small sample of the total population of 90 pupils with the risk of over-generalization by neglecting possible emergent phenomena and instead relying on our conceptual framework for interpretation. 2) Lack of gaming experience can prevent teachers from intervening in situations such as putting buildings on fire, explosion, invisible avatars, and fireworks. These elements can be turned off to avoid classroom disturbances, but it requires Minecraft expertise. The teacher students had learned Minecraft in a social-studies class the previous semester, but in the heat of the moment for a pre-service teacher it turned out to be a challenge. 3) The time spent obtaining reliable knowledge and ensuring accuracy in historical buildings and events, varied considerably among the groups.

The dilemma of providing relevant games vs. interesting educational tasks is not straightforward to resolve, and our tentative hypothesis is the former is easier than the latter. We have argued that finding the right balance of generic skill practice and domain knowledge as part of a dynamic process of developing intersubjectivity toward shared knowledge in parallel with building in Minecraft to gain practical experience is a step forward. Future work is a final iteration of DBR in same school to address some of the limitations.

References

Baek, Y., Min, E., & Yun, S. (2020). Mining educational implications of Minecraft. *Computers in the Schools*, *37*(1), 1-16.

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.

Callaghan, N. (2016). Investigating the role of Minecraft in educational learning environments. *Educational Media International, 53*(4), 244–260.

Craft, J. (2016). Rebuilding an empire with Minecraft: Bringing the classics into the digital space. *The Classical Journal, 111*(3), 347-364.

Engen, B. K., Giæver, T. H., & Mifsud, L. (2018). It's a fairy tale: Using tablets for creating composite texts. *Journal of Interactive Learning Research*, 29(3), 301-321.

Fischer, G. (1994). Domain-oriented design environments, *Automated Software Engineering, 1*(2), 204-213.

Fugelli, P., Lahn, L., & Mørch, A.I. (2013). Shared prolepsis and intersubjectivity in open source development: Expansive grounding in distributed work. In Proceedings CSCW 2013 (pp. 129-144). New York: ACM.

Hoadley, C. (2002). Creating context: Design-based research in creating and understanding CSCL. In Proceedings CSCL 2002 (pp. 453-462). Mahwah, NJ: Lawrence Erlbaum Associates.

Karsenti, T., & Bugmann, J. (2018). The educational impacts of Minecraft on elementary school students. In Research on e-Learning and ICT in Education (pp. 197-212). Cham, Switzerland: Springer.

Ludvigsen, S., Rasmussen, I., Krange, I., Moen, A. & Middleton, D. (2010). Intersecting trajectories of participation: Temporality and learning. In S. Ludvigsen, A. Lund, I. Rasmussen & R. Säljö (Eds.), Learning across sites: New tools, infrastructures and practices (pp. 105-121). London, UK: Routledge.

Mead, G.H. (1910). Social consciousness and the consciousness of meaning. *Psychological Bulletin, 7,* 397-405.

Mead, G.H. (1929). The nature of the past. In J. Coss (Ed.), Essays in honor of John Dewey (pp. 235-242). New York: Henry Holt & Co.

Mørch, A.I., Mifsud, L., & Eie, S. (2019). Developing a model of collaborative learning with Minecraft for social studies classrooms using role-play theory and practice. In Proceedings CSCL 2019, Vol. 1 (pp. 272-279). Lyon, France: International Society of the Learning Sciences.

Mørch, A.I. & Thomassen, I. (2016). From wooden blocks and legos to Minecraft: Designing and playing with blocks to learn in a 3D virtual world. In Proc. of the 4th CoPDA workshop: From "have to" to "want to" participate. In CEUR Workshop Proceedings, vol. 1776 (http://ceur-ws.org/Vol-1776/), pp. 61- 67.

Reichertz, J. (2014). Induction, deduction, abduction. In U. Flick (Ed.), The SAGE handbook of qualitative data analysis (pp. 123-135). London, UK: SAGE Publications.

Resnick, L.B. (1987). Learning in school and out. *Educational Researcher, 16*(9), 13-20.

Resnick, M. (2017). *Lifelong kindergarten: Cultivating creativity through projects, passions, peers, and play*. Cambridge, MA: The MIT Press.

Rommetveit R. (1976). On the architecture of intersubjectivity. In L.H. Strickland, F.E. Aboud & K.J. Gergen (Eds.), Social psychology in transition (pp. 201-214). Boston, MA: Springer.

Stahl, G. (2016). From intersubjectivity to group cognition. *Computer Supported Cooperative Work, 25*, 355-384.

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.

Suthers, D.D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *Int. J. of Computer-Supported Collaborative Learning, 1*(3), 315-337.

Taylor, K. (2020). Resuscitating (and refusing) Cartesian representations of daily life: When mobile and grid epistemologies of the city meet. *Cognition & Instruction*, 1-20.

Thorsteinsson, G., & Niculescu, A. (2016). Pedagogical insights into the use of Minecraft within educational settings. *Studies in Informatics and Control, 25*(4), 507–516.

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. San Francisco, CA: John Wiley & Sons.

Analysing Teacher Learning in Online Networks: An Outline for Methodological Decision-Making

Nick Kelly, Queensland University of Technology, nick.kelly@qut.edu.au
Jen Scott Curwood, The University of Sydney, js.curwood@sydney.edu.au
Kate Thompson, Queensland University of Technology, kate.j.thompson@qut.edu.au
Peter Reimann, The University of Sydney, peter.reimann@sydney.edu.au
Fredrik Mørk Røkenes, Norwegian University of Science and Technology, fredrik.rokenes@ntnu.no

Abstract: This paper contributes an outline to guide methodological decision making in analysing teacher learning in online networks, with a view towards research that produces actionable knowledge. It suggests that key decisions must be made in areas of research questions, units of analysis and observation, frameworks for analysis, data sources and methods of analysis, and reporting of context. It provides a concise example of decision making in the analysis of an online network of secondary design teachers in Australia.

Introduction

An *online network of teachers* has qualities of both a network—a group of teachers connected to one another and to knowledge objects through communications technology—and a community—a group of teachers with a shared identity around a common subject (Wenger, Trayner, & De Laat, 2011). Prior research has established that teachers make widespread use of online networks to access peer support and ad-hoc professional development (Lantz-Andersson, Lundin, & Selwyn, 2018; Macià & García, 2016). A range of methods have been established by researchers to understand *how teachers learn* within these networks, namely survey, interview, ethnographic observations, discourse analysis (of online interactions), social network analysis, epistemic network analysis, and process analysis (of online traces). In this paper we suggest that methodological innovation can advance this domain of research, particularly through combining quantitative and qualitative methods to move towards *actionable knowledge*—knowledge that can be used by someone in service of a desired outcome in the world—that can inform the design, implementation, and facilitation of online spaces to support teacher professional learning.

Professional learning supports teachers in adapting their pedagogy in response to the rapidly changing social, cultural, and economic environment in which they live and work. For this reason, it is essential that teachers update their skills and advance their practices in order to meet students' complex and evolving learning needs (Curwood, 2011; Desimone, Smith, & Phillips, 2013). Despite significant financial investments at local, national, and international levels, a substantial body of research challenges the effectiveness of traditional approaches to teacher *professional development*, which understands learning as a progression through stages and a series of learning opportunities that are frequently designed and administered by an outside expert (Walshe & Hirsch, 1998). It instead emphasises the importance of *professional learning*, which involves an active, collaborative, iterative, and ongoing process based on a teacher's personal interests, professional goals, and sociocultural contexts (Darling-Hammond & Sykes, 1999). As Borko (2004) argued, "To understand teacher learning, we must study it within these multiple contexts, taking into account both the individual teacher-learners and the social systems in which they are participants" (p. 4). To this end, the idea of teachers being in control of their own *professional learning network* has been theorised (Trust, 2012), and recent scholarship has situated teachers as *self-generators* of their own professional learning curriculum through social media (Prestridge, 2019).

Actionable knowledge for teacher educators, policy makers, school leaders, and practitioners is needed to improve teacher learning in this medium of online networks. Due to the lack of actionable knowledge about how to create and facilitate teacher learning within online networks, there is an urgent need to move beyond "conducting isolated studies focused on new things rather than significant problems" (Reeves & Reeves, 2015, p. 29). Prior research has described the scope, benefits, and potential of teacher online networks, yet recent scholarship has concluded that much is still not known about teachers' online learning (Lantz-Andersson et al., 2018) and "it is not clear to what extent participation [in online networks] contributed to the development of new skills or fostered teachers' reflection on their practice" (Macià & García, 2016, p. 305). This matters because there are more than 300,000 teachers in Australian schools, all of whom are required to meet and maintain national standards for the profession through post-qualification and in-service learning. At the same time, outside of school contexts, a growing number of teachers are using social media tools and accessing online resources in an effort to improve their teaching and support their students' learning.

This paper responds to Reeves and Lin's (2020) call for educational design research that addresses significant problems in education by considering how to identify units of analysis and lines of inquiry within online teacher networks and how quantitative and qualitative approaches can provide insight into the ways interactions and knowledge objects impact professional learning. It aims to: (1) explicate an outline for methodological decision-making that can inform the investigation of teacher learning in online networks; and (2) provide a concise example of how this outline can be applied through analysis of a network of teachers.

Background

Teacher professional learning

Professional learning extends beyond government-endorsed workshops or school-based initiatives to include self-directed and self-regulated activities, which are often invisible to accrediting organisations and schools. As Coburn (2001) stated, "Informal networks among teachers are largely unacknowledged by the policy world. Yet they have enormous potential to play an influential role in teacher sense-making" (p. 163). Informal networks give teachers access to new strategies, activities, and perspectives, and they provide a space to share any fears or frustrations that they may not feel comfortable expressing to immediate colleagues. Prior research has identified the core features of effective professional learning experiences for teachers: duration, content, and active learning repeatedly emerge as aspects of teacher learning that improve content knowledge, positively influence pedagogy, and promote student achievement (Desimone, 2009). Contrasted with passive forms of learning, such as attending a lecture, active learning is linked to more positive outcomes for teachers (Desimone, Porter, Garet, Yoon, & Birman, 2002). Active professional learning in online networks—where teachers have this kind of involvement—allows teachers to develop their knowledge of specific subject content and to share resources and ideas (Curwood, 2013).

Analysis of online networks of teachers

The analysis of online learning networks takes place within a complex set of nested contexts (Jones, 2015). Teachers who use networks are situated in a physical context (e.g., a school within an educational system within a nation), and the technology for both hosting (e.g., web technology) and accessing (e.g., devices) online networks frequently change, altering the affordances of online networks. Further, social norms influence the way that online networks are used and understood. Within these contexts, numerous studies of teacher learning in online networks have been conducted over the past two decades, mostly grounded in sociocultural learning theories, which have analysed teachers in a wide range of contexts, using diverse data sources, instruments, and analytical frameworks (Lantz-Andersson et al., 2018). Yet, despite this activity, the actionable knowledge that can be used by practitioners is limited. This includes the knowledge available to teachers, teacher educators, or school networks wishing to design, convene, or facilitate a network to support teachers or government bodies seeking to make decisions about deploying resources to support teacher learning in online networks.

One issue is that the diversity of analytical frameworks and variation in reporting makes it difficult to make comparisons across studies or to translate knowledge from one context to another. One, largely successful, convergence within the literature has been around the development of the Activity Centred Analysis and Design (ACAD) framework (Goodyear & Carvalho, 2013), which provides a general and widely adopted framework for the analysis and design of online learning networks (e.g., the 15 studies in Carvalho & Goodyear, 2014). The ACAD framework describes the four elements of learning networks as social design (roles and rules), set design (tools, digital, and physical learning environment), epistemic design (processes of knowledge building, tasks), and design for co-configuration (affordance for co-creation of the learning environment over time). The use of ACAD as a design framework assumes that learning (measured by outcomes or changes over time) is mediated by activity. *Activity* is the focus of the design and it can be considered to be emergent and influenced by these designable elements (Alhadad & Thompson, 2017).

Methodologically, online networks of teachers provide a great deal of latitude for researchers, as shown by recent review papers (Lantz-Andersson, et al., 2018; Macià & García, 2016). The fact that the activity in most online networks *leaves traces* makes it possible for researchers to easily gather data from conversations (e.g., teacher-teacher relationships) and activities (e.g., teacher-knowledge interactions), which have been effectively analysed using methods such as thematic analysis, discourse analysis (automated or manual), social network analysis, and process mining. Additionally, the *individuals* within a network are often involved in research, providing their perceptions and stories of lived experience, often through interviews or surveys. The aim of this paper is to move towards research that has alignment between these different elements: towards research that aims at actionable knowledge, through alignment of questions, methods, analysis, and reporting.

Methodological decision-making in studying teachers in online networks

Our framework for methodological decision-making seeks to aid researchers in designing and implementing studies. It suggests five points at which methodological questions ought to be asked, with examples for each, as relating to: (1) actionable knowledge; (2) units of analysis and observation; (3) analytical framework; (4) breadth, depth, and thickness of data sources and methods for analysing them; and (5) reporting and inclusion of context. This framework builds upon existing understandings of educational research design, where points (2), (3), and (4) are already widely recommended for most studies (e.g., Fraenkel, Wallen, & Hyun, 1993). The contribution of the framework is to guide researchers in applying these notions to the analysis of teachers in online networks. Critically, it promotes the importance of considering the actionable knowledge arising from the work and, relatedly, how this this knowledge will be communicated, through reporting that includes contextual information. The paper proceeds by describing these five points, then providing an example of the use of the framework for studying teachers in an online network. We recognise that research does not proceed in a linear fashion; it is typically cyclical and unpredictable. As such, this framework aims to elicit appropriate questions rather than to be prescriptive.

Firstly, how might the research lead to *actionable knowledge*? The potential for actionable knowledge needs to be considered any time that research questions are being formulated. The framing of the question as "how might we" is a tool to provoke thinking. While not all research needs to be directly relatable to actionable knowledge (i.e., there is a need for blue sky research too), there is a recognised need within teacher education research for more knowledge that is useful in addressing real problems (Reeves & Reeves, 2015). One way to approach this is to ask: who might this research be used by, and for what purpose? In studies of teachers in online networks knowledge tends to be actioned by teachers (e.g., guidance in developing professional learning networks), teacher educators (e.g., how to run support teachers' professional learning in online networks), and policymakers (e.g., data about the value of convening and facilitating online networks in relation to other models for professional learning).

Secondly, what are the *units of analysis and observation*? These questions are often more challenging in studying online networks than in other domains. Online networks of teachers (e.g., in Facebook) often take the appearance of *groups* within a *platform*, while a real-world clique (e.g., "English teachers in the state of Wisconsin") might be spread across many such groups and many such platforms. This can in turn be contrasted with an individual teacher who may be a member of many groups, many platforms, and many real-world cliques of different kinds. Given such confusion, there is a need for clarity around how units of both observation and analysis relate to actionable knowledge. One possibility is to have *teachers* as the unit of observation (what teachers are saying or doing) in order to make claims about the *network* as the unit of analysis ("what this network—and others like it—are good for", e.g., Carpenter & Krutka, 2015 in surveying teachers about perceptions of Twitter). Another is to have teachers as both the unit of observation and analysis ("what teachers do in online networks and what they learn from them"; e.g., Kelly & Antonio, 2016 in looking at teacher posts in Facebook to make claims about types of peer support that are valued). A third is for the network to be the unit of both observation and analysis (e.g., Macià & García, 2018 in comparing and analysing network topologies).

Thirdly, what *analytic framework* is being adopted? There has been a broad theoretical convergence in understanding teacher learning in online networks through the paradigm of sociocultural learning, and of framing analysis through notions of communities of practice, communities of inquiry, and learning networks (Lantz-Andersson et al., 2018; Kelly, 2019). The adoption of an existing framework for analysis—for example, the ACAD framework—allows a researcher to connect learning theory to research aims without needing to construct this bridge themself. Researchers may adopt elements of different frameworks—for example, a construct that has been used widely in studying is the *different forms of participation* in online networks (e.g., observers, peripheral contributors, and active members; Lave & Wenger, 1991). The analytical framework needs to match with the unit of analysis.

Fourthly, what are the *data sources and methods for analysing them*? This is a broad question that, as suggested in the Introduction, is aimed at suggesting combinations of research methods. Three ways of considering data sources are in relation to the breadth of data (e.g., big data is very broad, as determined by its velocity, variety, and volume; Sagiroglu & Sinanc, 2013), the thickness of data (the qualitative understanding that it is capable of providing), and the depth of data (how deep data sources go in demonstrating impact). For example, Homan (2014) conducts an ethnographic study of a single teacher using multiple online networks and its impact upon the teachers' teaching—this can be described as data that is thick and deep, but not broad. In contrast, Ranieri, Manca, and Fini (2012) surveyed 1107 teachers about their perceptions; their data has more breadth, but less thickness or depth.

A secondary question is to ask: How might methods be combined to broaden, thicken, or deepen data? An example can be seen in the study conducted by Lundin et al. (2018), who used computational techniques over a large corpus of data (Facebook posts) to identify 79 discussion threads, which were then subjected to a detailed qualitative analysis. This complementary combination of techniques used *automated analysis for inclusion of broad data* and then a subsequent phase of *qualitative analysis to thicken a target part of that data*. Strategies that have been used previously in studies of teachers in online networks are online surveys of participants and trace data from networks for broad data; combined with interviews and/or manual qualitative analysis of online discourse for thickening that data. The data sources and methods used within the research need to fit with research questions (and aims for actionable knowledge), units of observation and analysis, and the framework.

Finally, what *network context* needs to be included in reporting? When understanding the context of an online network the aim is to locate it *in relation* to other such networks. Given the large populations involved (e.g., teachers within a nation), changing technologies (e.g., social media platforms), and the variability in teacher groups within social networks, it is unlikely that any single study can make generalisable conclusions about 'how teachers learn within online networks'. It is thus important to have clearly established language and variables that allow for cross-study comparison. This is needed to permit convergent validity over time, through comparison of diverse studies by different researchers. Suggested basic information that should be provided about any network being included in a study is (following Kelly, 2019): the size of the network (number of members); the focus of the network if present (e.g., "History teachers sharing lesson ideas"); the regionality of the network if applicable (e.g., international, national, state, local, school); the privacy of the network (e.g., private/public); and the anonymity of participants (use of real names enforced/facilitated in any way?). Further qualitative information about the network context add to such data, and such contextual details permit findings to be more broadly useful in future studies (or meta-analysis).

As a part of a broader programme of research, this framework was implemented in our ongoing analysis of the online teacher network "Design Teachers Queensland" (DTQ; www.designteachersqld.org). Below, we describe the context for this network and our implementation of this decision-making process with partial results. The example is useful to realise why these five points of decision-making are rarely clear-cut, linear 'decisions' but rather the design of a 'best possible result' through trade-offs between competing priorities given limited resources.

Demonstrating the framework with Design Teachers Queensland network

Research context

A new syllabus for a senior high school subject "Design" was written in 2017 and released in 2018, to be taught for the first time in 2019, in the state of Queensland, Australia. This involved a range of teachers from varied backgrounds learning new content and new skills. Arts teachers, manual arts teachers, graphics teachers, home economics teachers, and technology teachers were all going to be teaching Design due to the absence of any existing specialisation in design theory, with the exception of a few teachers with a design background (e.g., as a trained architect).

DTQ is an online network convened in early 2018 to fill the need for a way to support teachers who now had to understand and implement this new Design syllabus. It was also an opportunity to put into practice the design principles and technologies developed in a prior project, TeachConnect, that aimed at exploring how to support teachers by designing and facilitating online networks (Kelly et al., 2018), which in turn built on prior theory such as Lave and Wenger (1991). As a result, DTQ was designed and facilitated according to well-established design principles in the literature (as reviewed by Kelly et al., 2018). These principles can be listed as (1) work with existing teacher communities, in this case, a range of teacher associations and official bodies; (2) have a clear focus for the network, in this case supporting the Senior Design syllabus; (3) ensure that the technology is comparable in terms of speed and feel with "best of breed" technologies, in this case through styling and implementation of the Discourse open source platform (www.discourse.org); (4) have a low threshold for participation, in that users could easily lurk and watch the community grow with minimum effort; (5) start with a core group of 10-100 users and grow from there, which was done with a two-day workshop with 80 design teachers to kick-off the online network; and (6) use a fractal design to allow for different levels of participation.

Actionable knowledge

The research question posed was: *To what extent (if any) is professional learning of teachers within DTQ changing classroom practice, in the context of the Queensland Senior Design Syllabus*? This focus upon the impact of a curated and researcher-facilitated online network is likely to lead to a kind of actionable knowledge: currently, it is difficult to make an evidence-based argument for the creation of such platforms. Because there are 'free' platforms for teacher networks available (e.g., Facebook, Instagram) that are widely used, there is a reluctance to access privately maintained networks. Addressing this question may be useful for policymakers, teacher organisations, and school leaders to make a case for (or against) convening similar networks in future scenarios.

Units of analysis and observation

The unit of analysis in this study is the DTQ network, to understand the impact of networks of this type and their potential use in different contexts. The unit of observation is the individual teachers involved within the network. Another considered option was to focus upon particular *knowledge objects* (topic threads) within the network as units of observation and attempt to follow the impact of these objects upon different teachers—we decided upon the former.

Analytical frameworks

In analysing DTQ, we used two complementary frameworks for analysis. The ACAD framework was used to consider relationships between the design of the network and the activity taking place within it. This fits well with the framework for "Promoting and Assessing Value Creation in Communities and Networks" (Wenger et al., 2011) which draws links between activities within a network (the "ground narrative") with the value represented by those activities (the "aspirational narrative"). Figure 1, which represents the analytical framework used in this study. Each framework has been described by its authors with valuable lists of *questions* to guide investigation, as well as *examples* of prior analysis; these will not be discussed in the interests of parsimony.

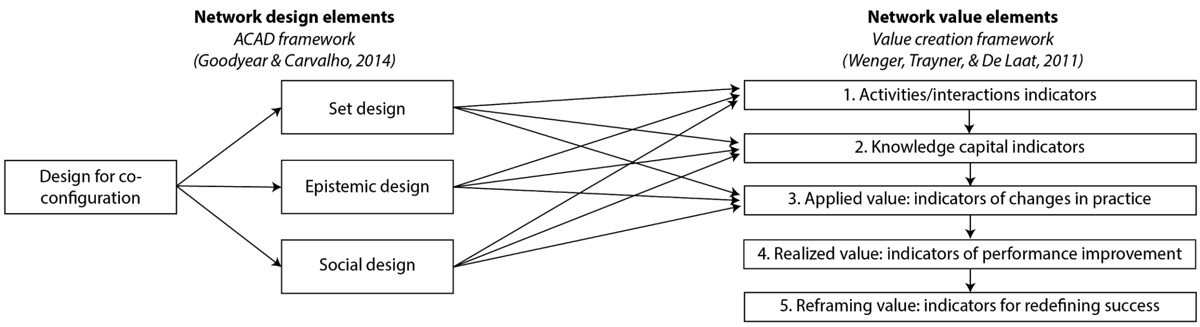


Figure 1. Analytic framework for the study, as synthesis of ACAD and value creation frameworks

Establishing network context

Establishing network context is a part of decision making for reporting—normally a fifth step. Here we have included it prior to the data sources to aid the reader. DTQ is a stand-alone, online network with its own domain name and styling (as opposed to being one part of a larger network). It is a partially private network of teachers. It was initially entirely public, but due to teacher concerns about students viewing teacher content, it was made member-only. To avoid the need to moderate all members at the time of application there are two circles of privacy: anybody can sign up and access the network, but a "DTQ Confidential" section within the network can only be accessed once approval has been granted. This area is used by teachers who wish to share discussions of sensitive content (e.g., examinations or student work). The network is targeted at teachers in Queensland, Australia (regionality) who are teaching into the Senior Design Syllabus (focus). Participants are encouraged to use their real-world identity, but anonymous presence is permitted if teachers choose to use it. Teachers are encouraged to share their details (type of teacher, subject area, career stage, location) at the time of sign up.

The platform has a modern—though minimal—styling, with a clear network identity. It is more text-heavy than many commercial social network platforms used by teachers (e.g., Instagram or Facebook). The network aims to be inclusive, firstly through intuitive design by implementing established social network norms (i.e., those used by Facebook) and secondly through semantic web design that is mostly accessible to visually impaired users. The size and usage of the network are established through two sets of data, the descriptive statistics for the life of the network, and those same statistics over a one-year period (a standardised window of time to allow for comparison between sites), the two Value columns in Table 1. Based on these values, the network is considered to be *large* in terms of teacher networks, but not massive (Kelly & Antonio, 2016)

Table 1: DTQ platform levels of interaction and activity (for one year and for the life of the platform)

Measure	*Value (Oct 2019-Oct 2020)*	*Value (Aug 2017-Oct 2020)*
Topics	81	260
Posts	434	1,300
Avg. page views per month	5,625	5,895
Daily average users / Monthly average users	12%	15%
Registered users	248	527
Registered users – participating	201	307
Registered users – non-participating	47	220

Data sources and methods of analysis

The focus upon the *impact* of the network led us to consider that—in the trade-off of breadth, thickness, and depth—the focus needed to be on depth and thickness. This depth will be achieved through a research design focused on *teacher interviews* with reflections upon lesson plans used and the role of the DTQ in preparing them. The challenge with such research—given limited resources and a large population—is that breadth of data becomes infeasible. Here we turned to *complementary combinations of methods* to leverage the benefits of each. The generalisability of the research comes in three forms: (1) statistical claims, based upon the sample size and methods for targeting individuals within that sample; (2) generalisation at a higher level of abstraction based upon theory; and (3) as one case that may come to be generalised through subsequent case comparisons and meta-analysis. The study design was guided by the analytical framework, Figure 2.

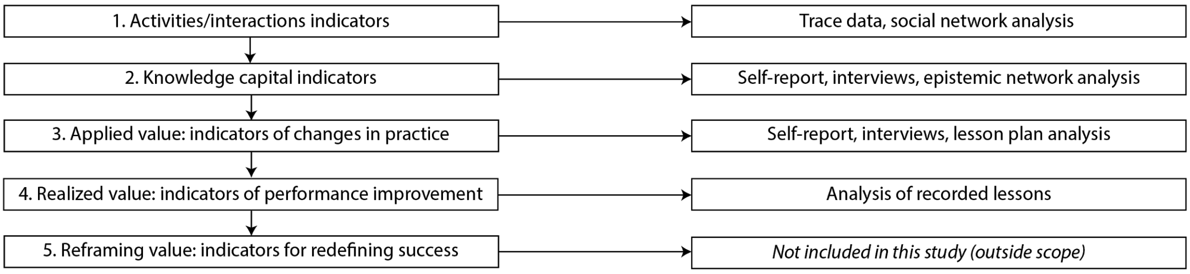


Figure 2. Links between analytical framework (value of network) and data sources

In an example of the first stage within this research, we stratified the network into three groups, based upon Lave and Wenger's (1991) description of legitimate peripheral participation: observers, active participants, and champions. Users were ranked for activity in the network in three different ways: (1) using social network analysis (analysis of a graph constructed using nodes as the individual teachers and evidence of their interactions as edges drawn between them) through the measure of betweenness centrality (a higher score indicates that a user is a "bridge" between other users in the network); (2) the quantity of their posts (raw number of posts by each teacher); and (3) the quality of their posts (calculated using weighted measures of the reply count, the number of likes, the number of links to it, the number of times it's been bookmarked, and the number of reads). This allowed for the identification of six teachers who were 'champions' within the network, who led the way in terms of active participation.

Qualitative research with these six teachers, and similar groups for active and observing users, will make up the next phase of research (yet to be conducted), addressing the depth and thickness of the data, focusing on establishing a link between activity in the platform and changes to teaching. Teachers will share lesson plans as a part of their reflection in an effort to move beyond reliance upon self-reports by teachers. We will then be able to explore links between lesson plans and teacher activity in the network based upon traces of teacher activity.

Reporting

In addition to the details provided in Network Context (which allow for comparison with other studies) the reporting will make links back to the analytical framework. In this case, the ACAD framework provides a basis for theorising a relationship between *design elements* and *observed value*. In addressing the research question this may allow for findings that, say, certain epistemic design features—for example, the way that the platform promotes the use of knowledge objects as a basis for ad-hoc discussion groups—contribute to the changes observed in teacher practices.

Discussion and conclusions

The methodological framework we propose is useful for developing the actionable knowledge that teachers, researchers, policymakers, and practitioners can use in designing, convening, and facilitating online communities of teachers. With the rapid shift to online teaching around the world due to the COVID-19 pandemic, many teachers were unable to access high-quality, face-to-face professional learning, let alone readily engage in informal dialogue with colleagues about pedagogy and practice. Consequently, there is an urgent need for researchers to consider how our methods can provide insight into how teacher learning is happening within interactions and through the exchange of knowledge objects in ever-evolving online networks.

This paper outlines a guide to methodological decision-making in analysing teachers in online networks, and it provides an illustrative example in the research planning for a study of the DTQ network. It argues, firstly, that actionable knowledge is lacking in this domain of research; and suggests that research questions might align with this goal of producing actionable knowledge. Secondly, it suggests that critical points of decision-making can be useful in addressing research questions. The domain of studying is maturing, with valuable analytical frameworks—such as the two used in this study—available to researchers, and the methods for analysing online networks are now well established. Further development of theory and to cases that can be more readily compared through combinations of methods, in research designs that are aimed at actionable knowledge, and through conventions in reporting this maturation might continue.

References

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15.

Carolan, B. V. (2013). Social network analysis and education: Theory, methods & applications. Thousand Oaks, CA: Sage.

Carpenter, J. P., & Krutka, D. G. (2015). Engagement through microblogging: Educator professional development via Twitter. *Professional Development in Education*, 41(4), 707-728.

Carvalho, L., & Goodyear, P. (2014). *The Architecture of Productive Learning Networks.* Routledge.

Coburn, C.E. (2001). Collective sense-making about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis, 23*(2), 145–170.

Curwood, J.S. (2011). Teachers as learners: What makes technology-focused professional development effective? *English in Australia, 46*(3), 68-75.

Curwood, J.S. (2013). Applying the design framework to technology professional development. *Journal of Digital Learning in Teacher Education, 29*(3), 90-97.

Darling-Hammond, L. & Sykes, G. (Eds.). (1999). *Teaching as the learning profession: Handbook of policy and practice*. San Francisco: Jossey-Bass.

Desimone, L.M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181-199.

Desimone, L.M., Porter, A.C., Garet, M., Yoon, K.S., & Birman, B. (2002). Does professional development change teachers' instruction? Results from a three-year study. *Educational Evaluation and Policy Analysis, 24*(2), 81-112.

Desimone, L., Smith, T., & Phillips, K. (2013). Linking student achievement growth to professional development participation and changes in instruction: A longitudinal study of elementary students and teachers in Title I schools. *Teachers College Record, 115*(5), 1-46.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (1993). *How to Design and Evaluate Research in Education*. New York: McGraw Hill.

Goodyear, P., & Carvalho, L. (2013). The analysis of complex learning environments. In H. Beetham & R. Sharpe (Eds.), *Rethinking pedagogy for a digital age: Designing and delivering e-learning* (pp. 49-63). New York:Routledge.

Homan, E. C. (2014). The shifting spaces of teacher relationships: Complementary methods in examinations of teachers' digital practices. Journal of Technology and Teacher Education, 22(3), 311-331.

Jones, C. (2015). *Networked learning: An educational paradigm for the age of digital networks*. London: Springer.

Kelly, N. (2019). Online networks in teacher education. In T. Allen (Ed.), *Oxford Research Encyclopedia of Education.* Oxford: Oxford University Press.

Kelly, N., & Antonio, A. (2016). Teacher peer support in social network sites. *Teaching and Teacher Education, 56*, 138-149.

Kelly, N., Doyle, J., & Parker, M. (2020). Methods for assessing higher education research team collaboration: comparing research outputs and participant perceptions across four collaborative research teams. *Higher Education Research & Development*, 39(2), 215-229.

Kelly, N., Russell, N., Kickbusch, S., Barros, A., Dawes, L. A., & Rasmussen, R. (2018). Online communities of teachers to support situational knowledge: A design-based study. *Australasian Journal of Educational Technology, 34*(5). doi:10.14742/ajet.3867

Lantz-Andersson, A., Lundin, M., & Selwyn, N. (2018). Twenty years of online teacher communities: A systematic review of formally-organized and informally-developed professional learning groups. *Teaching and Teacher Education*, 75, 302-315.

Laumann, E. O., Galaskiewicz, J., & Marsden, P. V. (1978). Community structure as interorganizational linkages. *Annual Review of Sociology*, 4(1), 455-484.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Little, J. W. (2002). Locating learning in teachers' communities of practice: Opening up problems of analysis in records of everyday work. *Teaching and Teacher Education, 18*(8), 917–946.

Lundin, M., Lantz-Andersson, A., & Hillman, T. (2017). Teachers' reshaping of professional identity in a thematic FB-group. *Open and Interdisciplinary Journal of Technology, Culture and Education*, 12(2), 12-29.

Macià, M., & García, I. (2016). Informal online communities and networks as a source of teacher professional development: A review. *Teaching and Teacher Education*, 55, 291-307.

Macià, M., & García, I. (2018). Professional development of teachers acting as bridges in online social networks. *Research in Learning Technology*, 26.

Prestridge, S. (2019). Categorising teachers' use of social media for their professional learning: A self-generating professional learning paradigm. *Computers & Education, 129*, 143-158.

Ranieri, M., Manca, S., & Fini, A. (2012). Why (and how) do teachers engage in social networks? An exploratory study of professional use of Facebook and its implications for lifelong learning. *British Journal of Educational Technology*, 43(5), 754-769.

Reeves, T.C., & Lin, L. (2020). The research we have is not the research we need. *Educational Technology Research and Development, 68*, 1991-2001.

Reeves, T.C., & Reeves, P.M. (2015). Educational technology research in a VUCA world. *Educational Technology, 55*(2), 26-30.

Sagiroglu, S. and Sinanc, D. (2013) Big data: A review. International Conference on Collaboration Technologies and Systems (CTS), San Diego, 20-24 May 2013, 42-47.

Sherin, B. (2013). A computational study of commonsense science: An exploration in the automated analysis of clinical interview data. *Journal of the Learning Sciences, 22*(4), 600-638.

Trust, T. (2012). Professional learning networks designed for teacher learning. *Journal of Digital Learning in Teacher Education, 28*(4), 133-138.

Walshe, J. & Hirsch, D. (1998). *Staying ahead: In-service training and teacher professional development*. Paris: OECD.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge University Press.

Wenger, E., Trayner, B., & De Laat, M. (2011). Promoting and assessing value creation in communities and networks: A conceptual framework. *Rapport 18*, Ruud de Moor Centrum, Open Universiteit, The Netherlands

International Society of
the Learning Sciences

Towards Asynchronous Data Science Invention Activities at Scale

Rafael Shalala, Ofra Amir, Ido Roll
rafael.s@campus.technion.ac.il, oamir@technion.ac.il, roll@technion.ac.il
Technion, Israel Institute of Technology

Abstract: Invention activities are carefully designed problem-solving tasks in which learners are asked to invent solutions to unfamiliar problems prior to being taught the canonical solutions. Invention activities are typically used in the classroom setting. As online education becomes increasingly common, there is a need to adapt Invention activities to the asynchronous nature of many courses. We do so in the context of an introductory undergraduate data science course. Using an online programming environment, students work on the tasks in pairs, without instructor support. We analyze the invention process and outcomes from two Invention activities on the challenging topics of classification and clustering. Detailed analysis of recordings of six student pairs shows how activity design supports insights at three levels: nature of models (e.g., the need to normalize); domain concepts (e.g., types of errors), and procedural solutions (e.g., weighting errors). We describe the activities, their design, and their outcomes.

Introduction

In Data Science, methods and procedures are defined and implemented in order to extract information and knowledge from datasets. As students often have very little relevant prior knowledge and experiences in these areas, teaching these methods is challenging (Berman et al., 2018). Data science literacy requires knowledge of statistics, understanding of data, and often fluency in programming. Thus, while students often follow the given procedures, they fail to acquire meaningful understanding of relevant concepts.

To address this challenge, we evaluate the benefits of introducing Invention activities to an introductory data science course. In Invention activities, students are asked to develop naive methods to solve problems prior to being taught an expert solution (Loibl, Roll, & Rummel, 2017; Schwartz & Martin, 2004). Such activities help students acquire meaningful experiences, on which future instruction builds (Schwartz, Sears, & Chang, 2012). Invention activities and other similar approaches were shown to improve students' understanding and provide strong foundations for future learning, mainly in the domain of statistics (Holmes, Day, Park, Bonn, & Roll, 2014; Kapur & Bielaczyc, 2012; Loibl et al., 2017). However, it is unclear whether such an approach would also be effective for learning more complex data-science concepts, especially when requiring programming.

A second challenge that we address in this work is the facilitation of Invention activities asynchronously, without instructor support. As online education becomes increasingly prevalent, in both informal (such as MOOCs) and university settings, there is a growing need to support meaningful, active learning in this context (Hew, 2016; Roll, Russell, & Gašević, 2018). To this end, we design activities that are facilitated remotely and asynchronously, via Zoom, without teaching staff support.

We present a case study of designing and deploying remote, collaborative, Invention activities that engage students in problem-solving tasks prior to instruction. We focus on the invention process itself and its outcomes, and discuss lessons learned and implications for the design of asynchronous Invention activities at scale.

Background

In traditional forms of science and math instruction, teachers explain core concepts, and then ask students to apply them in practice problems. Problem-solving followed by instruction (PS-I) flips the traditional approach by first engaging learners in problem solving before the teacher explains the related concepts (Loibl et al., 2017).

Invention activities are a class of the PS-I approach. These are carefully designed problem-solving tasks (Schwartz & Martin, 2004) in which learners are asked to invent general solutions for the given problems. This process helps learners acquire an intuitive understanding of the main domain concepts prior to being taught expert solutions through instruction (Loibl et al., 2017). It is done through the use of contrasting cases which highlight specific features of the domain (Schwartz, Chase, Oppezzo, & Chin, 2011). Invention-based approaches have been shown to boost conceptual learning and transfer to novel situations (Kapur, 2016; Schwartz & Bransford, 1998; Schwartz et al., 2011; Schwartz & Martin, 2004).

Building on these successes, two main challenges motivate the current work. First, the effectiveness of PS-I approaches depends on the type of knowledge being taught (Chase & Klahr, 2017), and its applicability to data science education has yet to be evaluated. Data science is intrinsically complex, as it combines statistics, programming, and big data. Each of these topics is new and challenging for students (Berman et al., 2018). Thus, there is a concern that their combination is too cognitively demanding for engaging in a productive invention process. Second, Invention activities are typically used in classroom settings. Thus, the teacher is often available to support students in their learning (Kapur & Bielaczyc, 2012). Furthermore, students are likely to stay on task even when facing challenges. However, given the current global pandemic, and to support adoption at-scale, we sought to implement Invention activities asynchronously, as homework assignments, without teacher support.

We designed two Invention activities in which students were asked to invent and implement quantitative methods for evaluating the quality of classifiers and for evaluating the quality of clustering methods. We collected recordings of several students who worked on these activities in pairs, and analyzed them. Using this data, we tried to answer our main research question: how to design asynchronous Invention activities to support data science learning for undergraduate level students?

The main contributions of this work are twofold: (1) providing a design approach and rationale for asynchronous Invention activities that could support their adoption at scale, and (2) demonstrating the efficacy of Invention activities for data science by mapping the outcomes of students' invention process to design features of the activities.

Method

To better understand the outcomes of the invention process, and how these were afforded by our design choices, we focus on analyzing students' invention processes and outcomes while working on the activities.

Procedure and participants

We ran four Invention activities that were followed by lectures as part of an undergraduate level introductory data science course. The first two activities served as a pilot. The latter two activities covered the topics of classification assessment and clustering assessment. They were written using the Jupyter notebook (Perkel, 2018) web application and used Python as the programming language (Python was used for all programming activities in the course). The Jupyter notebook web application allows users to create and edit documents that contain code, text and visualizations. Students worked on the Invention activities in pairs, at their own time, and from their homes. Students were asked to submit their Jupyter notebooks, including their solutions, one day prior to the lecture. Students received the assignments about a week prior to the lecture and could choose when and for how long to work on the activities.

The lecture instruction began with an overview of the students' solutions, followed by teaching of the expert solution or solutions. The overview of the students' solutions included discussions on the differences and trade-offs between them.

All students in the course were asked to complete the activities and were invited to participate in the study. Those who consented were asked to record themselves while working on the activity (while sharing the screen where they edit their code) and share it with the study team, and were given a compensation of $15. Activities took on average 70 minutes (min: 50, max:87). Six student pairs participated in the study (6 males, 6 females). Two pairs participated in both activities. In total, four pairs participated in each activity. Participants had no prior experience with this teaching approach.

Materials

The activities were delivered using code and text embedded in Jupyter notebooks. Each Invention activity included five consecutive tasks:

1. *Introduction* – Students were given a context story. For example, a story about the need for classification of COVID-19 at-risk population according to their medical information, and the goal of a company to develop such a classifier.
2. *Contrasting cases* - Students were presented with two cases that supported intuitive comparisons. Students were asked to choose between these cases and explain their choice. For example, choosing between two classifiers according to their classification results.
3. *Invent a numeric measure* - Students were asked to create a numeric measure for the presented problem. For example, "Suggest a numeric measure to estimate the quality of a classifier, higher value indicates a better classifier".
4. *Implement the suggested measure* - Students were asked to implement their suggested measure. For example, "Implement your suggested measure by completing the following methods that get as input the classifier results and the real data".

5. *Test and reflect* - Students were instructed to test the measure using the examples given in the contrasting cases and reflect upon the outcomes. For example, "Use your suggested measure to examine the classifiers presented in task 2. Do the results support your choice?".

Classification assessment invention activity

The main goal of this activity was to deliver two core concepts in classification: (1) accuracy score is not sufficient for evaluating a classifier and might be misleading, and (2) recognizing the significance of the different types of classification mistakes, namely false positive (wrongly classifying a negative case as positive) and false negative (wrongly classifying a positive case as negative). The introduction described the need for a classifier to identify COVID-19 at-risk populations based on medical data (see Figure 1). The goal of this story was to get students to think about the quality of classifiers in the context of a real-world example. Next, students were asked to choose between two classifiers (contrasting cases) that were tested on data of 40 people, of which only two people were at-risk. One classifier was more accurate, but failed to classify the at-risk people correctly, while the other, though less accurate overall, succeeded in classifying one of the at-risk people correctly. These contrasting cases aimed to highlight the tension between a classifier's accuracy and its ability to avoid critical mistakes.

The next task was to suggest a measure for the quality of classifiers and implement it. Students were provided with code that computes the basic accuracy score of a classifier, i.e., the percent of instances that were correctly classified. They were asked to fill in new methods that propose other measures for the classifiers' quality. The purpose of providing code for basic accuracy was twofold: first, driving students to think of alternative, more elaborate, solutions. Second, basic code that could be edited, scaffolded the process and reduced the risk of time-consuming programming bugs, allowing students to focus on the conceptual challenge of the activity. We further provided students with the name of the method ("classifier_measure1") and its signature which specified the input for the method - two arrays, one for the predictions made by the classifier and one for the ground-truth classification of the test instances. Finally, students were asked to test their implemented measure on the classifiers and reflect upon the choice they have made when choosing their preferred classifier, as well as on the measure they invented.

Due to COVID-19 there is a high demand for identifying in-risk population to provide necessary aid. For this purpose, "Meditest" company recruited two teams to build classifiers for in-risk population. Each of these teams suggeted a classifier that was trained on data gathered from thousands of people including medical information such as pre-existing condition, blood pressure and pulse. For each given person, the classifier predicts whether he belongs to in-risk population or not.

To test the classifiers and choose the better one, "Meditest" company assembled real classification data on a group of 40 people in which it is known if they belong to in-risk population or not. This group of people was not included in the training data.

The real data is presented below: 0 - not belongs to in-risk population, 1 - belongs to in-risk population

```
import numpy as np
np.set_printoptions(linewidth=100) # for printing each array in a single line
```

```
# The real_data numpy array contain 40 items representing the real data gathered by "Meditest"
# each item in the numpy array represents the true and reliable classification of a person.
# 0 - does not belong to at risk population, 1 - belongs to at risk population
# for example, the person in index 0 is not at risk, while the last one (index 39) is at risk

real_data     = np.array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
                          0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1])
```

Figure 1. The Classification Assessment activity introduction story

Clustering assessment invention activity

The main goal of this activity was to help students develop an intuition for how to assess a clustering method and give an example of the utility of clustering. The introductory story described an attempt to help students choose academic courses by presenting information about the interest level and difficulty of the courses, and the intent of the students to divide the data into three groups. Next, students were asked to choose between two clustering methods (Figure 2). The clustering on the left provides better separation between the groups, as there is less overlap between groups B and C. Similar to the Classification Assessment activity, the students' next task was to suggest and implement a measure for the goodness of a proposed division of data points to clusters, and finally, test the implemented measure on the provided clustering methods and reflect upon the choice they have made when choosing their preferred clustering method. In contrast to the classification activity, this activity addressed an unsupervised learning setting, where there is no available ground truth categorization. To support students' coding, we provided them with two auxiliary methods - one that extracts all points belonging to a particular cluster, and one that computes the distance between two data points.

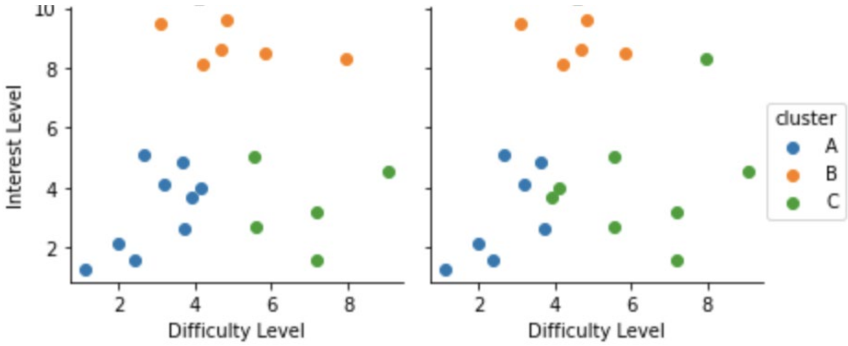


Figure 2. The Clustering Assessment activity choice between two clustering methods

Results

We describe students' outcomes and processes as identified from their video recordings.

Observed outcomes

The three authors analyzed video recordings from students' invention processes. Each invention process was segmented by turn-taking between the partners. Each segment then received a label (e.g., "identifying the importance of false negatives", "normalizing the measure according to sample size"), and these were clustered into themes. Pretty early it became clear that while students had different interactive patterns and participation models, they reached a finite set of outcomes. We repeated the procedure for two invention processes until we had reached saturation and no new themes were identified. Overall, three categories of students' outcomes were identified:

1. *Conceptual Insights* – insights that are related to the core domain, such as the distinction between different types of classification mistakes (e.g., false negatives vs. false positives).
2. *Design Approaches* - the students' approaches to formalizing their suggested measures, such as using the average between clusters centers to assess the quality of a clustering method.
3. *Nature of Model Insights* - insights that are related to the design of a quantitative measure that are not specifically related to the domain itself, such as considering a measure's boundaries.

Students' outcomes from the two activities are summarized in Table 1 and Table 2. The outcomes in each category are presented alongside examples from transcripts of the activities.

Classification assessment activity outcomes

All four pairs reached the conceptual insights presented in Table 1. That is, all pairs noticed that accuracy is not sufficient for assessing the quality of classification, and also noticed that misclassifying an at-risk person as not at-risk (false negative) is more critical than classifying a person who is not at-risk as at-risk (false positive). The design approaches were evenly distributed between the pairs. Two pairs chose to integrate the accuracy rate and the false negatives rate into a single measure. For example, one pair decided to reduce the false negatives rate from the accuracy score. The other two pairs decided to assign weights according to the type of mistakes that were made. For example, one pair assigned a higher weight (0.6) for false negatives and a lower weight (0.4) for false positives, so the total score was affected more by false negatives. Regarding the nature of model insights, all pairs reached the insight of the importance of providing a general solution that applies to different sample sizes by using normalization in their suggested measure. Two pairs that chose to integrate accuracy and false negative rate paid attention to a case in which the measure result might be negative and modified the measure such that its boundaries will be between 0 and 1.

Clustering assessment activity outcomes

All four pairs achieved the conceptual insight that a clearer separation between clusters indicates a better clustering method. All four pairs focused their design approaches on within cluster distances statistics such as the average of the within cluster distances averages, or the average of within cluster maximum distances. One of the pairs suggested a second measure that focuses on between-cluster distances statistic, and suggested using the average of distances between clusters' centers as a measure for a better clustering method. Another insight that relates to the nature of models was the distinction between choosing average or maximum distance as a statistic. The students realized that when choosing a worst-case approach such as choosing the maximum distance, a single extreme case determines the score for the entire cluster.

Table 1: Observed students' outcomes gained from the Classification Assessment Invention activity. The (PX) mark indicates which of the participating pairs is transcribed

Category	Outcomes	Transcripts
Conceptual Insights	Classification accuracy is not enough	"The second classifier is more accurate, but I don't think we should look at it that way. My opinion stays the same, I still prefer the first one, what do you think?" "I think we should do as we said earlier and consider the more critical mistakes" (P2)
	False negatives vs. false positives	"The question is what is more critical, classifying as at-risk while actually not at-risk, or vice versa?" "Let's think about it, the goal is to identify if you are at-risk. Basically, thinking that you are at-risk while you are not is less dangerous" (P1)
Design Approaches	Integrate accuracy score with critical mistakes	"Look, we can think of something that gives additional weight to specific mistake, but eventually we don't want to neglect the accuracy score, because having many mistakes, even if they are not critical, is not good either" "Maybe we should combine critical mistakes and accuracy somehow" (P1)
	Assign higher weights for critical mistakes	"Maybe we should just give a higher weight for more critical mistakes. I mean, maybe we'll give a 3/5 weight when missing at-risk person and 2/5 for missing not at-risk person" (P2)
Nature of Model Insights	Normalization	"There is something that bothers me, that this measure will be good only for a test group at the same size as in our case, but if I want a more general measure disregarding the test set size, it won't work as we want" "So, let's divide it on the size of the test set to normalize it" (P4)
	Boundaries	"Wait. what will we do if the false negative rate is higher than the success rate, it will lead to a negative result, no?" (P3)

Table 2: Observed students' outcomes gained from the Clustering Assessment Invention activity. The (PX) mark indicates which of the participating pairs is transcribed

Category	Outcomes	Transcripts
Conceptual Insights	Clear separation indicates better clustering method	"Intuitively I want to choose the first clustering method, since each cluster has its own boundaries and seems more clearly separated" "I agree, in this clustering method you can actually draw a clear separation line between the clusters" (P5)
Design Approaches	Within cluster distance statistic	"For each cluster, the points belonging to it should be closer to each other" "So, you mean that the average distances within each cluster should be lower to indicate a better clustering method" (P2)
	Between-clusters distances statistic	"We can calculate the center of each cluster and examine the distance between the clusters' centers" (P6)
Nature of Model Insights	Worst case vs. average	"If we choose maximum distance in a cluster, then a single case determines for the whole cluster" (P1)

Process analysis

As described earlier, the Invention activities in this study were composed of five consecutive tasks: read a context story, choose between a pair of contrasting cases, suggest a numeric measure for the given problem, implement the measure, and finally, test the measure on the examples given in the contrasting cases, and reflect upon the initial intuitive choice. Our process analysis focuses on linking between the activity structure and the phases that students went through while engaging with the activity. We break down the students' engagement into three phases: Analysis, Invention and Verification.

1. *Analysis* – in this phase students are introduced to the problem and develop their conceptual insights. In both activities, and for all pairs, the conceptual insights were supported by inviting students to engage with the pair of contrasting cases (task 2).
2. *Invention* - This is the main and longest phase of the activities in which the students' inventions are developed and solutions are designed and implemented. In both activities, and for all pairs, this phase occurred while engaging with tasks 3 and 4 - suggesting a measure and implementing the measure. This phase includes two intertwined components: Ideation and Implementation. Ideation is the process in which students discuss various aspects of their suggested solution and develop its fundamentals. Generally, the discussions refer to the parameters that should be taken into account, how they are formulated, and various nature of model aspects such as the measure boundaries. Implementation is the process in which students write code to implement their ideas. The Implementation phase helped the students to get into low-level details they have not paid attention to in the ideation phase, further refining their solutions. In two of the students' works, there was a clear distinction between these phases, the students developed their final solution before engaging with the implementation task, and only then started to implement it. In the remaining works, students went back and forth between those phases. For example, one pair decided to stop the ideation phase to implement false negatives and false positives counters, tested it, and then returned to develop their solution further.
3. *Verification* – in the final phase students test their suggested measure and reflect upon their work. This phase occurs while engaging with the final task (task 5) in which the students were asked to return to the beginning of the activity and test their implemented measure on the contrasting cases from task 2 and reflect upon their initial choice. This phase highlights the benefit of working in a code-based environment which enables implementation and testing cycles.

We demonstrate the invention process by describing in detail the Classifier Assessment activity of a single pair (P1). We describe the different phases, their duration and outcomes, supported by the activity transcripts.

Analysis (minutes 0-7) – In this phase, the students engaged with tasks 1 and 2. In the first two minutes they read the introduction story (task 1) and in the following five minutes they discussed the contrasting cases, attempting to pick the better classifier (task 2). Through these discussions they gained two conceptual insights: (1) *classification accuracy is not enough,* "The first classifier had three mistakes, the second had two, but the first did successfully classify one at-risk person while the second did not", and (2) *false negatives are worse than false positives*:

> Student 1: "The question is which mistake is more critical, classifying as at risk while actually not at risk, or vice versa?"
>
> Student 2: "Let's think about it, the goal is to identify if you are at risk. Basically, thinking that you are at risk while you are not is less dangerous"

Invention (minutes 7-37) – Based on the conceptual insights gained in the analysis phase, the students next tried to suggest a measure (task 3). In this work, there was a clear distinction between ideation and implementation - the students finalized their solution idea in the ideation phase and implemented it exactly as suggested in the implementation phase. In the *Ideation* part (minutes 7-27), the students discussed the parameters that should be considered and formulated the measure, while also addressing the general nature of model aspects. First, they came up with the idea of integrating the accuracy score with critical mistakes:

Student 1:	"Look, we can think of something that gives additional weight to specific mistakes, but eventually we don't want to neglect the accuracy score, because having lots of mistakes, even if they are not critical, is not good either".
Student 2:	"Maybe we should combine the critical mistakes with the accuracy somehow".

Next, they came up with a concrete way in which they can integrate the different parameters, "OK, so let's say we have our accuracy rate, the question is what exactly we do with the false negatives and false positives".

Finally, they raised the issue of the boundaries of the model (nature of models insight):

Student 1: “OK, so the accuracy rate is our upper limit”

Student 2: “Wait, but can it be smaller than zero if it is really bad?”

The *Implementation* (minutes 27-37) step was mostly technical. The student wrote code for their measure, making use of the provided auxiliary method for computing accuracy.

Verification (minutes 37-50) - After implementing the measure, the students moved to the final task of testing the measure using the classifiers presented in the contrasting cases (task 2) and reflected on their work. Interestingly, when *Testing* their measure (minutes 37-42), the students tried to predict the output they expect to get from the measure before running their code, “Let’s verify we know what to expect, to verify that [the code] is correct”. They also went beyond the required testing (of the two provided classifiers) and created a new test-case to further examine their measure by modifying the raw data to include more critical mistakes:

Student 1: “We can add more critical mistakes to the raw data and verify we get a reduced score; you want to try?”

Student 2: “Sure”

Finally, in their *Reflection* on the invented measure (minutes 42-50), the students expressed satisfaction from their work, “I am proud of us, this measure is not bad at all”, while acknowledging the limitations of their measure, “Generally, our measure is not suitable for small data such as we got here”. They further noted that there are likely other solutions:

Student 1: “Surely there are other ways to measure classification besides addressing critical mistakes, but eventually you have to give more importance to the type of mistake, so I think we did well with the given time we had, no?”

Student 2: “I think so…”

Discussion

We used carefully designed Invention activities to improve the teaching and learning experience on the topics of classification and clustering in an introductory Data Science course for undergraduate students. The Invention activities took place a couple days prior to the lectures, in which an overview of the students’ solutions was presented, and expert solutions were taught.

The engagement with the activity has led the students to impressive outcomes, including gaining important conceptual insights of the domain, providing valid and complete design approaches for solutions, implementing the suggested solutions while discussing various design aspects including those related to the nature of models (e.g., measure boundaries), and finally, reflecting and analyzing their work.

Design approach for data science invention activities

The activity design guided students through a process that was composed of three main phases: Analysis, Invention and Verification. Each of these phases contributed to the outcomes and insights students achieved.

We found that the stories encouraged students to use intuitive knowledge when analyzing the cases and concise contrasting cases helped students notice deep features of the domain. Notably, the contrasting cases were of small data, compared with the typical data science data, in order to enable sense-making. For example, in the Classification Assessment activity, the context story was used as an example in which it was easier to identify the critical mistake - at-risk person that was classified as not at-risk (false negative). The choice between the contrasting cases was used to highlight the problem. The sole method that students knew was the accuracy score, but in the presented problem it was not enough.

Data science makes heavy use of code-based environments. Code-based environments can be useful since they provide tools for easier exploration and testing. In our study, we found that they supported students in iterative ideation-and-implementation. However, they also add a challenge (and extraneous cognitive load) since coding requires technical skills that are not the main focus of the activity. This might lead students with lower coding skills to frustration. To reduce frustration, we added auxiliary methods that could serve as building blocks for the students’ implementation, such as providing a method that extracts all points in a specific cluster, or a method for calculating the distance between two given points. In the trade-off between open exploration and detailed support, through the use of generic methods (rather than developed answers), we tried not to channel students towards specific solutions.

Asynchronous invention activities

Working in an out-of-class online environment opens the opportunity for running large scale Invention activities. On the other hand, it raises various challenges, such as the lack of instructor presence and the total freedom given to the students. When there is no instructor present to guide the students through the activity, the students might get stuck and disengage since they cannot seek help. The converse is also true - students may search google for answers, and thus get too much support that short-circuit the invention process. In our study, we designed several elements that reduced the risks of dropping out or googling for answers. First, the activities were highly structured, providing clear guidance to students as they work. This helped prevent students from 'getting lost'. Second, we found that having students work in pairs was instrumental. Students consulted with each other, challenged each other, and completed each other's ideas. To reduce the risk of students googling answers, we used general language and did not specifically refer to formal domain terms. For example, we did not name the classification activity 'classifier assessment activity', rather we named it 'Meditest Project'. Another measure we took to prevent these "shortcuts" was to highlight to students in lecture that they do not need to find an optimal solution, and that often there is no such single solution. Instead, they were encouraged to come up with alternative measures.

The study has two main limitations. First, the sample size is fairly small, and the students self-selected to the study. Second, we did not evaluate the learning outcomes beyond the activity itself. Future work will address these limitations by evaluating the efficacy of Invention activities and follow-up instruction with a larger sample.

Conclusion

We put forward a design approach for asynchronous Invention activities for learning challenging concepts in data science. Analysis of the outcomes highlighted key insights that students reached through the invention process. Succeeding in implementing Invention activities at scale can add much-needed interactivity to online education, and specifically to data science education.

References

Berman, F., Rutenbar, R., Hailpern, B., Christensen, H., Davidson, S., Estrin, D., … Szalay, A. S. (2018). Realizing the potential of data science. *Communications of the ACM*, *61*(4), 67–72. https://doi.org/10.1145/3188721

Chase, C. C., & Klahr, D. (2017). Invention Versus Direct Instruction: For Some Content, It's a Tie. *Journal of Science Education and Technology*, *26*(6), 582–596. https://doi.org/10.1007/s10956-017-9700-6

Hew, K. F. (2016). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCS. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.12235

Holmes, N. G., Day, J., Park, A. H. K., Bonn, D. A., & Roll, I. (2014). Making the failure more productive: scaffolding the invention process to improve inquiry behaviors and outcomes in invention activities. *Instructional Science*, *42*(4), 523–538. https://doi.org/10.1007/s11251-013-9300-7

Kapur, M. (2016). Examining Productive Failure, Productive Success, Unproductive Failure, and Unproductive Success in Learning. *Educational Psychologist*. https://doi.org/10.1080/00461520.2016.1155457

Kapur, M., & Bielaczyc, K. (2012). Designing for Productive Failure. *Journal of the Learning Sciences*, *21*(1), 45–83. https://doi.org/10.1080/10508406.2011.591717

Loibl, K., Roll, I., & Rummel, N. (2017). Towards a Theory of When and How Problem Solving Followed by Instruction Supports Learning. *Educational Psychology Review*, *29*(4), 693–715. https://doi.org/10.1007/s10648-016-9379-x

Perkel, J. M. (2018). Why Jupyter is data scientists' computational notebook of choice. *Nature*. https://doi.org/10.1038/d41586-018-07196-1

Roll, I., Russell, D. M., & Gašević, D. (2018). Learning at Scale. *International Journal of Artificial Intelligence in Education*, *28*(4), 471–477. https://doi.org/10.1007/s40593-018-0170-7

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475–5223. https://doi.org/10.1207/s1532690xci1604_4

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing Versus Inventing With Contrasting Cases: The Effects of Telling First on Learning and Transfer. *Journal of Educational Psychology*, *103*(4), 759–775. https://doi.org/10.1037/a0025140

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. https://doi.org/10.1207/s1532690xci2202_1

Schwartz, D. L., Sears, D., & Chang, J. (2012). Reconsidering prior knowledge. *Thinking with Data*, (650), 319–344. https://doi.org/10.4324/9780203810057

 © ISLS

My Partner was a Good Partner: Investigating the Relationship between Dialogue Acts and Satisfaction among Middle School Computer Science Learners

Gloria Ashiya Katuka, University of Florida, gkatuka@ufl.edu
Richard T. Bex, University of Florida, rbex@ufl.edu
Mehmet Celepkolu, University of Florida, mckolu@ufl.edu
Kristy Elizabeth Boyer, University of Florida, keboyer@ufl.edu
Eric Wiebe, North Carolina State University, wiebe@ncsu.edu
Bradford Mott, North Carolina State University, bwmott@ncsu.edu
James Lester, North Carolina State University, lester@ncsu.edu

Abstract: Collaborative dialogue provides a rich information source for understanding the effectiveness of student interactions. While many studies emphasize the importance of productive dialogue behaviors, the impact of those behaviors on learners' perceptions of their partners is not yet understood. This paper examines a dialogue corpus of 18 pairs of middle school students as they engage in block-based coding activities. We tagged the corpus with a collaborative dialogue act taxonomy and identified sequences of one to two dialogue acts (*n*-grams) that are significantly associated with partner satisfaction during collaborative learning. Six *n*-grams were found to be significant predictors: *n*-grams that were positively associated with satisfaction included some questions and clarifications. In contrast, *n*-grams that were negatively associated with satisfaction included off-task utterances, pairs of consecutive questions, and unexpectedly, positive feedback. These findings contribute to our understanding of how learners prefer to interact with their partners and how that interaction impacts collaborative experiences.

Introduction

Collaborative dialogue constitutes one of the main channels for students to exchange information and co-construct knowledge (Wegerif, 2011; Mercer et al., 2019; Major et al., 2018) and has attracted considerable interest among computer-supported collaborative learning (CSCL) researchers (e.g., Madaio et al., 2017; Stahl, 2015; Rosé et al., 2008). Dialogue provides numerous cues and opportunities for understanding the effectiveness of collaboration, and thus there is a growing body of research concerning the types of dialogue behaviors that lead to better learning (Chi & Wylie, 2014). On the other hand, analyzing collaborative dialogue is a challenging process due to the dynamics and complexity of group interactions. There is still a need for developing instruments and methodologies to understand how certain dialogue moves occur and how they impact students' learning (Howe, 2017; Hennessy et al., 2016). In recent years, CSCL research has investigated collaborative dialogue for understanding students' socio-metacognitive dialogue patterns (Borge et al., 2019), dialogue transactivity and epistemic quality (Schmitt & Weinberger, 2017), reasoning processes (Snyder et al., 2019), and how students express and address uncertainty (Rodríguez et al., 2017).

In this paper, we extend this body of research by investigating the collaborative dialogue patterns that lead to higher partner satisfaction among middle school students in the context of pair programming. In pair programming, students take on structured roles: the driver's role is to control the mouse and keyboard and focus their effort on building and editing code, while the navigator's role is to observe the work being done by the driver to identify potential errors, provide suggestions, and ask clarification questions (Williams & Kessler, 2003). Pair programming holds great promise for supporting students' learning and engagement in K-12 settings (Campe et al., 2020; Denner et al., 2014), yet, several studies have reported that the demanding nature of collaborative learning can lead to challenges for younger learners, who lack effective collaboration skills (Deitrick et al., 2016; Lewis & Shah, 2015). If these challenges are not addressed, students may develop negative dispositions toward collaboration in the future (Schultz et al., 2010). Thus, it is crucial to deeply examine student dialogues during collaborative learning activities to reveal what kind of dialogue patterns are present and how those dialogue patterns are related to learners' perceptions of their partners. Identifying dialogue patterns that are predictive of learners' satisfaction with their partners can help researchers and educators to understand and facilitate more positive collaborative learning experiences. This study focuses on two research questions: (1) *What dialogue acts emerge during collaborative dialogue within pairs of middle school students during coding activities*, and (2) *How are the dialogue acts associated with outcomes related to partner satisfaction?*

To investigate these research questions, we first developed a taxonomy consisting of 15 dialogue acts, which provides a high-level representation of the underlying meaning of student dialogues, based on a corpus of collaborative dialogue from 36 middle school students who completed a coding activity in pairs. Next, we examined sequences of dialogue acts of length one to two (*n*-grams) and generated a linear regression model which used the frequency of the *n*-grams to predict partner satisfaction, with the *n*-grams as predictors of a derived *satisfaction* outcome (the average of partner-related post-survey items). The results showed significant associations between six dialogue *n*-grams and the learners' satisfaction with their partners. Learners reported higher partner satisfaction when their partners were more engaged, such as by asking questions, seeking clarifications, and actively talking about the task. The results also showed that when their partner frequently responded with positive feedback or when both collaborators engaged in off-task dialogues, learners reported lower partner satisfaction. These findings provide us with a better understanding of how learners would prefer their partners to interact with them and how they prefer to interact with their partners when participating in pair programming activities.

Background: Dialogue analysis

Within dialogue analysis, representing conversations at the utterance level, such as through *dialogue acts*, has long been studied. Dialogue acts are a higher-level representation of the intention of the user (Austin, 1975), and dialogue act tagging involves labeling an utterance with a predefined dialogue act that provides information about that utterance. Each dialogue turn is considered as one utterance; thus, an utterance can serve as a smaller unit of communication that describes a single event (Polanyi et al., 2004). An utterance can be an incomplete or grammatically broken sentence but still have a role in conversation depending on the context (Bakhtin, 2010). A dialogue act expresses the nature of a communicative behavior between a sender and addressee that has an effect on the context of understanding the behavior (Bunt, 2005). Previous research has investigated the ways in which dialogue acts are associated with learning outcomes (Dubovi & Lee, 2019; Olsen & Finkelstein, 2017) and motivation (Meier et al., 2007). In this paper, the goal of dialogue act tagging is to classify the utterances to show collaborative patterns that are associated with partner satisfaction.

Methods

Participants and context

This work is part of a larger project aimed at developing computer science knowledge and deepening understanding of science concepts through computationally rich science activities for middle school students (Celepkolu et al., 2020). To achieve this goal, the research team collaborated with a middle school science teacher to implement a series of computer coding lessons as part of their regular classroom activities. The students learned about the fundamentals of coding, such as loops, conditionals, and variables, and applied their coding knowledge to create science models and simulations, such as homeostasis and evolution, using the Snap*!* block-based programming environment. The researchers explained the driver and navigator roles in pair programming, the expectations for each role, and reminded students to switch roles regularly (12-15 minutes). Data was collected as part of an IRB-approved study that included written parental consent and student assent. The researchers implemented the activities during a science class in two semesters (Spring and Fall 2019), which was taught by the same teacher and followed the same structure. Out of 204 students, 145 students provided assent and parental consent, and we randomly selected 19 pairs (38 students) to audio/video record their interactions during the coding activities (24 students in Spring 2019 and 14 students in Fall 2019). Out of these students, there were 23 girls (60.5%) and 15 boys (39.5%). The distribution of race/ethnicities was 14 White (36.8%), 2 Hispanic (5.3%), 7 Asian (18.5%), 10 Multiracial (26.2%), and 5 Other (13.2%). The mean age was 12.1, with ages ranging from 11 to 13, and 53% of students reported having had some prior coding experience at the beginning of the semester.

Procedure

In every class, researchers assisted the teacher by presenting an introduction to the science topics and providing students with a copy of the written instructions. Next, students worked on activities for 35-40 minutes with a randomly assigned partner. During these activities, the teacher and researchers were available to help students with their questions. After pairs participated in the collaborative work sessions, students were asked to individually complete a post survey. We developed the post survey items because, to the best of our knowledge, there is no existing survey that captures partner satisfaction within the pair programming context. From the post survey, this paper utilizes the following six questions for analysis: *(1) "My partner answered my questions well," (2) "My partner listened to my suggestions," (3) "My partner often cut my speech" (which was reversed scored* *for accurate computation), (4) "My partner was comfortable asking me questions," (5) "My partner asking questions helped me think about things differently," and (6) "Overall, my partner was a good partner."* Responses followed a 5-point Likert scale, with 1 representing "strongly disagree" and 5 representing "strongly agree." Figure 1 shows the distribution of the survey responses from both studies. Most students agreed or strongly agreed that their collaboration with their partner was successful.

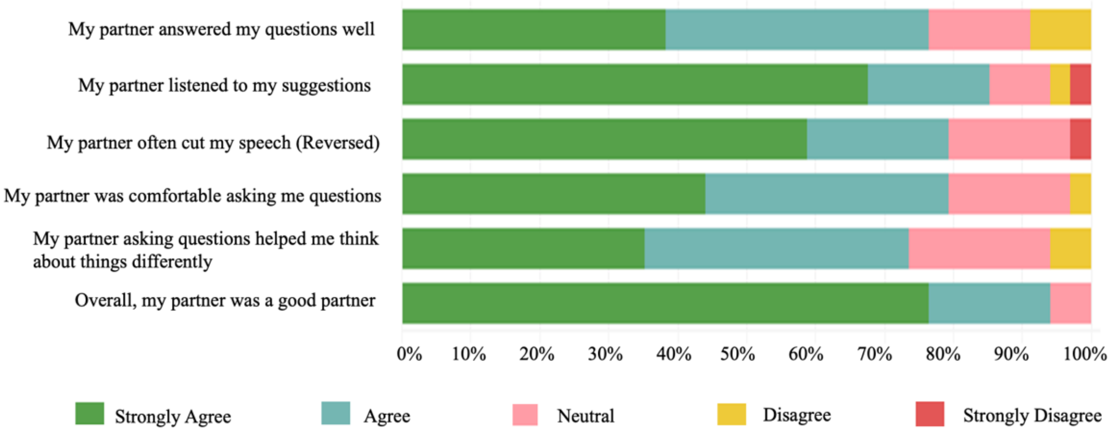


Figure 1. Post survey responses related to partner satisfaction

Dialogue corpus and annotation

We manually transcribed the 19 video recordings of students collaborating, which resulted in 8,940 dialogue utterances. Next, we tagged each utterance based on the function of the information in the dialogue. Prior to tagging our dataset, we filtered and removed utterances directed toward anyone other than the learner's partner (teachers, researchers, and other students). Next, we removed one session that contained large amounts of chatter and indistinguishable utterances. Lastly, we also removed all utterances that were untranscribable due to audio quality. Our final student-student dialogue corpus included 18 sessions (36 students) and 4,859 utterances with a mean of 242 utterances per session (*SD* = 118, *Min* = 93, *Max* = 526) and a mean of 121 utterances per student (*SD* = 114, *Min* = 42, *Max* = 264).

To develop a taxonomy for our corpus, we reviewed the existing taxonomies within closely related fields and age groups and considered relevant taxonomies (Core & Allen, 1997; Rodríguez et al., 2017; Tsan et al., 2018). Previous research has established dialogue act tags for pair programming among college students (Rodríguez et al., 2017) as well as among elementary school students (Tsan et al., 2018), and our work took the union of these two taxonomies as its starting point, producing 19 initial tags. Several iterations of dialogue tag application and refinement revealed, as expected, that some of the tags from the taxonomies were not present in the current middle school corpus, and that some new or modified tags were needed. A process of iterative refinement of the tagging scheme in several rounds of collaborative and then independent tagging produced a final dialogue act taxonomy of 15 tags (Table 1). Eight of these tags were adopted from Rodríguez et al. (2017): *Statement*, *Acknowledgement*, *Uncertain*, *Meta comment*, *Positive Feedback*, *Non-Positive Feedback* and *Off-task*. From the Tsan et al. (2018) scheme, three tags were adopted: *Make Suggestions*, *Acceptance*, and *Rejection*. The newly developed tags are *Next Step*, *Seek Clarification*, *Question*, and *Seeking Attention*. Both annotators independently tagged 23% of the dataset and achieved an inter-rater agreement score Cohen's kappa of .83 (Landis & Koch, 1977) indicating "almost perfect" agreement. The two annotators then each tagged half of the remaining utterances so that the entire corpus was tagged.

Data analysis

Our next goal was to discover the ways in which student dialogue acts were related to the outcomes reported on six post survey items. To determine whether to treat these six post survey items as a single item or multiple items, we conducted a principal component analysis (PCA). The results of PCA suggested proceeding with only one derived outcome variable, which we refer to as *satisfaction*. The single component explains 52% of the variation across all six survey items with eigenvalue 3.15. The distribution of the *satisfaction* outcome shows 76% of the learners agreed or strongly agreed that they were satisfied with the overall interaction with their partner.

Our overarching goal was to identify the ways in which dialogue acts (or sequences of them) were associated with partner satisfaction. From our tagged dialogue corpus, we proceeded to extract sequences of dialogue acts, known as *n*-grams, which will be treated as predictors within a regression model. To extract the sequence of *n*-grams, we applied standard practices from previous dialogue analyses (Forbes-Riley & Litman, 2005). In our work, we generated *n*-grams of dialogue act tags for each learner's dialogue using a sliding window of *n*=1 (unigrams) and *n*=2 (bigrams). We used a sliding window approach and assigned each dialogue act tag a student or partner subscript (*e.g.*, $Statement_{stu}$, $Question_{par}$) to indicate whether the utterance originated from the student or their partner. Every learner was tagged as a student to ensure that we extracted each *n*-gram from each learner's perspective. Each row in the resulting dataset corresponds to a student whose own dialogue moves contain the subscript "*stu*" and whose respective partner's dialogue moves contain the subscript "*par.*" Each student played the role of a "driver" as well as a "navigator" during the pair programming task, and these roles are not indicated within the bigrams. We extracted 563 *n*-grams, 30 distinct unigrams and 533 distinct bigrams. Unigram frequencies are shown in Table 1. The most frequent bigrams were ($Statement_{stu}$, $Statement_{par}$), ($Question_{stu}$, $Statement_{par}$), and ($Statement_{stu}$, $Question_{par}$), which occurred 677, 291, and 235 times, respectively.

Table 1: Taxonomy of Dialogue Acts

Dialogue Act	**Frequency**	**Description**	**Example(s)**
Statement	1622	Makes a statement of information, an explanation, or a response to an inquiry	*"This looks like it's not moving at all."* *"Oh, we forgot to put repeat forever."*
Off-task	733	Interacts with someone other than their partner or off-topic conversations with their partner	*"I have a, we have an orchestra test today."* *"You like my new look?"*
Question	604	Asks partner for help or information seeking some feedback from the partner with regards to the task.	*"Do I put this in here?"* *"Do we just have to put it together now?"*
Directive	405	Provides an explicit instruction to their partner	*"Push the restart button."* *"Click on the amplitude variable."*
Acknowledgment	216	Accepts or acknowledges the previous statement or utterance	*"Okay."*
Meta Comment	193	Makes a meta response to something relating to the task	*"Um, uh…"* *"Oh my gosh."*
Uncertainty	133	States an opinion or indication of uncertainty or confusion	*"Maybe. I don't know."* *"I'm a little confused."*
Seek Clarification	112	Asks for further clarification on something mentioned earlier or referred to in the text	*"What?"* *"Which one?"* *"What do you mean?"*
Positive Feedback	88	Provides positive feedback related to a task action completed by themselves or their partner	*"There! We finally did it."* *"Oh, ours is good."* *"Yeah that's good, it's good."*
Make Suggestion	81	Makes a suggestion or contributes an idea without explicitly asking the partner to do something	*"Maybe make a new forever loop just for that."* *"Let's go back to the directions because it will tell us what code to use."*
Non-positive feedback	54	Provides negative feedback on the task or something incorrectly done by themselves or their partner	*"Wait, try the, oh that's not gonna work."* *"We don't need that."*
Next Step	52	Makes a suggestion for what they believe should be the next step to be completed in the near future	*"And then I think you're supposed to put it in."* *"And then change variables.."* *"And then we can do the operators."*
Acceptance	44	Accepts or acknowledges their partner's idea, suggestion, or directive. (Follows a MS, NS, or D)	*"Yes."* *"Right."*
Seeking Attention	12	Seeks partner's attention while working on task	*"Hello?"* *"Bro."*
Rejection	10	Rejects a direct instruction or idea or suggestion	*"No."*

Results

To determine the *n*-grams that were significant predictors of the *satisfaction* outcome, we conducted a regression analysis using the JMP statistical software. To mitigate the problem of a large number of *n*-gram predictors for a smaller sample size (563>>36) we included only *n*-grams that occurred in at least half of the sessions. The remaining 78 predictors included 30 unigrams and 48 bigrams. We provided these 78 *n*-grams as predictors and the derived *satisfaction* variable as the outcome variable to a generalized regression model. We selected the best subset estimation method, which uses an exhaustive algorithm that fits and assesses all possible models and chooses the best subset to predict the outcome variable. We used the AIC (Akaike information criterion) statistic as the goodness-of-fit measure. Table 2 shows the regression results, including the six *n*-grams that satisfied the test for statistical significance ($p < .05$). The regression model passes the test for multicollinearity with all resulting variance inflation factor (VIF) values less than 2 (VIF values greater than 5 often indicate multicollinearity). The adjusted R^2 of .74 shows that the model explains 74% of the variance in partner satisfaction.

Table 2: Generalized Regression Model (Best subset method) of *n*-grams as predictors of partner satisfaction

Dialogue Act *n*-gram	Estimate	Standardize Estimate	Std Error	VIF
Intercept	3.853	1.477	0.126	0
Question_{stu}, $\text{Seek Clarification}_{par}$	1.945	0.601	0.091	1.99
Directive_{stu}, Question_{par}	1.092	0.209	0.025	0.878
Statement_{par}	2.119	0.681	0.002	1.122
$\text{Positive Feedback}_{par}$	-1.406	-0.938	0.019	1.817
Question_{par}, Question_{stu}	-1.736	-0.941	0.018	1.162
Off-task_{stu} Off-task_{par}	-1.413	-0.941	0.002	1.064

Note: The model only contains significant *n*-gram predictors with $p<.001$.

As the parameter estimates in Table 2 show, three *n*-grams are positively related to partner satisfaction: (1) a question by the student followed by their partner seeking clarification, (2) a directive by the student followed by a question from their partner, and (3) a statement from their partner. In contrast, the model also revealed that three *n*-grams are negatively associated with partner satisfaction: (1) the student initiates a conversation not related to the task and their partner responds and continues with the unrelated conversation, (2) positive feedback from their partner, and (3) a question from their partner followed by a question from the student.

Discussion and implications

The overarching goal of this study was to develop a better understanding of how dialogue acts are associated with partner satisfaction for middle school students during collaborative coding. The model identified six statistically significant *n*-grams, including three bigrams and one unigram that may indicate an interactive partnership: ($\textit{Question}_{stu}$, $\textit{Seek Clarificaton}_{par}$), ($\textit{Directive}_{stu}$, $\textit{Question}_{par}$), ($\textit{Question}_{par}$, $\textit{Question}_{stu}$), and $\textit{Statement}_{par}$. Conversely, ($\textit{Off-task}_{stu}$, $\textit{Off-task}_{par}$) and—perhaps counterintuitively—$\textit{Positive Feedback}_{par}$, may suggest a tendency for reduced engagement or distractions in this context. This section discusses these findings in turn.

Three of the significant *n*-grams within the model include asking questions. The literature has clearly established the role of questions in collaborative learning as a means of establishing and sustaining mutual understanding (Spada et al., 2005). Asking questions also elicits a constructive engagement between collaborators by presenting an avenue to generate new ideas (Chi & Wylie, 2014). When learners ask their partner questions, they create a channel for dialogue interaction by taking the first step to access information and resolve confusion (Chin & Osborne, 2008). The analysis results indicate that question-related dialogues are significant indicators of partner satisfaction. Higher occurrence of bigrams where learners ask their partner a question followed by their partner seeking clarification are associated with that learner reporting higher partner satisfaction. This finding is likely related to the importance of understanding a question before attempting to answer it. For example, one student said, *"Why is the wavelength N-A-N?"*($\textit{Question}_{stu}$) and their partner replied, by *"Nan what?"*($\textit{Seek Clarification}_{par}$). Here, the partner is making an effort to better understand the student, and the student subsequently reported higher satisfaction with that partner. Similarly, the results show a positive correlation between partner satisfaction and higher occurrences of instances when a partner asks a question after receiving a directive/instruction from the student. For example, a student said, *"Okay. Now, create a variable."*($\textit{Directive}_{stu}$) and their partner asked, *"Named what?"*($\textit{Question}_{par}$). Here the question *"Named what?"*($\textit{Question}_{par}$) refers to seeking new information. This is different from seeking clarification, which refers to a question or information already stated previously.

Not all occurrences of *Question* were positively associated with the satisfaction outcome. The more often collaborators asked back-to-back questions without a response to the first question, the less likely the student reported a high satisfaction rating. For example, a student asked, *"Why did it set the generation to zero?"(Question$_{stu}$)* and their partner asked, *"So, wait is this what we're supposed to do?"(Question$_{par}$)*. Unanswered questions and unresolved uncertainty have been linked to less positive outcomes in other work on collaborative coding for dialogue as well (Rodríguez et al., 2017).

Another dialogue act whose frequency was highly predictive of satisfaction is a statement from the partner. Statements are one of the most prominent conversational dialogue moves in the corpus. The findings in this study are consistent with previous results where statements were shown to be associated with effective collaboration (Rodríguez et al., 2017). Statements can indicate more active engagement by the partner, which improves learning outcomes by facilitating advancement from *constructive* to *interactive* behavioral modes (Chi & Wylie, 2014). In the current corpus, the most frequent occurrences of statements were in response to a question, directive, or acknowledgment by the student. For example, a student asked, *"Right or does it not get longer?" (Question$_{stu}$)* and their partner responded with *"It doesn't get longer."(Statement$_{par}$)*. The second most common occurrence of a statement from the partner is as a response to a directive from the student. For example, one student said, *"Wait, increase the clone counter by one,"(Directive$_{stu}$)* and their partner responded, *"I think they do the same thing."(Statement$_{par}$)*. In the third most common occurrence of a statement, a statement is followed by an acknowledgment. For example, an *"Okay" (Acknowledgement$_{stu}$)* from the student was followed by *"Um, when a clone is spawned it should increase the clone generation counter by one, uh clone generation counter…" (Statement$_{par}$)* by their partner.

In addition to the previously mentioned sequence of back-to-back *Questions*, two others emerged as negative predictors of partner satisfaction: *Positive Feedback$_{par}$*, and (*Off-task$_{stu}$, Off-task$_{par}$*). Feedback within peer collaboration has been shown to positively enhance interpersonal behaviors and social performance (Phielix et al., 2010), but a potential explanation for the negative association of a partners' positive feedback with a student's perception of that partner might be the possibility of the partner compensating for lower participation, which eventually becomes apparent (Prinsen et al., 2007). In the corpus, the most common occurrence of positive feedback by the partner followed a statement. For example, one student said, *"Now, we're going to do this. There we go." (Statement$_{stur}$)* and their partner responded with *"Yay."(Positive Feedback$_{par}$)*. It is also possible that a positive feedback response might function as the partner doubtfully accepting the student's assertions. This could be due to the partner not fully understanding their role, the task, or their ability to effectively contribute to the collaboration.

As for the off-task bigram's role in predicting partner satisfaction, recent CSCL research investigated the impact of off-task exchanges during collaborative problem solving such as lower participation and distraction from the task (Cheng et al., 2020). In the corpus, we see threads of *Off-task* utterances that can pose a distraction to the collaborators. This may result in the collaborators not completing their tasks and lower satisfaction in their interaction. For example, the utterance, *"It's so surprising because my parents don't believe in bath and body works."(Off-task$_{stu}$)* by the student followed by *"Really? They don't believe in bathworks."* (*Off-task$_{par}$*) by their partner sets the tone for more off-task dialogue. This exchange shows a mutual distraction between collaborators that can deviate the conversation from the task at hand. These pairs of off-task utterances are associated with lower partner satisfaction in the current context.

Implications

The findings discussed here hold several potential implications for research and practice. The findings have shown that dialogue moves indicating an interactive give-and-take, including questions, clarification questions, and elaboration, are positively associated with partner satisfaction while other phenomena such as sequences of off-task dialogue acts are negatively associated. Some seemingly positive moves, such as positive feedback from the partner, were negatively associated with a learner's satisfaction with that partner, and these phenomena warrant deeper investigation for several reasons. For example, they tell us that as we move toward using natural language processing to automatically analyze and support real-time collaboration, we must take great caution in interpreting utterances at face value: positive sentiment, whether in on-task or off-task utterances, may express a wide variety of underlying states and different levels of engagement. Additionally, while a tremendous body of literature shows the importance of certain dialogue moves including question asking, the results here suggest that the ways in which these questions are incorporated into collaborative dialogue could have a significant impact on outcomes.

Conclusion

The overarching goal of this study was to explore the relationship between dialogue patterns and partner satisfaction during pair programming activities. The findings suggest that collaborative dialogue acts that reflect interactive partnerships and active participation between learners are associated with higher satisfaction ratings, whereas dialogue acts that reflect lower participation and distraction during collaborative activities are associated with lower satisfaction ratings. This research contributes to a better understanding of the ways in which learners' and their partners' interaction during CSCL activities impact the collaborative learning process. Several limitations of this work are important to note. First, our resulting model uses partner satisfaction as its primary outcome, rather than measures of learning or process-oriented metrics of collaboration. This intentional choice is due to the importance of learners' affective and motivational states during collaboration, for which satisfaction with a partner is an important component. Second, a limitation of this work is that the relationship between dialogue acts and partner satisfaction is correlational and not causal. Finally, an additional limitation is that the studies were only conducted with middle school students from the southeastern United States and important cultural differences in other contexts may influence the generation of dialogue moves and findings.

There are several promising directions for future work. First, while this work investigated the relationship between dialogue patterns and partner satisfaction, it is also important to examine whether these patterns are also associated with learning outcomes or process-level collaborative metrics. Secondly, the relationship between partner satisfaction and effective learning outcomes should be further examined. Moreover, there is a need for examining the dialogue patterns for different pair compositions by characteristics such as gender, experience level, and personality. Additionally, deeper qualitative analysis can shed further light on how these dialogue patterns influence partner satisfaction. Finally, these findings can inform the design of adaptive support for computer-supported collaborative learning technologies, which use rich data from student dialogues.

References

Austin, J. L. (1975). *How to Do Things with Words* (2nd ed.). Clarendon Press.

Bakhtin, M.M. (2010). *Speech Genres and Other Late Essays*. University of Texas Pres.

Borge, M., Aldemir, T., & Xia, Y. (2019). Unpacking Socio-Metacognitive Sense-Making Patterns to Support Collaborative Discourse. *In Proceedings of Computer Supported Collaborative Learning (CSCL),* 320-327.

Bunt, H. (2005). A framework for dialogue act specification. *In Proceedings of SIGSEM WG on Representation of Multimodal Semantic Information,* Tilburg, Netherlands.

Campe, S., Denner, J., Green, E., & Torres, D. (2020). Pair programming in middle school: variations in interactions and behaviors. *Computer Science Education, 30*(1), 22–46.

Celepkolu, M., Fussell, D. A., Galdo, A. C., Boyer, K. E., Wiebe, E. N., Mott, B. W., & Lester, J. C. (2020). Exploring Middle School Students' Reflections on the Infusion of CS into Science Classrooms. *In Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (pp. 671-677).

Cheng, X., Fu, S., de Vreede, G.-J., & Li, Y. (2021). Using Collaboration Engineering to Mitigate Low Participation, Distraction, and Learning Inefficiency to Support Collaborative Learning in Industry. *Group Decision and Negotiation, 30,* 171-190.

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist, 49*(4), 219–243.

Chin, C., & Osborne, J. (2008). Students' questions: a potential resource for teaching and learning science. *Studies in Science Education, 44*(1), 1-39.

Core, M. G., & Allen, J. (1997). Coding dialogs with the DAMSL annotation scheme. *In AAAI Fall Symposium on Communicative Action in Humans and Machines* (Vol. 56, pp. 28-35).

Deitrick, E., Shapiro, R. B., & Gravel, B. (2016). How do we assess equity in programming pairs? *In Proceedings of the 12th International Conference of the Learning Sciences (ICLS)* (pp. 370–377).

Denner, J., Werner, L., Campe, S., & Ortiz, E. (2014). Pair Programming: Under What Conditions Is It Advantageous for Middle School Students? *Journal of Research on Technology in Education, 46*(3), 277–296.

Dubovi, I., & Lee, V. (2019). Comparing the Effectiveness of Supports for Collaborative Dialogic Sense-Making with Agent-Based Models. *In Proceedings of Computer Supported Collaborative Learning (CSCL)* (pp. 88-95).

Forbes-Riley, K., & Litman, D. J. (2005). Using bigrams to identify relationships between student certainness states and tutor responses in a spoken dialogue corpus. *Proceedings of the 6th SIGDIAL Workshop on Discourse and Dialogue* (pp. 87-96).

Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., García-Carrión, R., Torreblanca, O., & Barrera, M. J. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction, 9*, 16–44.

Howe, C. (2017). Advances in research on classroom dialogue: Commentary on the articles. *Learning and Instruction, 48,* 61-65.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Lewis, C. M., & Shah, N. (2015). How Equity and Inequity Can Emerge in Pair Programming. *Proceedings of the Eleventh Annual International Conference on International Computing Education Research,* 41–50.

Madaio, M., Cassell, J., & Ogan, A. (2017). The impact of peer tutors' use of indirect feedback and instructions. *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning (CSCL),* 383–390.

Major, L., Warwick, P., Rasmussen, I., Ludvigsen, S., & Cook, V. (2018). Classroom dialogue and digital technologies: A scoping review. *Education and Information Technologies, 23*(5), 1995–2028.

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.

Mercer, N., Hennessy, S., & Warwick, P. (2019). Dialogue, thinking together and digital technology in the classroom: Some educational implications of a continuing line of inquiry. *International Journal of Educational Research, 97*, 187–199.

Olsen, J. K., & Finkelstein, S. (2017). Through the (thin-slice) looking glass: An initial look at rapport and co-construction within peer collaboration. *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning (CSCL)* (pp. 511–518).

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.

Polanyi, L., Culy, C., van den Berg, M., Thione, G. L., & Ahn, D. (2004). Sentential structure and discourse parsing. *In Proceedings of the 2004 ACL Workshop on Discourse Annotation - DiscAnnotation '04*.

Prinsen, F., Volman, M. L. L., & Terwel, J. (2007). The influence of learner characteristics on degree and type of participation in a CSCL environment. *British Journal of Educational Technology, 38*(6), 1037–1055.

Rodríguez, F. J., Price, K. M., & Boyer, K. E. (2017). Exploring the Pair Programming Process: Characteristics of Effective Collaboration. *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education* (pp. 507–512).

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.

Schmitt, L. J., & Weinberger, A. (2017). Collaborative learning on multi-touch interfaces: scaffolding elementary school students. *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning (CSCL)* (pp. 9–16).

Schultz, J. L., Wilson, J. R., & Hess, K. C. (2010). Team-Based Classroom Pedagogy Reframed: The Student Perspective. *American Journal of Business Education, 3*(7), 17–24.

Snyder, C., Biswas, G., Emara, M., Grover, S., & Conlin, L. (2019). Analyzing students' synergistic learning processes in physics and CT by collaborative discourse analysis. *Proceedings of the 13th International Conference on Computer Supported Collaborative Learning (CSCL)* (pp. 360–367).

Spada, H., Meier, A., Rummel, N., & Hauser, S. (2005). A new method to assess the quality of collaborative process in CSCL. *Proceedings of Computer Supported Collaborative Learning (CSCL)* (pp. 622–631).

Stahl, G. (2015). A decade of CSCL. *International Journal of Computer-Supported Collaborative Learning, 10*(4), 337–344.

Tsan, J., Lynch, C. F., & Boyer, K. E. (2018). "Alright, what do we need?": A study of young coders' collaborative dialogue. *International Journal of Child-Computer Interaction, 17*, 61–71.

Wegerif, R. (2011). Towards a dialogic theory of how children learn to think. *Thinking Skills and Creativity, 6*(3), 179–190.

Williams, L., & Kessler, R. R. (2003). *Pair Programming Illuminated*. Addison-Wesley Professional.

Acknowledgments

This material is based upon work supported by the National Science Foundation through grant DRL-1640141. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Teacher Support of Emergent Shared Regulation for Dynamic Collaborative Inquiry

Simona Pesaresi and Jianwei Zhang
spesaresi@albany.edu, jzhang1@albany.edu
University at Albany, State University of New York

Abstract: A collaborative inquiry approach to science learning has the potential to develop students' higher-order thinking skills, deep understandings and induce into productive dispositions. To realize this potential, students need to gain control over their knowledge advancement in terms of adapting collaborative goals and reinventing their inquiry practices as problems evolve temporally. To this purpose, we consider the teacher scaffolding moves in turning higher levels of agency to students and enhancing their shared regulation of the inquiry. Going beyond existing research on scaffolding, the study examines such moves to see how teacher support plays out to help the classroom community co-construct, adapt and expand collective inquiry structures and trajectories. We explore this role in the context of a classroom of fourth-graders investigating light using knowledge building pedagogy and technology. Qualitative analysis of classroom discourse and teacher reflection identifies patterns of interaction by which the teacher leverages students' emergent shared regulation.

Introduction

As the world enters a new era featuring rapid changes, extraordinary challenges, and radical social and technological transformation, educators face a heightened demand to cultivate new cultures of learning that prepare students for productive participation in the ever-changing environment. Aligned with this need, research in the learning sciences has investigated collaborative, inquiry-based practices by which students develop higher-order thinking, deep understandings, and productive dispositions and identities. However, existing research and classroom practices have focused on pre-structured inquiry and collaboration in which the goals, tasks, procedures, and group structures are set by the teacher (or designer). Within the emergent field of shared regulation of collaborative learning (Järvelä et al., 2016), future research needs to understand students' strategic adaptation of collaborative goals and processes in temporally evolving learning situations (Järvelä, Järvenoja & Malmberg, 2019). Thus, the current study explores shared regulation in more dynamic and transformative forms of collaborative inquiry that are critically needed for the new social contexts. The goal is to understand how the teacher works to enhance students' shared control in co-constructing their ever-evolving inquiry practices and trajectories in light of emergent problems and opportunities.

Authentic knowledge-creating practices by nature require adaptive processes to solve complex problems while pursuing emergent agendas. The recent developments of collaborative knowledge work in the real world have further shifted toward highly dynamic configurations to support evolving goals, flexible collaboration, cross-boundary idea contact, and distributed leadership (Hagel, Brown, & Davison, 2010; Sawyer, 2007). This vision drives classroom innovations using the knowledge building pedagogy in which students are expected to take collective cognitive responsibility for advancing their shared knowledge and to exert epistemic agency, that is, to assume control on setting knowledge goals and strategies, monitoring and evaluating progress in understanding (Scardamalia & Bereiter, 2006; Zhang et al., 2009). These expectations raise the question of how students can become independent in determining what they need to inquire, and how, in order to continually advance their knowledge. A way to address this problem is to consider the nature and function of the teacher role in turning higher levels of agency to students and enhancing the community's shared regulation of the inquiry processes. Past research theorized and investigated new teacher roles that focused on scaffolding in classroom interactions. The mentor participant structure described a teacher who is concerned with helping students growing into the subject matter and uses slot-like scaffolding and prompting to activate their cognitive engagement while retaining a high-level control of the inquiry process (Scardamalia and Bereiter, 1991; Tabak & Baumgartner, 2004). However, for students to develop epistemic agency, teachers should do more than asking stimulating questions that deepen students thinking. More symmetric participant structures have been identified in which the teacher articulates his discourse moves to help student formulate their questions and goals, build their rationale and monitor their understanding. Acting as a partner, the teacher investigates side-by-side with students, by modelling cultural tools of scientific inquiry (Tabak & Baumgartner, 2004) and supporting a dialogic discourse form in which students are encouraged to assume intellectual authority to deepen their understanding (Ford & Forman, 2015). Previous studies analyzed teacher support in discrete episodes of interaction. Further work is needed to connect the different units and timescales of analysis to show how discourse in the here and now is embedded in a history of interactions (Tabak & Baumgartner, 2004), and how it embodies the accrued knowledge of past practices that participants have gained jointly or separately (Mercer, 2008).

This study builds on previous findings on these new teacher roles and goes beyond, by giving a more explicit attention to the teacher support in the cross-temporal adaptation of collective inquiry structures. We draw on the emergent structuration framework proposed by Zhang and colleagues (2018) to develop an analysis of teacher scaffolding that captures the interplay between the historical and the dynamic dimensions of co-constructed inquiry structures. This framework suggests that the inquiry practice of a classroom community builds on existing structures (e.g. curriculum content, epistemic work practices, theories, models, etc.) emerged from joint previous activities or from the past experience of other communities within the larger school context to which participants have been exposed. These initial structures mediate and steer participants' actions and interactions giving emergence to a shared space of knowledge and new lines of inquiry. The emergent inquiry agenda, in turn, creates the conditions for further elaboration or modification of pre-existing structures to address the evolving knowledge goals and support related actions. Thus, the inquiry structures that are so dynamically co-constructed and in a continuous state of flux help students foster their epistemic agency for directing their inquiry and cognitive responsibility for advancing their community's knowledge. We position and interpret teacher moves in the context of both these historical and dynamic aspects of the inquiry trajectory. The purpose of the study is to see how the teacher dialogic support plays out as part of the community's co-constructing of inquiry structures to enhance its regulation of shared practices. To this aim, our analysis attends to the way that discourse, both talk-in-interaction and inner discourse, can be used to provide an accumulative and continuing frame (Mercer, 2008) to enable students participation on equal footing in the control of these practices. We explore the teacher role in a classroom of fourth-graders who investigated light using a knowledge building design and an open-ended inquiry approach.

Method

Classroom context

The data we analyzed are from a fourth-grade classroom at the Dr. Eric Jackman Institute of Child Study Laboratory School of Toronto participating in a three-month study of light. The unit of study was carried out using the knowledge building pedagogy and Knowledge Forum (Scardamalia & Bereiter, 2006), an online technological platform for writing, co-authoring and building-on others' notes through a set of scaffolds that help students develop, monitor and deepen their understanding. As part of the knowledge building curriculum and design principles, students investigated light by engaging in whole-classroom discussions, small-group practical activities, including experiments, demonstrations and reciprocal reading and by contributing their knowledge to the online discourse in Knowledge Forum (KF). They were encouraged to exert epistemic agency by identifying questions to investigate as knowledge goals, generate ideas and leverage empirical evidence or authoritative sources of knowledge to make progress in deepening their understanding. Guided by the principle of collective cognitive responsibility, students engaged in sustained efforts of improving knowledge and deepening problems of understanding by bringing their ideas to the collective space of classroom and online discourse and subjecting them to questioning and refinement. At the time of the study, the teacher had several years of experience in implementing knowledge building designs and principles. To conduct the inquiry unit on light, he adopted an opportunistic collaboration design, encouraging students to participate flexibly in the different inquiry areas identified by the classroom. He tried to not direct the study both in terms of problem goals and collaboration strategies, allowing for spontaneous grouping and re-grouping in relation to the evolving students' foci. He sat in circle with students during whole class talks or circulated among small groups to engage them in discussions about progress with their inquiry. This context, rather than in traditional lessons, grants students an initial advantage to exert agency and authority in the knowledge building process.

Data sources and analysis

The whole data set includes video recordings of 11 class episodes, online notes in Knowledge Forum (168 discussion notes and 48 personal reflection notes) and a teacher's reflection journal (17 daily entries). Observation and data collection followed a three-year long design-based research comprising three iterations of Knowledge Building Communities design (Zhang et al., 2009). Throughout this period, a trustful relationship developed between the principal investigator (second author) and the teacher leading the latter to progressively step in to participate in the design interventions. The analysis in this study was conducted holding awareness of this contextual background. In examining how the teacher scaffolding played out to support the community's construction and shared regulation of its inquiry we mainly considered two types of data sources: Excerpts of classroom interactions (names of participants have been replaced by pseudonyms) and the reflection journal. The latter documents a teacher practice conducted beyond classroom interactions for the purpose of supporting the ongoing inquiry. In this case, the process of scaffolding involves, first, observing and assessing an activity or situation; then, reflecting upon what observed and connecting it to the desired goals; and finally, responding to opportunities or constraints that emerged from the reflection by designing, planning or adapting subsequent actions (Bakker et al., 2015). Thus, we examined how the teacher support to build on and adapt shared inquiry structures was carried in his practice of reflection and his interaction with students.

Data were analyzed using the method of discourse analysis with a focus on the temporal dimension. In particular, we drew on the ethnographically grounded approach to discourse analysis proposed by Gee and Green (1998), which offers analytic guidelines for exploring what semiotic and sociocultural connections are proposed, recognized, or made relevant in discourse to ideas, people, practices and interactions located in the past or the future. We combined this approach with the multiple timescale analytic perspective of Lemke (2000), who suggests that every moment-to-moment action or activity be viewed as interdependent with one another and adding up to a longer timescale process that constrains what is more suitable or likely as consequent actions. We started with a broad-stroke exploratory analysis to look for patterns that displayed features of discourse moves as informed by the conceptual framework. We understood the purpose of a specific teacher move and how it allows the emergence of structures by moving back and forth through the sequence of episodes and trying to make sense of it as part of a dialogic, cohesive process (Mercer, 2008) generative of cross-temporal connections. We then identified a few meaningful illustrations of these patterns and unpacked the historical and dynamic features. We later discuss how these patterns introduce opportunities to students for taking on a regulating role in co-structuring their inquiry.

Findings

Structures leveraging connections between past, present and future

The teacher draws students' attention on discussion topics by linking these to memories of past shared knowledge and events as a way of justifying, legitimizing and providing the resources for orienting the future activity. In this way, he creates a meaning-making structure that helps students carry on actions that build into continuity.

Connecting the topic of light to prior students' knowledge embodied in artifacts

During the first classroom discussion episode at the beginning of the inquiry, the teacher showed notes written in a former view of Knowledge Forum when students were in Grade one. One of these notes, in particular, contained a theory about how animals' fur color adapts to light. What follows is the teacher's journal entry on this episode:

> Using the data-projector (…) we looked at the "Adaptive Weirdos" view in the old Grade one database created by the current Grade four students. The students enjoyed seeing their notes and illustrations. The last note we opened was Julien's. It contained a theory about grey fur "reflecting" light away from his creatures' eyes so that it can see better. This note generated an interesting discussion on how light responds to color. (…) It also happened to be snowing outside after a few weeks of very mild spring weather. The discussion progressed to snow and the color white. We asked the question if there was a reason why snow was white. The students had many theories to share (…) We posted the theories on the board.

From this reflection, it appears that an inquiry focus emerged and evolved spontaneously in response to students' renewed interest in their Grade one notes and also to the concomitant snowing event that generated meaningful connections between theories of animal adaptation, color and the new topic of light. The illustration of old notes was an intentional action designed to elicit such connections and use these as foundations for the future inquiry. The teacher built an initial interactional frame by activating prior knowledge and offering students the opportunity to develop an inquiry thread in connection to their shared previous knowledge. He did not determine upfront how students would link their past theories to new questions about light. Instead, during the interaction, students dynamically leveraged their previous theories as well as the occurrence of the snow fall to raise questions and offer new ideas as new inquiry lines. A further entry in his journal provides evidence for how the teacher positioned his role in relation to the inquiry work, recognizing that the emergence of future inquiry directions built on students' initial interests, pre-existing knowledge and contextual circumstances:

> In terms of light, I have no preconceived notions as to how the study will take place other than that the students have identified it as a topic of interest and that our introduction to it seems to have originated from an archived Grade one note on adaptation, and the coincidence of a snowy spring day.

The emergent collective focus – i.e. study of the light-color relationship – became an inquiry structure that influenced students in identifying the problems they wanted to investigate. This focus was embodied in a new view in Knowledge Forum that the teacher created as reported in his journal following the first whole classroom episode:

> I created a new view in the database called "Grey Fur and White Snow". I copied portions of Julien's Grade one note as background of the new view. A brief paragraph explains the origin of the view.

Figure 1 shows the new view "Grey Fur and White Snow" with the note from Grade one inserted in it. The title chosen by the teacher for this view together with the inscribed former note constitute a symbolic trace connecting past students' knowledge to the present focus of discussion and the generation of further questions to investigate. The view thus embodies the historical and dynamic aspects of a structure functional to sustain students' inquiry. Its design reflects the teacher's intention to involve students in a cumulative and progressive inquiry journey highlighting the continuity of their learning experience.

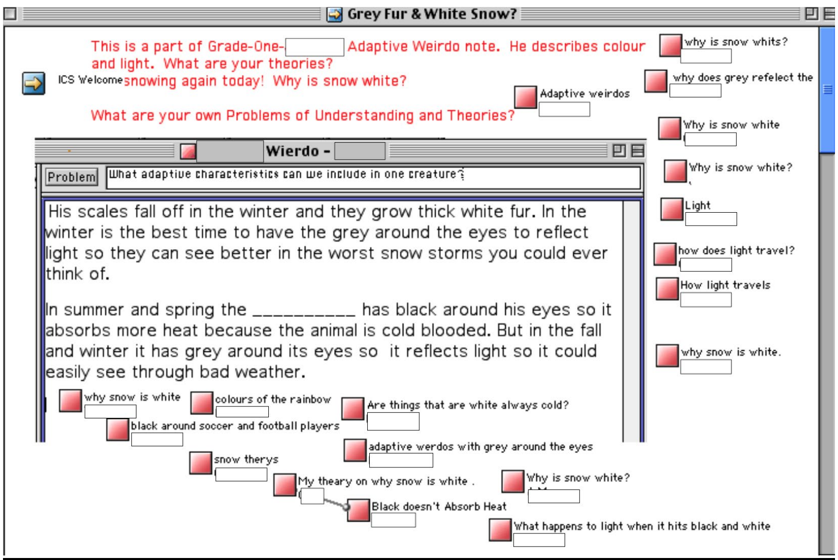


Figure 1: Knowledge Forum view called "Grey Fur and White Snow".

Leveraging students' accumulated experience with Knowledge Forum to support a reflective use of the tool.

The teacher reflected further on his role in the inquiry by analyzing his and students' experience of knowledge building pedagogy. He noticed that not all Grade four students in the previous years were keen on working with Knowledge Forum and that their feeling about the technology would influence their preference for Grade five teachers (knowing that one teacher would continue teaching with Knowledge Forum and the other would not). This reflection induced him to adapt his plan for the current year by letting students use the platform spontaneously as reported in the teacher's journal:

> My focus this year, is to let the students lead the direction of the study. (…) I noticed in the past three years, students either graduated from Grade four with a love or hate for KF. This feeling would direct their wish for which Grade 5 teacher to have the following year (…). This year, I wanted to downplay the use of KF, using it when only the students felt it was appropriate. Funny enough, this unique group of students from the onset of school this year, have directed the knowledge building to occur on the database. The original study of Greek Mythology was not intended to take place on the database but the students felt it would be the best place to store and share the information they were gathering about each God. (…) It was a student who suggested a math question be placed in the database to ensure its life span. Prior to this year, I had not attempted to "do math" on KF. This group of students was the first to challenge me in my description of the portfolio views as a "private place" in the database. They asked me why the knowledge building could not continue at this level and questioned if there really can be (or

> should be) a "private place" on KF. We have attempted to use the database as a means of peer editing our paragraphs, using scaffolds to comment on writing rather than frame it.

In this excerpt of his reflection journal, the teacher reported several instances of the way students perceived Knowledge Forum and how they adapted, reshaped and reinterpreted the existing tool structures in flexible ways over time to support their activity. They extended the scope of the database to support the study of other curriculum units such as Mythology and Math. They challenged the private character of portfolio notes (personal notes where students write their own evolving understanding of the problems investigated) and suggested that they could be the place for others to contribute knowledge and, finally, they explored other scaffolds to enable peers editing of someone's theories or ideas in online notes. We see how the teacher stepped back from his initial plan and envisaged shifting the regulating role to the students as a result of a temporal analysis. He invoked memories of other students experience in previous years and related these to the students' enhanced competence with the tool accumulated since the beginning of the school year. The teacher reflection displayed internal persuasiveness, a characteristic of dialogic discourse (Bakhtin, 1981), when he related two competing personal considerations of how students communities positioned themselves towards Knowledge Forum over time.

Co-designing open and flexible inquiry and collaboration strategies built on past teacher's experience

As a second reflection step, the analysis of the different approaches to knowledge building taken in the previous three years informed the teacher on the potential of a flexible inquiry structure in terms of having students decide what problems to investigate and with whom working:

> In analyzing the data from the past three years, it seems true (contrary to my original hypothesis), that there has been progress each year in the significant change from pre-test to post-test, significantly more activity each year, and even the portfolio notes themselves seem to suggest that the students have been demonstrating epistemic agency. Building from last's years success – a year with students working organically in any study group they were interested in – with less structure, I think this year, I will continue to test the boundaries by consciously trying to not influence the direction of the study. (…).

Teacher's reflection about how students grew into the knowledge building activity and out of external control in directing their inquiry paved the way for giving them a higher level of agency in the knowledge building process. The teacher shared his reflection with students in a subsequent classroom discussion:

1	Teacher:	I know how we started looking at this [study]. We looked at a weirdo note, a weirdo kinda day connected to light that you identified at the beginning of the year. You have written problem notes and problems of understanding.
2	Sylvie:	If we want to study light, we split off into groups, one "rainbows", one "prisms", then close to end of study, we get together and tell them what we did and people won't only know about what they studied.
3	Chloe:	I remember in Grade 2, when studying the ancient Egypt, each group studied something different and groups did a presentation.
4	Sienna:	We should split into groups, do a presentation and an experiment to show what we have learned.
5	Teacher:	The first year we did Knowledge Forum, we had students work in groups. There was a lot of knowledge building only within the small groups (…). Students knew nothing of the other groups, they only learned about one thing instead of learning about six things. The second year, students worked in groups, but always knew about what the other groups were doing because were always reading and writing in all the views. (…). The third year, we divided questions into views, nobody was assigned to a view, you could work into any views that you wanted. I thought it wasn't going to work, because students need to work in one group to get a deep understanding. (…) Data [from the survey] showed that students in the third year understood problems best [because] they were given the option of choosing the view they wanted to study in.
6	Sylvie:	We could do that. Maybe we could be each in one group and every week each group could be working on another thing.
7	Teacher:	Knowledge building is about working with people who want to work...not just friends.

8 Greg: We can work with people who share the topic you want to work on.
9 Teacher: These are all your questions; I don't know which direction this will go in.

The teacher did not explicitly introduce the topic of how carrying out the inquiry nor did he present it as a plan he decided on upfront. Instead, he organized his discourse about the matter stepwise. He initially pointed students' attention to what the classroom as a whole accomplished so far and how past steps enabled them to identify problems to investigate (line 1). This introduction triggered a discussion on how to work on these problems, what sort of groups should be formed and how they should share their inquiry work (lines 2-4). The discussion yielded a symmetric interaction during which some students contributed ideas on the working strategy. Taking up these ideas, in line 5 the teacher recalled his past experience reporting on different approaches of organizing the inquiry work in other classrooms and highlighting how each approach affected learning. He provided factual elements, shared initial beliefs and weighed them against the evidence. By making visible in his reasoning how he builds connections to and critically analyzes past structures, the teacher modeled an interpretive framework to inform or guide future decisions. He thus offered students an opportunity to position with agency in appropriating and adapting previous strategies. His account may have probably influenced or constrained students' orientations but did not determine how or with whom students eventually would work in the classroom. The subsequent turns (lines 6-9) give evidence of how invoking of pre-existing strategies dynamically mediated students' interaction and built the foundations for designing future actions without relying on the direction of the teacher. We see students offering suggestions for group forming strategies with the teacher simply reminding them of the principle guiding collaborative work in knowledge building. Then, with a quick twist he connected the conversation back to the focal themes of the investigation and displayed a downgraded epistemic stance relative to students about the direction of the study. In this way, he subtly showed an effort to turn over to students the regulating role of the inquiry focus.

Structures leveraging connections across classroom communities

The teacher reflects on and creates opportunities to develop connections with the knowledge building work of other classrooms communities so that knowledge generated in other space-temporal contexts can be leveraged by students to orient their choices.

Past visual exposure to other class communities knowledge artifacts shaped students' choices

In the following excerpt, the teacher reflected on how students past experience and context may inform and steer students' choices. He first remarked that their Grade one work oriented their approach to investigate light and gave emergence to specific lines of inquiry, as discussed earlier. He also noticed how students' exposure to knowledge artifacts developed by former Grade four classroom and exhibited in a common space influenced their goals in terms of units they wanted to study during the school year.

> I think that their Grade one work on Adaptation is influencing their theories and how they are looking at light and color. This is a new approach to light. The students had identified "Light" as a topic they wanted to research when asked in the beginning of the year along with Mythology, the Middle Ages, Shakespeare, the Holocaust. As the Grade three room shares the same hallway, there is exposure to Grade four exhibited work, and thus I think the children picked many of the units they knew the Grade fours had studied last year: Medieval Times and Light.

Acting as boundary texts (Lemke, 2000), these artifacts allowed students to interact with their prior knowledge and with practices developed by others and to leverage these resources in orienting their interests. The teacher recognized the agentic role of former Grade four community in providing guidance to students. The work exhibited in the hallway acted as an initial structure, an area of content of the curriculum (Zhang et al., 2018) that students adopted to outline their plan of research topics. Yet, as we indicated earlier, they engaged with the selected topics flexibly by reinterpreting the use of Knowledge Forum for supporting the inquiry and proposing new ways to share and contribute to the knowledge building process.

Creating opportunities to interact with knowledge accumulated by other classes in Knowledge Forum

As the inquiry evolved, students identified 'shadows' as a focal theme to investigate. Thomas noticed that a view in Knowledge Forum called 'Shadows' was already created by Grade five/six students and proposed the class the use of the existing view. His suggestion opened up a discussion among students on the possible future implications of taking on this opportunity:

> We looked at the clutter of notes in the original view: Grey fur & White Snow. The students suggested we organize the notes into new views. (…). Next we attempted to create titles for the new views. Once again we have come up with 6 views! (…) One of the suggested views was "Shadows". I told the class that there already is a view on the database created by the 5/6 class called shadows and I asked them if we should simply use this view. Some students felt that was not knowledge building "we would have their answers before we developed our questions". Others felt this might upset the 5/6 students and we should ask permission. Others felt we should continue where the 5/6 students left off.

As evidenced from students' responses reported in the teacher's journal, some students were concerned about emotional and practical aspects (i.e. how Grade 5/6 students' reactions), whereas others exhibited higher-order thinking about what type knowledge they would build if it leveraged existing work by others. Students were not only projecting the consequences of taking a certain strategy but also reflecting on the educational point of adopting an existing structure as a knowledge practice, "some students felt that was not knowledge building… ." The preceding excerpt also indicates a substantial progress in organizing and structuring the inquiry work. Starting with the initial view "Grey Fur & White Snow" that the teacher created in Knowledge Forum, students contributed several notes reflecting knowledge advancement in the form of new theories and problems of understanding. The teacher was intentional in showing students the groupings of notes he discovered in the existing view (he had planned this action as reported in an earlier note) as an occasion for reshaping the current structure and creating additional views that would better serve the needs of the inquiry. We see how the teacher did not give explicit guidelines in this regard. Instead, he elicited students' reaction by showing them the original view and conferred them the authority to organize notes into new views and propose new titles.

Following the classroom discussion, the teacher reported his own reflection about the opportunity of connecting to and using a shared inquiry space of inquiry across communities:

> What would be the ramifications of working in a view created by another class? My gut feeling is that this would be great opportunity to examine knowledge building outside of the classroom. (…) The Grade 5/6 class was focusing on sun rays and the tilt of the earth. This is not yet an area of interest for the Grade four students. I wonder what the impact will be for the Grade fours and the Grade 5/6?

These considerations reflected a teacher's contemplative stance towards students' interaction with the inquiry work initiated by other students. Again, he harnessed existing structures to project future directions of the inquiry and imagine new practices of working creatively with knowledge across classroom boundaries, such as building knowledge using inquiry spaces shared by more communities. He did not constrained students to use these existing structures but sought to create an opportunity in which they could leverage them dynamically to orient their inquiry focus.

Discussion

We have identified two main patterns that characterize the way the teacher role played out to support the students' community in regulating its knowledge building processes. The analysis of these patterns showed how the teacher's discourse strategically draws on histories of events and relationships to project future actions and directions. The teacher was intentional in reflecting on his practices as a long-term trajectory and discovering students' accumulated experience and ideas as means to shape their current inquiry directions and ways of inquiring and collaborating. He was able to discern emerging directions from individual contributions and bring what he had noticed to collective meetings to reshape the community's inquiry foci and guide the subsequent work of the classroom community as reflected in the Knowledge Forum views. We saw how he established a symmetric relationship with students, positioning himself on "the same plane of participation as the students" (Tabak & Baumgartner, 2004, p. 420) and allowing them to take control over the problem at hand, as in the discussion about group work strategies. He empowered students by positioning them as authoritative sources of knowledge as in the case of their Grade one notes or by allowing them to flexibly use Knowledge Forum or an existing view created by other students. However, the teacher did more than simply acting as a partner. A partner-teacher is mainly concerned with inducing students into asking their own questions and monitoring their understanding to regulate their learning. In this study, there was a concern with helping the community leverage shared knowledge and make a reflective and flexible use of existing structures (Zhang et al., 2018), in connection with the ongoing agenda, to organize and guide their inquiry. In this sense, the teacher was rather playing the role of a participatory co-designer, who elicited, reflected on and interpreted relevant past knowledge to mediate ongoing interactions and project future directions and possible students' practices. The creation of a view in Knowledge Forum invoking a former Grade one note and outlining the new inquiry focus emerged from the classroom interaction is illustrative of this design approach. Teaching to co-construct shared structures that enhance students' regulation requires, as seen in this study, a systematic effort to engage in a reflective practice whereby teachers attend to prior experience, existing knowledge resources and structures in the larger school context and envision possibilities for students to harness them. Pursuing this effort enables them not only to provide cohesiveness and continuity to students experience but also to empower them to appropriate and adapt proposed working practices.

In our future work, we need to develop a full case picture that examines other forms of building and eliciting temporal connections, how they are constituted through talk or other discursive spaces and how they contribute to foster students' shared regulating role. We also need to learn more about what teaching as co-designing means in the context of inquiry-based collaborative learning, showing how this role is embedded in a larger and longer view spanning multiple classroom communities and timescales.

References

Bakhtin, M. M. (1981). The dialogic imagination. Four essays. University of Texas Press.

Bakker, A., Smit, J., & Wegerif, R. (2015). Scaffolding and dialogic teaching in mathematics education: Introduction and review. ZDM, 47(7), 1047–1065. https://doi.org/10.1007/s11858-015-0738-8

Ford, M. J., & Forman, E. A. (2015). Uncertainty and scientific progress in classroom dialogue. In L. B. Resnick & C. S. C. Asterhan (Eds.), Socializing intelligence through academic talk and dialogue (pp. 143–156). AERA.

Gee, J. P., & Green, J. L. (1998). Discourse Analysis, Learning, and Social Practice: A Methodological Study. Review of Research in Education, 23, 119. https://doi.org/10.2307/1167289

Hagel, J., Brown, J. S., & Davison, Lang. (2010). *The Power of Pull: How Small Moves, Smartly Made, Can Set Big Things in Motion*. New York: Basic Books.

Järvelä, S., Järvenoja H., & Malmberg, J. (2019). Capturing the dynamic and cyclical nature of regulation: Methodological progress in understanding socially shared regulation in learning. *International Journal of Computer-Supported Collaborative Learning, 14*, 425-441. https://doi.org/10.1007/s11412-019-09313-2.

Järvelä, S., Kirschner, P. A., Hadwin, A., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual- and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning, 11*(3), 263-280. https://doi.org/10.1007/s11412-016-9238-2.

Lemke, J. L. (2000). Across the scales of time: Artifacts, activities, and meanings in ecosocial systems. Mind, Culture, and Activity, 7(4), 273–290.

Mercer, N. (2008). The Seeds of Time: Why Classroom Dialogue Needs a Temporal Analysis. Journal of the Learning Sciences, 17(1), 33–59. https://doi.org/10.1080/10508400701793182

Sawyer, R. K. (2007). *Group genius: The creative power of collaboration*. New York, NY: Cambridge University Press.

Scardamalia, M., & Bereiter, C. (2006). Knowledge Building: Theory, Pedagogy, and Technology. In K. Sawyer (Ed.), Kel (pp. 97–118). Cambridge University Press.

Scardamalia, M., & Bereiter, C. (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.

Tabak, I., & Baumgartner, E. (2004). The teacher as partner: Exploring participant structures, symmetry, and identity work in scaffolding. Cognition and Instruction, 22(4), 393–429.

Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for Collective Cognitive Responsibility in Knowledge Building Communities. *Journal of the Learning Sciences*.

Zhang, J., Tao, D., Chen, M.-H., Sun, Y., Judson, D., & Naqvi, S. (2018). Co-Organizing the Collective Journey of Inquiry with Idea Thread Mapper. *Journal of the Learning Sciences*, *27*(3), 390–430. https://doi.org/10.1080/10508406.2018.1444992

Acknowledgements

This research was sponsored by U.S. National Science Foundation (#1441479). We thank our collaborating teacher and students for their creative classroom work enabling this research.

International Society of the Learning Sciences

Designing for Equitable Participation in Collaborative Game-Based Learning Environments

Karyn Housh, Asmalina Saleh, Tanner Phillips, Cindy E. Hmelo-Silver, Krista Glazewski
khoush@iu.edu, asmsaleh@iu.edu, tanphill@iu.edu, chmelosi@indiana.edu, glaze@indiana.edu,
Indiana University

Seung Lee, Bradford Mott, James Lester
sylee@ncsu.edu, bwmott@ncsu.edu, lester@ncsu.edu
North Carolina State University

Abstract: With the growing popularity of game-based learning, researchers should take steps to ensure that our designed technology-enhanced environments reflect our desire to implement equitable learning environments for all our students. Does our implementation of technology to motivate and encourage learning, at times result in some of our students having less of an opportunity to engage in the learning activity? This study focuses on a STEM-based game-based learning environment designed to facilitate equitable participation of learners. Through the analysis of game interaction log files, we explore time series plots and determine patterns of student participation. We highlight our findings using a case-study approach in which we focus on the interaction of a middle-school group collaborative activity as they engage in solving a problem embedded in a game-based learning environment.

Introduction

Game-based learning environments provide a rich avenue to support collaborative inquiry learning and in turn, can provide key insights into designing and analyzing group processes and interactions (Mislevy et al., 2014; Rupp et al., 2010). Although empirical research has focused on how these learning environments can increase motivational, cognitive, and affective outcomes (Connolly et al., 2012), there has been less attention on how to intentionally design for equitable participation. Game-based learning environments are often viewed as highly engaging for students since they represent a means through which rich learning activities can be accomplished with sustained student participation. However, there remains a question of equitable participation within these learning environments, as they may inadvertently allow those more experienced with gaming a greater advantage, or it may be presented as more attractive to particular groups of students (Buffum et al., 2016). In our work, equitable participation refers to the provision of equal opportunities to participate in classroom activities (Rasooli et al. 2018). Several authors have espoused similar sentiments, advocating for more equitable forms of participation to better reflect the capabilities of our learners (Poehner, 2011). However, educators and researchers alike often grapple with the implementation of these considerations. This is evident in game design where aspects of equity and inclusion should be addressed.

This paper seeks to explore these issues by examining the following research question: *How can we design for equitable opportunities to participate in a collaborative game-based learning environment?* We draw on an activity theory framework to guide our design and analysis of ECOJOURNEYS, a collaborative game-based learning environment designed for middle school ecosystems learning. We highlight how activity theory can be used to design equitable tasks and mediators that act as active modifiers that aid in creating new actionable pathways to learning (Poehner, 2011). Active modifiers are tools which serve to actively enhance students' understanding. A brainstorming board with rules for participation (designed to require each group member's participation to vote) and in-game chat (collaborative discussion of tasks where students are free to participate at will), both serve as active modifiers in this study. Drawing on students' in-game interactions as captured in log files, we adopted a social learning analytics approach and explored time series plots to discern patterns in how groups of students interacted in the designed tasks. Using these plots, we used an instrumental case study approach (Stake, 2008) to the extent to which designed opportunities learning supported student participation.

Equitable participation and activity theory

Classroom activities that adopt sociocultural perspective should involve participatory tasks and authentic inquiry (Danish & Gresalfi, 2018). Additionally, classroom activities that attend to equitable practices should focus on dimensions such as procedural elements (i.e., practices that give rise to equitable outcomes) and interactional justice (Rasooli et al., 2018). However, the question remains, how do we design for equitable opportunities to participate in these collaborative game-based learning environments?

Our work centers on the concept of activity theory and assumes equity as access to mediators (Poehner, 2011). Collaborative inquiry learning is assumed to be a joint cooperative activity wherein instruction and learning co-occurs through mediational means i.e., the resources used to construct knowledge (Holzman, 2018). Additionally, these frameworks should include multiple levels of activities and participation building towards the formulation of an argument. Thus, we draw on activity theory as a theoretical framework to designing and analyzing equitable tasks. In activity theory, the object or the collective goal plays a vital role in mediating and organizing interactions (Engeström, 1987). In working towards a collective object, an individual's activity is artifact-mediated, or influenced by tools, the division of labor and rules associated with each community. These mediators are historically and culturally shaped and transform the way that individuals can perform tasks. In the design of ECOJOURNEYS, the object or designed goal was to support students' collaborative inquiry. In this form of inquiry, students are expected to share their thoughts, and reflect on other students' ideas (Liu et al., 2016). However, this can be problematic when students are unaware of the steps in the inquiry process (Quintana et al., 2004). In our work, mediators are meant to encourage actions that students may not otherwise engage in and we attended to three elements: 1) mediational tools, 2) division of labor, and 3) rules that guide the inquiry process. Below, we unpack the design of these mediators.

How mediators were embodied in the design

In ECOJOURNEYS, students participated in a cultural exchange program to learn about tilapia farming in the Philippines. There, the locals requested students to investigate why tilapia at the hatchery fell sick. Students worked in groups of four and engaged in two inquiry cycles that consisted of (1) collecting data from the in-game environment and talking to in-game stakeholders, and (2) sharing and negotiating their ideas with one another using a collaborative space (see Figure 1). The brainstorming board was an in-game collaborative space providing structure for students to share their observations and to reach a shared understanding about the problem they are facing (Saleh et al., 2019).

Mediational tools

As students explored the game-based learning environment, individual students collected information in the form of notes. After collecting these notes, students collaboratively used the brainstorming board to share notes. The board highlighted the components that tilapia fish need to survive (e.g., temperature, air, etc.). The main task was for students to move the notes onto the board and determine the extent to which the information in the note was relevant to the component. After moving the notes, students clicked on their peers' notes to evaluate the relevance of the note. A visual indicator denoted when students reached agreement on whether a note was relevant to the component in the system (i.e., green when all students agreed, red when one disagrees, orange for default, see Figure 1). Students also used an in-game chat to negotiate their ideas, especially if there was disagreement over where the notes should be placed. The brainstorming board thus actively modified how students could participate equitably in collaborative inquiry, by encouraging multiple opportunities for students to 1) share notes and engage with the information, 2) demonstrate their thinking about the relevance of the notes and negotiate with their peers.

Rules

We also designed several rules that supported the collaborative inquiry process. First, inquiry milestones consisted of individual data collection and collaborative sensemaking. Collaborative sensemaking at the board was triggered after all group members completed data collection. Second, all members were prompted to share their ideas during the process. This task was formalized as notes that students collected, but students were also encouraged to share information as they explored the environment. Third, the placement of the notes was also a crucial step in the process. By placing notes on the board, students demonstrated that the note was relevant to the component on the board. Fourth, students voted to indicate how the notes may or may not be relevant to the component. Finally, students were also required to come to a consensus on how the information should be sorted and whether they were ready to move on to the next phase of the inquiry. These rules supported equitable participation because each student was expected to engage in these actions in the collaborative game-based learning environment.

Division of labor

Each student had the role of sharing their individually collected notes and evaluating each other's ideas. This equal distribution of roles ensured that each student had explicit ways that they can contribute to collaborative inquiry. We also accounted for the role of the facilitator as part of the community. The facilitator supported the inquiry process by prompting for contributions in chat *and* to ensure that the tool-based interactions also occur.

Facilitators provided support by marking information and questioning students, by asking for evidence and inviting participation. Additionally, facilitators and students engaged in socially shared regulation of learning, discussing norms for collaboration, deciding goals and planning actions related to solving the problem. These roles, however, were not explicitly designed interactions, but expected to emerge in collaborative discourse.

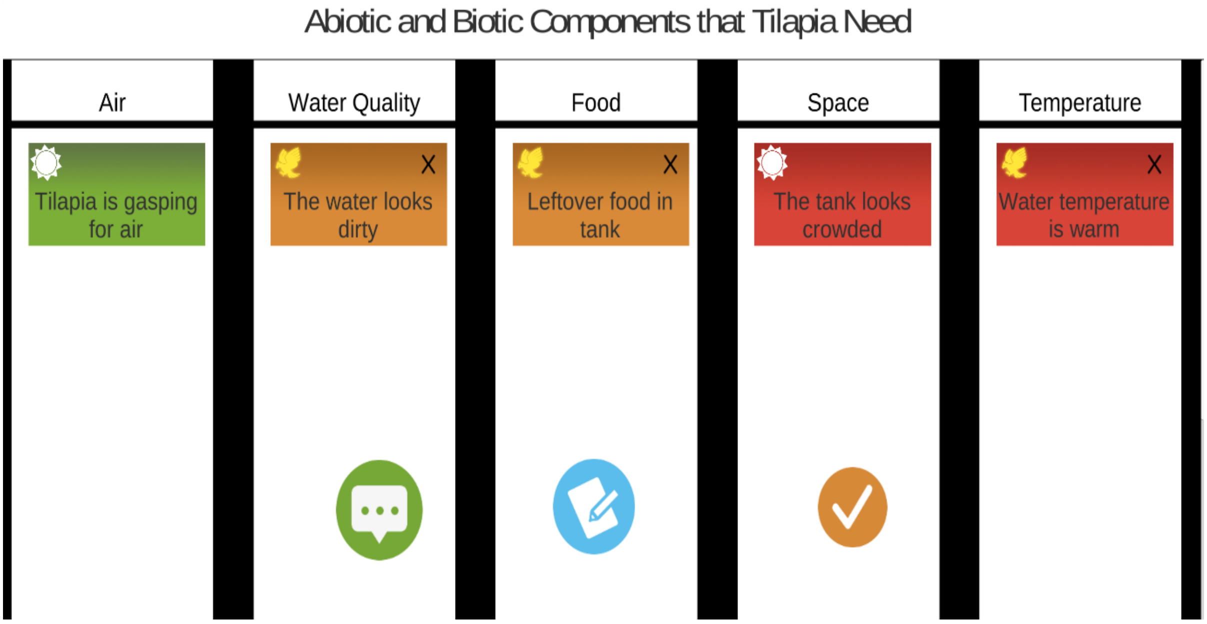


Figure 1. Overview of the brainstorming board

Methods

We drew on data from a classroom study with 29 students ranging from 11 to 13 years old (10 females, 19 males). Student demographics were as follows: 4 students identified as African/Americans, 4 as multi-racial, 2 as Asians/Pacific Islanders, 1 as Hispanic/Latinx, 1 as Native American/American Indian, and 17 as White. Students worked in groups of four and played ECOJOURNEYS in place of their science lesson. They participated in two sessions which were two-hours long. Students first took a survey where they shared demographic information and answered questions about familiarity with games. In the implementation, students engaged in two inquiry cycles of the game. Each group was assigned a facilitator who played the role of wizard or helper and supported group inquiry using the in-game chat. All students' in-game interactions were logged. In the last session, students were asked to draw a model explaining why the problem was occurring. Students also took a pre- and post-test that focused on their ecosystems understanding.

To understand the nature of participation, we examined the log files to focus on individual actions across time and in relation with other students. Individual in-game actions while using the board included 1) creating notes, 2) moving the notes, 3) voting on the notes, and 4) contribution to chat. These were used as indicators of equitable participation across students. To understand tool use, we aggregated group interactions at the board and chat, obtained frequency counts of individual game actions using the notes and examined the amount of time that students spent on reading notes, and quality of contributions to the in-game chat. To understand the division of labor among students and (in)equitable interactions among the different groups (i.e., rules), we generated time series plots. The time series plots feature the proportion of interactions contributed by each student over 50 events (i.e., 1 unit of X is 50 events). If student A contributed 20 of the last 50 interactions, their contribution value for that event index would be 40% (i.e., Y axis). We created two plots for each group: a plot of chat contributions, and a plot of board contributions. We qualitatively inspected these plots and engaged in qualitative case study analysis to further examine these patterns (Stake, 2008). Additionally, we also reviewed student chat utterances to determine which were off- or on-task. On-task utterances referred to instances when students discussed topics related to the science content or game-based learning environment whereas off task utterances were categorized based on whether students were socializing and discussing topics other than science or tasks related to the game-based learning environment.

Results

Paired t-test comparing the pre-test (M = 27, SD = 3.6) and post-test (M = 28.4, SD = 3.7) scores demonstrated that there were significant learning gains, $t(26) = 2.13$, p =.04. To understand how the design of the learning environment supported equitable participation, we present an overview of student interaction with our designed tools. We highlight how students use the notes at the brainstorming board and the chat tool, and then present the distribution of students' participation across the brainstorming board and chat activities (i.e., division of labor). Finally, the average frequency counts of student actions with creating, moving, and voting on the notes indicate whether the designed rules for interactions supported equitable student actions at the board. Table 1 provides summary statistics for each group's interactions at the brainstorming board, and contributions to chat.

Table 1; Summary statistics for each group.

Group	Total board actions	Total mins on notes	No. of created notes	No. of moved notes	Total count of votes	Student chat contributions
1	385	53	28	93	264	896
2	509	68	46	133	330	665
3	276	49	27	56	193	457
4	358	67	33	85	240	514
5	335	66	23	73	239	473
6	238	26	24	55	159	277
7	155	57	19	31	105	482
Mean	322	55	29	75	219	538
SD	113.6	14.8	8.85	32.9	73.6	194.5

How does tool use differ across the groups?

In terms of students' activities at the brainstorming board, summary statistics indicate that groups had an average of 322 actions at the brainstorming board, with group 2 recording the highest contributions, and group 7 the lowest. Group 2 similarly spent the highest amount of time on the notes. In terms of chat use, groups contributed an average of 538 lines, with group 1 recording the highest contributions to chat, and group 6 with the lowest. As we will demonstrate in our analysis later, the multiple ways of interacting with the designed tools indicate that there may be more opportunities to participate, thereby supporting equitable participation among students.

How did the designed rules influence student interactions at the board?

To better understand how the students participated in their groups, we inspected the time series plots for all groups as they interacted at the board and used the in-game chat. Because of space constraints and to ground our findings, we showcase the results from one team to provide a rich description of the findings. Group 2 was selected for this case study because (a) students' pattern of board use was relatively similar and had the highest amount of board interactions but (b) were diverse in terms of chat use, demographic data, and video game experience (see Table 3). The students in the group also scored in the lower range in their pre-test (see Table 1). The contrasting profiles in how the students in the group engaged in commercial video games and in the game-based learning environment was also another reason why the group was selected. Moreover, the students' interactions in the brainstorming and chat activities provided a rich illustration of how different students participated and how the mediators did or did not affect students' actions.

Table 2: Demographic data of Group 2 members

Name (Pseudo nyms)	Age	Gender	Race	Weekly hours gaming	Time spent on in-game notes (mins)	Chat	Pre-test score	Post-test score
Jacob	11	Male	White	10-20	9.6	17%	21	29
Olivia	11	Female	White	3-5	20.5	3%	26	26
Ethan	11	Male	White	5-10	3.2	65%	28	30
Rakesha	12	Female	African-American / Black	0-2	27.0	17%	25	30

Figure 2 illustrates the pattern of interactions at the brainstorming board among students in Group 2. The blue vertical line indicates the initial use of the board, whereas the yellow vertical line represents the end of the first brainstorming session. Each participant is represented by a different color horizontal line within the plot. The activities captured by use of the board include creating or placing a note on the board, moving a note from the board, voting on the relevance of the notes to the associated components (e.g., air, temperature, food, water quality etc. (see Figure. 1)) and voting on whether the irrelevant component should be removed as a possible explanation for the tilapia being sick.

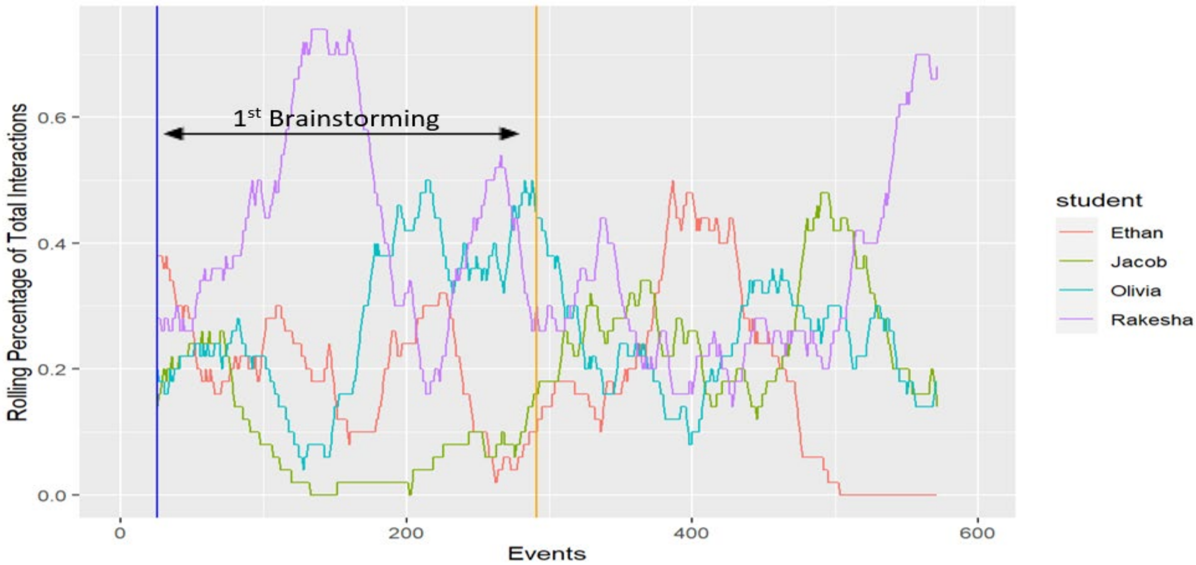


Figure 2. Group 2's use of the brainstorming board

In the first brainstorming session (i.e., after the blue line to about 250 events), Rakesha (pink line) appears to be the most active on the board. However, during the second session all students display approximately the same level of participation. This plot is representative of all groups' use of the board, there may be a few students who are more active in the first session. However, there is relatively similar board use in the second session across all groups. The difference observed between the two sessions is likely because all students collected the same five notes during their first exploration and students like Rakesha, who were quick to place items on the board tended to be more active. Moreover, once these notes are placed on the board, only the owner, in this case, Rakesha, will be able to move the notes. In Figure 2, student interactions during the second brainstorming session were relatively similar, apart from Ethan (in red). However, despite his late start, Ethan's actions mirror those of his peers (i.e., between 300-400 events) in their earlier interactions with the board (i.e., increase of actions before trending down). Although Ethan's actions reduce after the 400-event mark before logging out, the rest of his peers continue to engage with the board, repeating the pattern of upward and downward trend. The pattern of increased and decreased activity at the brainstorming board is likely triggered by the design features that are logged as these events, 1) creating the notes on the board, 2) the number of votes recorded, 3) placement of the notes, 4) consensus or lack of among the group, and 5) topics of discussion in the chat. Although we highlighted group 2's plot, plots for all groups depicted comparable patterns and symmetry across individual student's interactions at the board.

How was labor distributed among students across the chat and board activities?

These findings, however, are a contrast to the use of the chat feature of the game. Figure 3 represents student frequency and use of the chat feature with ECOJOURNEYS. Just as Figure 2, the blue and yellow vertical lines indicate the initial and subsequent use of the board. Based on the observed patterns in Figure 3, student participation in group 2 varied in frequency for both sessions of playing ECOJOURNEYS. Ethan appeared to be the most active in the chat for both whiteboard sessions. In both instances, Ethan, a self-identified white male who describes himself as a frequent video gamer, participates in the chat the most. In contrast, the student that contributes the least to chat is Olivia, a self-identified white female who plays video games occasionally. Comparatively, Rakesha, an African American female, who rarely plays video games and Jacob, another male student, had somewhat moderate and similar contributions to chat. Notably, both Rakesha and Jacob had the highest gains in their pre-post test scores (5- and 8-point gains), whereas Olivia maintained her score, and Ethan scored 2 points higher in the post-test.

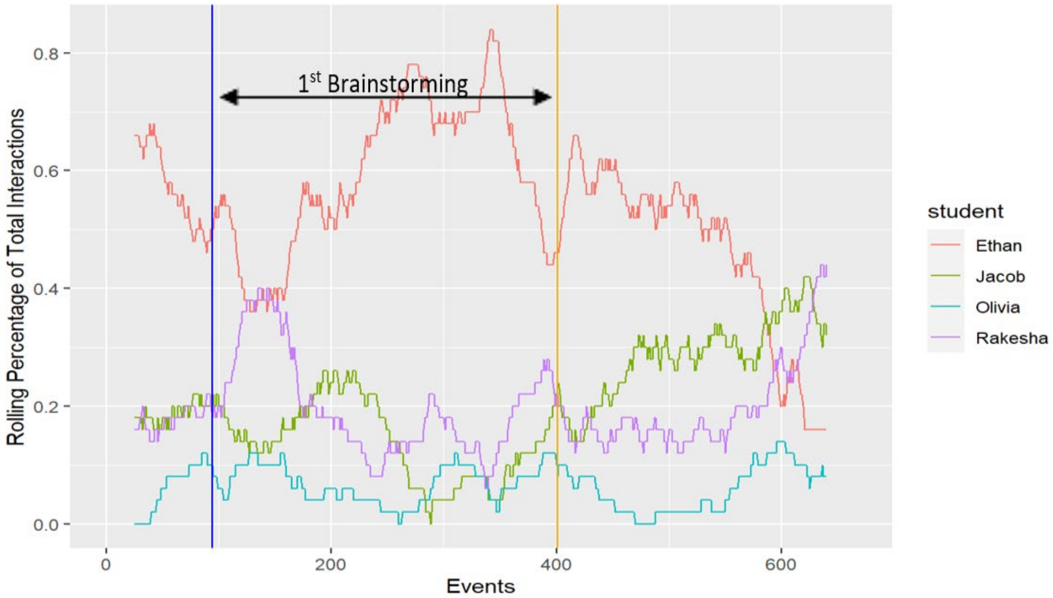


Figure 3. Group 2's contribution to chat when using the brainstorming board

Students' chat statistics also mirror the students' participation across time, corroborating students' quantitative contributions. However, when analyzing students' contributions, Ethan's utterances were on task 73% (i.e., content and game-oriented talk) and 27% off-task of the times, whereas Olivia and Rakesha were on task for all their contributions. Jacob's contribution on the other hand, was approximately distributed equally between on-task and off-task utterances. To illustrate students' conversations, consider their contributions to talk in excerpt 1 below.

Table 3: Group 2's in-game discussion in chat about water quality

	Time	**User**	**In-game chat**
1	12:56:44.8	Jacob	so water quality is pretty good i think
2	12:56:59.7	Jacob	i dont really think theres any problems
3	12:57:05.1	Ethan	dude,
4	12:57:06.9	Ethan	last time
5	12:57:08.1	Ethan	remember
6	12:57:14.4	Ethan	theres to much of whatever its called
7	12:57:18.2	Olivia	it said the water looked cloudy
8	12:57:18.2	Ethan	and its making the water
9	12:57:21.9	Ethan	yea
10	12:57:24.0	Ethan	what olivia said
11	12:57:34.0	Facilitator	Okay, water is cloudy
12	12:57:34.8	Rakesha	cynabacteri
13	12:57:43.1	Facilitator	What makes water cloudy?
14	12:57:52.7	Ethan	to much cynabacteri
15	12:57:52.9	Rakesha	cynabacteria
16	12:58:23.4	Facilitator	what is cyanobacteria?
17	12:58:37.8	Ethan	its a thing thats good for tiapia
18	12:58:40.5	Ethan	but to much of it
19	12:58:44.0	Ethan	polutes the water
20	12:58:50.4	Jacob	yea that
21	12:59:03.7	Jacob	water gets sick fish get sick sick fish die

Jacob begins by positing that there was no problem with water quality but was countered by Ethan and Olivia (lines 3-10). Although Olivia's contribution is succinct, she gets her point across, and Ethan agrees with her assertion (line 8). Rakesha then extends this by noting that cyanobacteria are the cause of the cloudiness (lines 12 and 15). Ethan and Jacob were then able to build on these contributions to explain how the cyanobacteria can affect the fish (lines 17-21). This excerpt highlights that despite her lower contributions to chat, Olivia provides critical information for her peers. Closer inspection of Olivia's participation at the brainstorming board moreover indicated that she spent approximately 20 minutes reviewing the notes, compared to Ethan, who spent about 3 minutes on the notes (see Table 2). This additional data along with the board participation (Figure 2), suggests that the chat data only provides one aspect of engagement. In designing ECOJOURNEYS, we intentionally created numerous ways in which students would be able to participate within their groups to help build an argument and effectively solve the problem at hand. Equitable participation was encouraged through the design and implementation of these variable pathways for participation in the learning environment. This allowed individuals equal opportunities to share and showcase their knowledge through diverse means. To help us gain a better idea of student understanding of the system, consider the students' representations of what may be causing the tilapia illness. Figures 4 and 5 below illustrate the models drawn by Olivia, and Ethan, respectively.

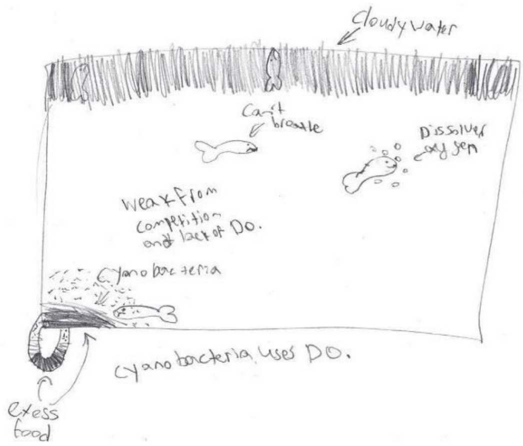


Figure 4. Olivia's model

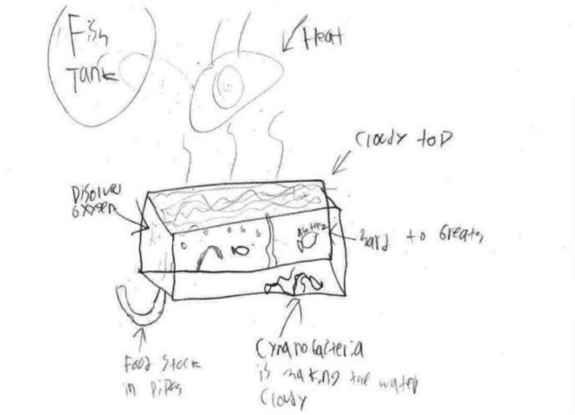


Figure 5. Ethan's model

Olivia's model (Figure 4) depicts several key facts of the system. It highlights the use of dissolved oxygen by both cyanobacteria and fish, which results in fish being weak (or ill) due to competition of resources (lack of dissolved oxygen), and presumably how the fish waste and food adds to the build-up of material in the water. Ethan's model (Figure 5) on the other hand, depicts components (heat, cloudy water, dissolved oxygen, food and cyanobacteria), but no indication of relationships among the components other than that the cyanobacteria make the water cloudy. Based on their patterns of participation, it is likely that the designed tools such as the brainstorming board facilitated student interactions with the learning material in their own ways. For example, it is likely that Olivia represented her knowledge of the system based on her use of the in-game notes, whereas Ethan may have benefitted more from his in-game interactions with his peers. In this way, the design considerations of this game may have encouraged and supported different forms participation among group members. It is clear from the data obtained from this study that low participation in chat features of ECOJOURNEYS, is not indicative of student engagement in learning activities and that equitable participation can be achieved through various means within game-based learning.

Discussion and implications

In designing this collaborative videogame, we focused on design features which would help promote equitable participation in each group. Although we have adopted a narrow definition of equitable participation, it is a crucial initial step in designing various activities through which students could engage in multiple pathways toward problem solving and work collaboratively with peers. From the structured design of the board, to the free use of the chat, to provisions of content material within the game, we designed with various student preferences and comfort levels of gaming in mind. Working collaboratively, allows students to bring their strengths, weaknesses, knowledge, and misconceptions to problem-based learning, so that together group members can build a strong argument and solve the problem at hand. However, if students are not afforded the opportunity to be encouraged and to feel comfortable enough to participate, then equitable student participation within collaborative game-based learning would be difficult to achieve.

With the popularity of game-based learning, designers and researchers need to attend to equity and inclusion. We should design learning spaces for all students, in which they are encouraged to participate through various forms of collaborative activity. Equitable participation should be at the forefront of collaborative game-based learning design as we seek to design for all learners. The diverse features of the game-based learning environment and the rules employed to foster equitable participation amongst group members, facilitated a learning environment in which all students were able to engage in the collaborative activities. Because some students may not use chat, we needed to design an alternative pathway for these students to express their understanding and contribute to the group. Working from the socio-cultural perspective, we need to consider backgrounds and preferences of learners when designing, which includes the development of multiple activities for participants. Learning activities and design features must engage students as we seek to make their learning and skills visible and valued to encourage equitable participation for all learners.

References

Buffum, P. S., Frankosky, M., Boyer, K. E., Wiebe, E. N., Mott, B. W., & Lester, J. C. (2016). Collaboration and gender equity in game-based learning for middle school computer science. *Computing in Science & Engineering*, *18*(2), 18-28.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661-686. https://doi.org/https://doi.org/10.1016/j.compedu.2012.03.004

Danish, J. A., & Gresalfi, M. (2018). Cognitive and sociocultural perspective on learning: tensions and synergy in the learning sciences. *International handbook of the learning sciences*, 34-43.

Engeström, Y. (1987). Learning by Expanding: An Activity-Theoretical Approach to Developmental Research. Orienta-Konsulti Oy.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. *Handbook of educational psychology*, *77*, 15-46.

Holzman, L. (2018). Zone of proximal development. In J. P. Lantolf, M. E. Poehner, & M. Swain (Eds.), *The Routledge handbook of sociocultural theory and second language development*. Routledge.

Liu, L., Hao, J., von Davier, A. A., Kyllonen, P., & Zapata-Rivera, J.-D. (2016). A tough nut to crack: Measuring collaborative problem solving. In *Handbook of research on technology tools for real-world skill development* (pp. 344-359). IGI Global.

Mislevy, R. J., Oranje, A., Bauer, M. I., von Davier, A., Hao, J., Corrigan, S., Hoffman, E., DiCerbo, K., & John, M. (2014). Psychometric considerations in game-based assessment. *GlassLab Report*.

Poehner, M. E. (2011). Dynamic Assessment: fairness through the prism of mediation. *Assessment in education: Principles, policy & practice, 18*(2), 99-112.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, 337-386.

Rasooli, A., Zandi, H., & DeLuca, C. (2018). Re-conceptualizing classroom assessment fairness: A systematic meta-ethnography of assessment literature and beyond. *Studies in Educational Evaluation, 56*, 164-181.

Rupp, A. A., Gushta, M., Mislevy, R. J., & Shaffer, D. W. (2010). Evidence-centered design of epistemic games: Measurement principles for complex learning environments. *The Journal of Technology, Learning and Assessment, 8*(4).

Saleh, A., Hmelo-Silver, C. E., Glazewski, K. D., Mott, B., Chen, Y., Rowe, J. P., & Lester, J. C. (2019). Collaborative inquiry play: A design case to frame integration of collaborative problem solving with story-centric games. *Information and Learning Sciences*, 120(9/10), 547– 566.

Stake, R. E. (2008). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (pp. 119-150). Sage.

Acknowledgments

This research was supported by the National Science Foundation through grants DRL-1561486 and DRL-1561655. Any opinions, findings, conclusions, or recommendations expressed in this report are those of the authors, and do not necessarily represent the official views, opinions, or policy of the National Science Foundation.

International Society of
the Learning Sciences

Taking Transactivity Detection to a New Level

James Fiacco, Ki-Won Haan, Anita Williams Woolley, Carolyn Rosé
jfiacco@cs.cmu.edu, khaan@andrew.cmu.edu, awoolley@andrew.cmu.edu, cprose@cs.cmu.edu
Carnegie Mellon University

Abstract: Transactivity is a valued collaborative process, which has been associated with elevated learning gains, collaborative product quality, and knowledge transfer within teams. Dynamic forms of collaboration support have made use of real time monitoring of transactivity, and automation of its analysis has been affirmed as valuable to the field. Early models were able to achieve high reliability within restricted domains. More recent approaches have achieved a level of generality across learning domains. In this study, we investigate generalizability of models developed primarily in computer science courses to a new student population, namely, masters students in a leadership course, where we observe strikingly different patterns of transactive exchange than in prior studies. This difference prompted both a reformulation of the coding standards and innovation in the modeling approach, both of which we report on here.

Introduction

Research shows that students benefit from rich discussions with other students in learning environments (Ferschke et al., 2015). Consequently, for more than a decade, researchers in the field of learning science have developed many frameworks for automated analysis of student discussion, as it has repeatedly been shown to be valuable for assessing student learning (McLaren et al., 2007; Dascalu et al., 2015; Joshi & Rosé, 2007; Rosé et al., 2008; McLaren et al., 2007; Ai et al., 2010; Gweon et al., 2013; Fiacco & Rosé, 2018), supporting group learning (Kumar et al., 2007), and enabling effective group assignments (Wen et al., 2016). Some work has explicitly addressed the issue of whether these frameworks generalize across domains (Mu et al., 2012; Fiacco & Rosé, 2018), which is critical to enabling educators in a variety of fields to leverage these tools. While cross-domain generalizability may sound like a purely technical problem, what we find in the current study is that the characteristics of different student populations and their learning processes, as well as the interplay between the two, are critical components. Here we report on the elements required to generalize technology developed for automated analysis of transactivity from one student population to another.

In this work, we focus on transactivity as a quality of conversational behavior where students explicitly build on ideas and integrate reasoning previously presented during the conversation. Transactivity stems from the Piagetian theory of learning. While its earliest formulations comprised a set of 18 different codes (Berkowitz & Gibbs, 1979), applications within the CSCL community, aiming to achieve success with automation, have targeted much simpler operationalizations defined by the presence of two requirements. First, the speaker must demonstrate a reasoning attempt. Second, the speaker must reference ideas or concepts presented earlier in the conversation. Prior datasets for transactivity typically identify the presence or absence of these requirements in a binary fashion (Joshi & Rosé, 2007; Wen et al., 2016). This approach has been successful in extant work which has focused on assignments that made use of very short collaborative discussions (Fiacco & Rosé, 2018), informal posts in discussion forums (Nelimarkka & Vihavainen, 2015), or team-based project support (Wen et al., 2016). However, the complexity of the language articulated within these previous works was limited, with some studies finding transactive exchange in only 60% of posts in a discussion forum (Sankaranarayanan et al., 2018). By contrast, in the current dataset, which includes masters level students in social science courses, about 80% of posts were rated as transactive by human coders using the simple definition, resulting in a lack of useful differentiation. Therefore, we turned our attention to adding more nuance to our operationalization of transactivity for automated analysis in order to better differentiate students and conversations.

Here we present a new dataset for transactivity detection based on a more detailed conceptual framework and measure. We then answer the following research questions targeted at automatic transactivity detection with respect to this new operationalization: First, can previous state-of-the-art models of transactivity detection apply to the domain of current event discussion forums in social science courses, and what phenomena exists in that domain that distinguish it from transactivity datasets? And, second, how can we capture these differences in a model that can better detect transactivity on this new dataset?

We show that, despite being highly functional on simpler datasets, the existing state-of-the-art model fails on our new dataset owing to the higher degree of abstractness in the conversations analyzed. We then present a new model based on this dataset that leverages the structure of the source data to more accurately predict more nuanced transactivity phenomena. The work illustrates how variations in student and course content result in different expressions of transactivity and that successful models must reflect those differences.

Transactivity coding

We collected communication data from 198 students in a master's degree program from a university in the Northeastern U.S. As part of a leadership course assignment, instructors provided students with a weekly article related to some of the topics learned from class for seven weeks. Students were instructed to post their thoughts on an online discussion board and also provide a response to at least three other classmates' posts. We extracted the response data, nested in each student's post thread, from the course platform in a json file format for each discussion topic. Across six different discussion topics, 3,415 replies in total were collected.

Building on prior work on transactivity coding (Berkowitz & Gibbs, 1979; Gweon, Jain, McDonough, Raj, & Rosé, 2013), we operationalized and coded the transactivity of the responses. Overall, the responses were found to be more elaborative than previous work that coded transactivity using a binary approach (i.e., transactive or non-transactive) and automatically annotated transactivity (Joshi & Rosé, 2007; Wen et al., 2016). But more importantly, while in prior studies the prevalence of transactivity was fairly low, in this population nearly 80% of contributions satisfied the simple definition of transactivity. The software infrastructure and nature of the task did not differ significantly from prior studies. Thus, we concluded that the substantial shift in conversational practices was due to the different student population and assignment, namely masters students in a social science course discussing current events.

To develop an appropriate framework, we proceeded to regroup the transacts of Berkowitz and Gibbs (1979) according to their roles (functions) in collaborative learning, develop a new coding scheme, and measure the level of transactivity or the extent to which an individual expended effort to represent and operate on their partners' reasoning. In their original framework, Berkowitz and Gibbs (1979) identified 18 types of transacts or dialogue behaviors, which are classified as higher or lower order transacts. Higher-order forms are operational transacts (e.g., counter consideration) that work on partners' reasoning through logical analysis and integration. Lower-order forms include representational (e.g., juxtaposition) and elicitational transacts (e.g., feedback request), which do not entail any transformations of partners' reasoning. Moreover, the transacts feature either competitive (e.g., competitive paraphrase) or non-competitive (e.g., paraphrase) forms, which can be focused on either partner and the dyad's positions.

Building from this framework, we developed new transacts and grouped them into three dimensions (functions): *active listening* (acknowledgment), *idea extension* (elaboration), and *challenging views* (qualification). First, we focused on *active listening* as it is conducive to creating an environment of mutual respect (Itzchakov, Kluger, & Castro, 2017) and psychological safety where the partners feel their contributions are valued and respected (Azmitia & Montgomery, 1993). In examining *active listening*, we coded whether the responders put in the effort to acknowledge their partners' ideas and thoughts by paraphrasing and/or asking them for further explanation. Second, more learning happens when discussions are disequilibrating, where individuals are exposed to something new from the interactions and experience cognitive perturbations (Berkowitz & Oster, 1985; De Lisi & Golbeck, 1999). Accordingly, *idea extension* evaluates the extent to which the individuals were elaborative in presenting their own reasoning processes in relation to their partners' original ideas or asked thought-provoking questions about their counterpart's contribution. Third, considering that cognitive perturbations could be more salient when there are conflicting views, thereby increasing the likelihood of transactive exchanges, we evaluated *challenging views* to assess the strength and clarity of the partners' challenging of their counterparts' argument.

In coding the data, each dimension was rated independently, although multiple dimensions might apply to a particular response. Specifically, for *active listening*, a binary rating was used, while for *idea extension* and *challenging views*, a 3-point scale (0: Not at all, 1: A little, 2: A lot) was used. In short, each statement was evaluated for the focal individuals' perceived exertion of effort to make sense of the meaning of their partners' argument (*active listening*) and build on their partner's reasoning (*idea extension, challenging views*).

Extended operationalization of new transactivity dimensions

Active listening

Active listening involves the focal participant's (Ego) acknowledgment of their partners' (Alter) contribution as is, in a non-judgmental manner. The transacts for *active listening* include "paraphrasing" and "soliciting clarification." That is, evaluators code whether Ego made the effort to paraphrase Alter's message and asked for further explanation to better understand Alter's point of view. Importantly, in soliciting clarification, Ego is not asking Alter to justify their reasoning or explore the ideas Ego proposed. The main criteria is: "Did Ego attempt to identify Alter's ideas and thoughts?" *Active listening* is coded "Yes (1)" when there is paraphrasing and/or soliciting clarification transacts; otherwise, it is coded "No (0)".

Coding examples for *active listening*

- Example of "Yes (1)"
  Alter: "… their job is to return value to the shareholders… That being said I don't think that only extroverted, or introverted people can do this. It just changes the way the company is set up and the culture that inherently stems from the leadership."
  Ego: "I agree with you that the main focus of hiring new leaders should be whether they can return value to the shareholders, but as you say introverted and extroverted leaders will set up different cultures. While these two cultures may be able to return the same value…"

Explanation: Ego paraphrased Alter's ideas in a clear way.

- Example "No (0)"
  Alter: "… Hiring new talent is an excellent way to gain access to these skills, but this should be in addition to retraining current staff, not in lieu of training. Some companies are able to fully hire a new staff, but many won't be able to do this…"
  Ego: "I think you make great points but there is one to add…"

Explanation: Here Ego did not demonstrate the way Ego understood Alter's argument.

Idea extension

In evaluating *idea extension,* coders annotated the extent to which Ego elaborated Alter's ideas by (1) exploring parallel lines of thought (i.e., agreement-based idea extension) and/or (2) qualifying Alter's argument (i.e., disagreement-based *idea extension*). Notably, Ego may demonstrate both forms of extension. First, for agreement-based extension, the transacts include "exploration," "exploration request," and "application." That is, Ego can provide additional evidence and thoughts either declaratively (exploration) or interrogatively (exploration request), and apply Alter's ideas to different contexts (application), such that Alter's argument becomes clearer and more generalizable. Second, for disagreement-based extension, the transacts include "critique" and "counter-argument." Ego can critically evaluate Alter's reasoning in a declarative or interrogative way (critique) and present opposing arguments (counter-argument)—uncovering the assumptions and exploring alternatives—such that Alter's argument becomes more robust and competitive. In evaluating this dimension, coders ask: "To what extent did Ego demonstrate their reasoning process?" Specifically, agreement-based *idea extension* is coded "A lot (2)" when Ego illustrated their argument with examples, demonstrated logical thinking, and/or integrated multiple ideas. Moreover, disagreement-based extension is coded "A lot (2)" when Ego explicated why Alter's argument may not be supported and/or provided clear evidence to support their counter-argument.

Coding examples for *idea extension*

- Example of "A little (1)"
  Alter: "…that extroversion became a cultural ideal and if extroversion is indeed the perceived ideal, maybe we have CEOs who learned how to be extroverted on the job because that is what is expected of them…"
  Ego: "You make a very interesting point. CEO extroversion could be a result of society's perception that it is crucial or more important than the other aspects and traits you mention. I agree that these are equally if not more important to hiring decisions. It would be interesting to see how this cultural ideal varies across countries/societies."

Explanation: Ego provided an additional thought, which needs to be developed.

- Example of "A lot (2)"
  Alter: "I don't believe that standardized tests should be used for college admissions, hiring, or anyplace else. Different people may have different skill sets that standardized tests don't take into account. Moreover, people may not have the same opportunity to be as prepared as they can for these tests."
  Ego: "The problem with eliminating standardized testing to remove bias is that there isn't a less biased criteria to replace it with. Ultimately, the bias shown on standardized testing is the result of general disadvantages that impact all parts of the student's application. In fact, when you consider recommendation letters…, essays…, and extracurricular activities that low income students simply can't afford, standardized testing is actually one of the less biased parts of the application…We also should do what we can to reduce the inequalities that cause all of these problems."

Explanation: Ego raised an alternative perspective and provided supporting pieces of evidence.

Challenging views

Challenging views gauges whether Ego was clear and extensive in stating their opposing position to Alter. Coders focused on the choice of words/phrases and the sentence structure to evaluate the clarity and strength of the challenge. Notably, coders do not evaluate if Ego's argument is relevant and well-reasoned, which is the focus of *idea extension*. The main criteria here is: "To what extent did Ego qualify Alter's argument"?

Coding examples of *challenging views*

- Example of "A little (1)"
  Alter: "I believe that firms should include retraining initiatives as they transform their businesses… Retraining is a difficult journey, but it is one that can be mutually beneficial for companies and their employees."
  Ego: "I agree that re-training employees will typically be worthwhile. But, should re-training be available to all employees?…"

Explanation: Ego agreed with Alter's view in general; Ego qualified one aspect of the argument.

- Example of "A lot (2)"
  Alter: "… I would have to assume that the team would be next to impossible to rectify in due time to complete the deadline for WS1, and I would respectfully decline the position…"
  Ego: "You have only described opportunities for James. The bar has been set low by the poor performance of the group which has been operating without a strong leader. James can be the new spark that keep everyone on track..."

Explanation: Ego qualifies Alter's argument directly, explaining how it can be interpreted in a different way.

Evaluation of interrater reliability

Two independent raters coded three dimensions of transactivity for a sample of 180 responses. To be comprehensive, interrater reliability was assessed using three measures, including intraclass correlations (ICC), Krippendorff's alpha, and weighted Cohen's kappa. For ICC, ICC(2, k) or a two-way random effects model was used (McGraw & Wong, 1996; Shrout & Fleiss, 1979). For ordinal variables, ICC is recommended (Hallgren, 2012). ICC is also suitable for nominal and continuous variables. Krippendorff's alpha (Hayes & Krippendorf, 2007) was also computed as a measure for assessing inter-rater reliability for all types of variables. Moreover, whereas Cohen's (1960) kappa is only suitable for nominal or categorical variables, weighted Cohen's (1968) kappa allows estimating the reliability for ordinal variables.

The results demonstrated excellent absolute-agreement ICC values for all dimensions: *active listening* (.89), *idea extension* (.87), and *challenging views* (.91). Krippendorff's alpha was found to be acceptable across dimensions: *active listening* (.80), *idea extension* (.76), and *challenging views* (.80). Last, Cohen's kappa showed moderate to strong levels of agreement: *active listening* (.80), *idea extension* (.69), and *challenging views* (.77). Given these values, we were confident in moving forward with our plan to have only one of the two raters code the responses that are required to train the machine for automatic detection of transactivity. A sample of 910 comments, consisting of a similar number of comments for each discussion topic, were coded to be used for deep learning, as discussed in the following section.

Automated transactivity detection experiments

Our goal was to find a model that most accurately predicts the various facets of transactivity that we have defined in our dataset. To this end, we started with an implementation of the previous state-of-the-art in transactivity detection to evaluate its ability to detect our more nuanced operationalization of transactivity in our data. We analyzed the data to identify reasons for the discrepancy in performance of the baseline model on each dataset. Our findings informed a new detector for transactivity to address the shortcomings of the baseline model. Below we provide an evaluation of the new transactivity detector.

Results for each experiment for transactivity detection were obtained via a 10-fold cross-validation where each fold was randomly assigned but consistent throughout the different conditions.

Baseline: Transferable attention model for transactivity detection

The model, called the Transferable Attention Model by Fiacco & Rosé (2018) is a variant of the Decomposable Attention Model for Recognizing Textual Entailment by Parikh et al. (2016), where the model is pre-trained on the RTE task after which the final layers are re-randomized and the model is allowed to fine-tune on the small transactivity dataset. Full implementation details of the model are discussed in Fiacco & Rosé (2018).

While the entailment task takes in a premise and a hypothesis statement to train the model with the hypothesis statement being the statement to be determined if the entailment relation holds, in the transactivity task, the premise is replaced by the context and the hypothesis is replaced by the message. The message is the text that is to be labeled as transactive and the context is the text for which the message is responding to.

For experiments on our dataset, the message was the post that is to be determined to show one of the aspects of transactivity while the context is the post to which that message responded. Note that the message and context may not be temporally adjacent as determination for message response was made via the forum response tree and participants can respond directly to prior posts.

Comparisons of transactivity data with respect to transferable attention model

The first research question we sought to address stems from a comparison between the data used by Fiacco & Rosé (2018) to train the Transferable Attention Model and our new dataset from class discussions. We noted that previous datasets used far more concrete language, while we observed more abstract language in our new dataset. Concreteness of language is characterized by referring to specific objects, people, or actions while abstractness is defined as language referring to concepts and ideas.

Table 1: Abstractness for datasets relevant to transactivity detection; scale 0 (concrete) to 1 (abstract)

Dataset	Text Abstractness
SNLI (Bowman et al., 2015)	0.334
MultiNLI (Williams et al., 2018)	0.530
Powerplant Transactivity Corpus	0.538
Masters Student Corpus	0.583

In Table 1, we present the abstractness of each dataset based on the average abstractness of inputs using the methodology from Brysbaert et al. (2014). We evaluate the transferable attention model using an alternative entailment pre-training dataset, the Multi-genre Natural Language Inference corpus (MultiNLI; Williams et al., 2018) which we found to be considerably more abstract than the Stanford Natural Language Inference (SNLI) corpus (Bowman et al., 2015) which was the pre-training corpus for the original Transferable Attention Model. This pretraining corpus was hypothesized to improve the model's performance by better representing the more abstract text found in the masters student data.

Transformer model for transactivity detection

One of the key shortcomings of the Transferable Attention Model is its inability to take into account word order. This is especially relevant to the *challenging views* dimension as negation is common within examples of that dimension and the meaning of a negation is highly word order dependent. To address this, we propose to use a class of models from the Natural Language Processing literature called transformers (Vaswani et al., 2017). The benefit of this type of model is that it combines the capability for self-attention with sequential reasoning to build a numerical representation of a sequence of text that can be used to classify that sequence.

Specifically, we use the pretrained transformer model, RoBERTa (Liu et al., 2019) which incorporates some optimizations of the BERT transformer model (Devlin et al., 2019). This model, like the GloVe embeddings used in the Transferable Attention Model, was pretrained on an enormous amount of general text data and will be fine-tuned on both the entailment pretraining task and the transactivity task. The model was then fine-tuned on the Recognizing Textual Entailment task similarly to the Transferable Attention Model. This fine-tuned model was the based model for each of the cross-validation folds. For each fold, the model was further fine-tuned on the transactivity data with a separate classification head as the entailment classifier.

Evaluation

We evaluated the potential to automate analysis using the extended transactivity definition proposed here, beginning with the best published approach from Fiacco & Rosé (2018), and comparing its approach to three other variants. From Table 2, it is evident that pretraining the Transferable Attention Model on the MultiNLI dataset had a large positive effect ($p < 0.05$) on all of the dimensions of transactivity. The increase was most notable for *active listening* while there were only modest improvements for *idea extension*, and *challenging views*.

Table 2: Cohen's kappa scores of transactivity detection models on 10 fold cross-validation

Model	Active listening	Idea extension	Challenging views
Transferable Attention Model (Fiacco & Rosé, 2018)	0.239	0.399	0.316
Transferable Attention Model+MultiNLI	0.314	0.429	0.340
Transferable Attention Model+MultiNLI+Self Attn.	**0.715**	0.656	0.461
RoBERTa+MultiNLI	0.651	**0.660**	**0.668**

Even more dramatic is the increase in performance from the redefinition of inputs for the Transferable Attention Model to make the model perform self-attention rather than attending between context and content. Furthermore, the RoBERTa model is able to significantly improve upon the performance on the *challenging views* dimension. However, it did not significantly improve on *idea extension* and underperformed on *active listening.* All differences between rows in Table 2 are significant ($p < 0.05$).

Discussion and conclusions

In this study, we uncovered some important considerations that must be taken into account when modeling approaches are used for automated detection of constructs such as transactivity. The line of experimentation reported here was prompted by an observation that a previously published demonstration of domain generality could not be generalized to a substantially different student population with distinctive discourse practices. Our findings point to necessary adjustments, first at the level of operationalization of the construct and then at the level of modeling approach -- with synergistic considerations between the two -- in order to achieve success.

In particular, our findings reveal a larger increase in performance for the *active listenin*g dimension between the baseline Transferable Attention Model and the version that used the MultiNLI pretraining as compared to the *idea extension* or *challenging views* dimensions. We attribute this largely to the vocabulary of the NLI datasets as compared to the masters student data. The masters student data is far more verbose and abstract than the SNLI dataset as compared to MultiNLI dataset. *Active listening* is a relatively simple task as compared to *challenging views* or *idea extension* as it is frequently signaled by agreement or disagreement. As the SNLI dataset is based off of image descriptions, there is little opportunity for that kind of language to occur. The MultiNLI corpus pulls data from a far broader range of genres and may expose the model to more relevant sentence forms. For *idea extension* or *challenging views*, the limiting factor was not as much the vocabulary, but how the model was able to use the data it had.

There was a large jump in performance across all dimensions of transactivity by redefining the Transferable Attention Model as a self-attention model as opposed to attending between the content and its context. While in data with less abstract contributions, the important factor for detecting transactivity may be ensuring that there are aligned phrases between the content and the context, in our masters student dataset, it appears to be more important for the model to understand what the responder is contributing. This result aligns with our qualitative observations that the masters students had deeper contributions and more structured responses as compared to the contributions in prior datasets. Detecting transactivity, in this case, is more about evaluating how the response is formed, regardless of the context.

This insight is reinforced by the performance of the RoBERTa based model that also uses sequential information to preserve the word order and sentence structure within the embedding. For c*hallenging views*, word order is critical to understand the content of a contribution as challenging one's view often involves negation. Negation can drastically change the meaning of a text segment depending on where it occurs. Adding the capability to do word order allowed the RoBERTa model to perform comparably between *idea extension* and *challenging views* while the Transferable Attention Model demonstrated a large gap between the two.

However, an interesting result was that the RoBERTa-based model performed worse than the Transferable Attention Model on *active listening*. A possible explanation of this comes from a qualitative analysis of the data where many of the *active listening* examples (for both the positive and negative cases) had a consistent structure where a student would express agreement or disagreement and then give an example. For the cases that reflected *active listening*, the example used specific language referring to content in the previous post (e.g. "I agree that re-training employees would be worthwhile.") For the cases that did not show *active listening*, the examples tended to use generalization or non-specific language (e.g. "I agree with what you said.") This difference can be modeled very well by simple self-attention; the model only needs to determine if the words attended to are generic or specific. Adding considerably more information via the RoBERTa model may make the distinction less clear.

Finally, the work reported here, to investigate the transfer of a successful automated analysis approach for transactivity from one context to another, is important for the community if resources are to be used efficiently through sharing. We began by recounting some history in application of the construct of transactivity to research in CSCL and the reasons why automation is valuable to the community. We then presented the contrasting case of masters students in social science to illustrate how population differences may be associated with substantial differences in discourse practices which may render earlier definitions unable to differentiate between students. A more nuanced operationalization and corresponding automation approach was therefore needed, which we have presented along with an evaluation in this paper. In future work, it would be valuable to explore how population differences impact desiderata for operationalization of other constructs related to collaboration process; it could be fruitful to identify how differences in population characteristics such as personality, age, academic/professional field or discussion context necessitate changes in the analysis approach. These further point to the need and potential value for a more coordinated effort across the CSCL community to provide sharable resources for automatic collaborative process analysis.

Acknowledgments

This research was funded in part by NSF grants ACI-1443068 and IIS 1546393 and funding from the Schmidt Foundation.

References

Ai, H., Sionti, M., Wang, Y.C., & Rosé, C. (2010). Finding transactive contributions in whole group classroom discussions. In Proceedings of the 9th International Conference of the Learning Sciences-Volume 1 (pp. 976–983).

Ankur P. Parikh and Oscar Täckström and Dipanjan Das and Jakob Uszkoreit (2016). A Decomposable Attention Model for Natural Language InferenceCoRR, abs/1606.01933.

Azmitia, M., & Montgomery, R. (1993). Friendship, transactive dialogues, and the development of scientific reasoning. Social Development, 2, 202-221.

Berkowitz, M. W., & Gibbs, J. C. (1979). A preliminary manual for coding transactive features of dyadic discussion. Unpublished manuscript, Marquette University, Milwaukee.

Berkowitz, M. W., & Oser, F. (1985). Moral education: Theory and application. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bowman, S., Angeli, G., Potts, C., & Manning, C. D. (2015, September). A large annotated corpus for learning natural language inference. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (pp. 632-642).

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. Behavior research methods, 46(3), 904-911.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.

Cohen, J. (1968). Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychological Bulletin, 70, 213-220.

Dascalu, M., Trausan-Matu, S., McNamara, D., & Dessus, P. (2015). ReaderBench: Automated evaluation of collaboration based on cohesion and dialogismInternational Journal of Computer-Supported Collaborative Learning, 10(4), 395–423.

De Lisi, R., & Golbeck, S. (1999). Implications of Piagetian theory for peer learning. In O'Donnell, A. M., & King, A. (Eds.), Cognitive perspectives on peer learning (pp. 213–312). New York, NY: Routledge.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 4171-4186).

Ferschke, O., Yang, D., Tomar, G., & Rosé, C. (2015). Positive impact of collaborative chat participation in an edX MOOC. In International Conference on Artificial Intelligence in Education (pp. 115–124).

Fiacco, J., & Rosé, C. (2018, June). Towards domain general detection of transactive knowledge building behavior. In Proceedings of the Fifth Annual ACM Conference on Learning at Scale (pp. 1-11).

Gweon, G., Jain, M., McDonough, J., Raj, B., & Rosé, C. P. (2013). Measuring prevalence of other-oriented transactive contributions using an automated measure of speech style accommodation. International Journal of Computer-Supported Collaborative Learning, 8, 245–265.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. Tutorials in Quantitative Methods for Psychology, 8, 23-34.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. Communication Methods and Measures, 1, 77-89.

Itzchakov, G., Kluger, A. N., & Castro, D. R. (2017). I am aware of my inconsistencies but can tolerate them: The effect of high quality listening on speakers' attitude ambivalence. Personality and Social Psychology Bulletin, 43, 105-120.

Joshi, M., & Rosé, C. P. (2007). Using transactivity in conversation summarization in educational dialog. Proceedings of the SLaTE Workshop on Speech and Language Technology in Education

Kumar, R., Rosé, C., Wang, Y.C., Joshi, M., & Robinson, A. (2007). Tutorial dialogue as adaptive collaborative learning supportFrontiers in artificial intelligence and applications, 158, 383.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1, 30-46.

McLaren, B., Scheuer, O., De Laat, M., Hever, R., De Groot, R., & Rosé, C. (2007). Using machine learning techniques to analyze and support mediation of student e-discussionsFrontiers in Artificial Intelligence and Applications, 158, 331.

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 discussionsInternational journal of computer-supported collaborative learning, 7(2), 285–305.

Nelimarkka, M., & Vihavainen, A. (2015). Alumni & tenured participants in MOOCs: Analysis of two years of MOOC discussion channel activity. In Proceedings of the Second (2015) ACM Conference on Learning@Scale (pp. 85–93).

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 learningInternational journal of computer-supported collaborative learning, 3(3), 237–271.

Sankaranarayanan, S., Dashti, C., Bogart, C., Wang, X., Sakr, M., Rosé, C. (2018). When Optimal Team Formation is a Choice - Self-Selection versus Intelligent Team Formation Strategies in a Large Online Project-Based Course, *Proceedings of AI in Education 2018*

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin, 86, 420-428.

Turney, P., Neuman, Y., Assaf, D., & Cohen, Y. (2011, July). Literal and metaphorical sense identification through concrete and abstract context. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (pp. 680-690).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998-6008).

Wen, M., Maki, K., Wang, X., Dow, S. P., Herbsleb, J., & Rosé, C. P. (2016). Transactivity as a Predictor of Future Collaborative Knowledge Integration in Team-Based Learning in Online Courses. In Barnes, T., Chi, M., & Feng, M. (Eds.), Proceedings of the 9th International Conference on Educational Data Mining (pp. 533–538).

Williams, A., Nangia, N., & Bowman, S. (2018, June). A Broad-Coverage Challenge Corpus for Sentence Understanding through Inference. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers) (pp. 1112-1122).

International Society of the Learning Sciences

How Real-Time Shared Gaze Visualizations Can Benefit Peer Teaching: A Qualitative Study

Tonya Bryant and Bertrand Schneider
tonya_bryant@gse.harvard.edu, bertrand_schneider@g.harvard.edu
Harvard University Graduate School of Education

Abstract: Real-time Shared Gaze Visualizations (SGVs) are a compelling way to encourage effective virtual teaching and learning interactions as SGVs can help to re-establish non-verbal social processes such as the attentional focus of group members. In this study, we look at a subset of data from a larger study (N=75) in which learners applied newly acquired knowledge about a microcontroller and programming to physical tasks with an instructor present as a support. We conducted a constant comparative analysis (Glaser & Strauss, 1967) using dual eye tracking video footage gathered across three experimental conditions (i.e., SGV, Head-Mounted Camera, Webcam). This paper supports a key finding from the larger study (i.e., SGVs help instructors track learner cognitive state), and its contribution goes one step further to identify a property of tracking cognitive state: Just-in-time support (described in findings section). We discuss implications of SGVs in peer teaching and conclude with anticipated future work.

Keywords: Synchronous Shared Gaze Visualizations, Joint Visual Attention, Grounding, Tangible Computing, Remote Learning and Teaching

Introduction

Remote learning and teaching is an increasingly common mode of learning. While it comes with many benefits such as connecting with others at the same time while in different places, it comes with unique challenges as well. In particular, the rich non-verbal information that is generated in social interactions such as gestures, facial expressions, postures, and eye movements is largely diminished in remote teaching and learning settings. Eye movements play an especially important role in facilitating effective teaching and learning as it indicates the object to which a person is attending (Allopenna, Magnuson, and Tanenhaus 1998; D'Angelo and Gergle 2016). Traditionally, gaze has been studied through tedious analytic methods such as interviews, case studies and video analyses. In this paper, we explore how to do this more easily with the use of shared gaze visualizations.

Shared gaze visualizations, or the real-time sharing of social partners' gaze locations, are a compelling way to encourage effective remote teaching and learning interactions as they can help to re-establish non-verbal social processes such as the attentional focus of collaborators. In educational scenarios that implement shared gaze visualizations, a commonly studied relationship dynamic is between experts and novices where an expert explains concepts with their eye movements being recorded. Then the novice watches the resulting video and takes a test (e.g., Mason, Pluchino, and Tornatora, 2015). The effect is that learners follow the asynchronous gaze of an instructor in a screen-based activity such as technical reading, language acquisition, computer programming, or perceptual tasks. Asynchronous gaze sharing may come with its own challenges with respect to gaze placement and its effect on learning outcomes (Jarodzka et al. 2013, 2010). Given that teaching and learning scenarios primarily occur synchronously and use physically manipulable materials to facilitate learning activities—even in remote settings—more exploration of synchronous gaze sharing in teaching and learning scenarios whose activities use tangible materials is needed.

From a socio-cultural perspective, studies examining expert/novice teaching and learning scenarios place a particular emphasis on the learner receiving information from the expert. The effect to date is that little attention has been given to how shared gaze visualizations can be used during application-centered tasks where learners apply newly acquired knowledge as they might in a 1:1 tutoring scenario, office hours for a technical course, or receiving after-school support from their teacher. An interesting opportunity, therefore, presents itself to seek to understand shifts in core activities for experts and novices where experts shift their core activity from instructing to supporting, and novices shift their core activity from learning to practicing and applying. This shift in core activity for learners, in particular, may support more meaningful learning through meaning making (Stahl, 2007) and transference (Engle, 2006).

Given the current context of remote learning and teaching and outlined gaps in literature in the topical areas of synchronous shared gaze visualizations, expert/novice dyadic interactions, and learner-led activities involving physical tasks, we ask (RQ): *How does the presence or absence of synchronous shared gaze*

 © ISLS

visualizations influence the dyadic interactions between an expert and a novice programmer throughout a remote learner-led tangible computing task across three conditions: Shared Gaze Visualization (SGV), Head-Mounted Camera (HMC), and Webcam (WC)? Furthermore, (RQ1a) *What strategies do instructors use to establish/maintain grounding and what do these interactions look like?* (RQ1b) *At what points in time do instructors assert themselves into the learner's cognitive process and what do those interactions look like?* We begin exploring these questions by sharing the theoretical background underpinning the purpose of shared gaze visualizations: grounding in social interactions (Clark & Brennan, 1991; Clark, 1985).

Theoretical background

In dyadic interactions, the two people involved must coordinate on the *content* and the *process* of what they are doing to share a common understanding of their work (Clark & Brennan, 1991). When a teacher and student meet for a 1:1 tutoring session on conjugations of irregular verbs for a French class, they must both work from the same course materials such as the chapter in a textbook on conjugations of irregular verbs (i.e., coordination on content). They must also synchronize their actions at the beginning and ending of their time together and across the sequence of events that unfold over time (Jordan and Henderson, 1995), thereby coordinating on process. The wealth of information generated and received from each person while engaging in the content and process develops a common ground—a non-static process continually augmented with new information through the willingness of both people to continue interacting, perceive the messages sent by the other, to understand the messages, and react and respond adequately to the messages (i.e., accept or reject them) (Clark, 1985). Certainly, verbal communication facilitates grounding, but there is also a wealth of non-verbal information exchanged between the individuals that facilitates grounding such as gestures, postures, facial expressions, and eye movements. Eye movements in particular play a unique role in that they communicate information about what the other person is presently attending to. This type of attentional awareness is useful in various scenarios particularly with respect to assisting in the formation of joint visual attention—the tendency for social partners to focus on a common reference and to monitor one another's gaze to an outside entity such as an object, person, or event (Tomasello et al., 2005).

Joint visual attention (JVA) has been extensively studied across a variety of domains and is an active area of research in the social sciences observing, for example, the importance for individuals to learn how to socialize and develop social motivation (Bruinsma, Koegel, & Koegel, 2004; Salley & Colombo, 2015). In developmental psychology, JVA is highlighted as an important mechanism for social coordination between family members and young children (Scaife & Bruner, 1975; Bates, 1976; Mundy et al., 1990; McClure et al., 2018). In the learning sciences, JVA is used to understand collaborative small group activities between peers (Roschelle & Teasely, 1995; Dillenbourg et al., 2006; Schneider & Pea, 2013). Interest in JVA by Human-Computer Interaction researchers has led to the development of shared gaze visualization tools to support effective communication and collaboration in remote settings (D'Angelo and Gergle, 2016; Higuchi et al., 2016). Much of the focus of JVA in computer-supported collaborative learning (CSCL) looks at the application of shared gaze visualization tools in peer collaborations where learners work together to solve a problem or in expert/novice scenarios where the expert and novice take on traditional roles in educational spaces. An under-explored area of CSCL research relates to the expert/novice relational dynamic where practice and application of newly gained knowledge is the central activity and tangible computing objects are the content bringing learners and instructors together. This research aims to explore these areas through a qualitative study where we look at a subset of data from a larger study (N of dyads=6 of 75) in which learners applied newly acquired knowledge about a microcontroller and programming to tangible computing tasks with an instructor present as support. A general description of the study follows.

General description of the study

In a larger study (Sung, Feng & Schneider, 2021), a trained instructor guided a novice through a series of increasingly complex physical tasks designed to teach the novice how to read and interpret code, and use the components of a microcontroller (i.e., a GoGo board) to simulate real-world scenarios such as making a streetlight turn on at nighttime when a person walks by. The goal of the study was to help domain experts unpack the cognitive state of a less knowledgeable peer using synchronous shared gaze visualizations in a 1:1 teaching and learning interaction. Instructor participants (N=18) were recruited from a master's program in educational technology. Many had short teaching experiences, and all had prior experience in microcontrollers, programming, or both. Learner participants were 75 adults between the ages of 18 and 38, with little to no experience with microcontrollers or programming.

While the larger study consisted of two main activities, the present study (N of dyads=6) focuses on the second activity where the learner took the lead on three GoGo board tasks with the instructor present for support. The learning objective of the first task (i.e., circuit assembly) was to read a block-based computer program and assemble hardware to the GoGo board to produce the desired output (i.e., turn on an LED light). The learning objective of the second task (i.e., a visual search task) was to compare an error-free block-based program to an identical line-based program and identify five errors in the line-based program. The third task was excluded from analysis as most participants did not complete it. Instructors were encouraged to give hints only when they perceived it was clear the learner needed help. There were three experimental conditions: Shared Gaze Visualization feed (SGV), Head-Mounted Camera (HMC), and Webcam (WC). Each experimental condition had an equal number of sessions, and instructors contributed the same number of sessions to the three conditions through a randomized block design. This was a design choice to minimize instructor effect.

Methods

Data generated from this study used physiological and traditional data collection instruments. Traditional data collection instruments included pre/posttests, a post-study survey, instructor predictive ratings on learner post-test scores, and webcam video recordings. Physiological data collection instruments included Empatica E4 wristbands and Tobii Pro eye tracking glasses. For details on how these instruments were used for measurement in the larger study, please refer to Sung, Feng & Schneider (2021). This study uses the webcam recordings and video footage rendered from eye tracking glasses. Circles representing the synchronous gaze of the participants were overlaid onto eye tracking videos during post-processing and are referred to as shared gaze visualizations (SGVs) in this paper. The goals of observing webcam recordings (which were not augmented with SGVs) were to: (1) gain familiarity with data; observations were made blind (i.e., the conditions were not known to the analyst); and (2) simulate the experience of researchers who were present during data collection. While those researchers were privy to the conditions, they (like the analyst) were unaware of the eye movements of the participants as they did not view the screens of participants during the study. Methodology for analysis is as follows:

Data organization

Webcam videos used in the present qualitative study were previously cleaned and analyzed for a prior research study whose research question was also interested in activity 2, tasks 1 (i.e., circuit assembly) and 2 (i.e., a visual search task). To organize webcam video data, three analysts timecoded videos (n=75) to identify the beginning of task 1, and then removed sessions with lost files and crippling technical errors (i.e., serial equipment failure). This resulted in a total of 59 sessions. We further cleaned the data to consider only sessions where participants carried out activity 1, tasks 1 and 2 with fidelity. This resulted in a total of 41 viable sessions. To account for instructor effect during analysis, we distributed instructors across strata of analysis such that each instructor only appeared once in each stratum (n=6). Approximately eight hours were spent to conduct data organization with timecoding taking roughly 75% of the total time.

Data cleaning, key observations, and sub-research questions

We systematically grouped dyads by learning gains scores to assist the data selection process. We found the mean of the absolute difference between pre/post test scores (i.e., learning gains) for all sessions (n=59) rounding up to the nearest whole number (i.e., n=22 points for pre/post-tests measured on a 100-point scale). A learner who increased their pre-test score by 23 points or more was considered a high achieving learner (i.e., HAL); those with 22 points or less were considered low achieving learners (i.e., LAL). 64% (n=25) of the sessions were categorized as HAL sessions and 36% (n=16) of the sessions categorized as LAL sessions, with HALs representing roughly two-thirds of the viable sessions (n=41). We took care to select different instructors. The six sessions selected for observation were **12** (Gaze x LAL) and **67** (Gaze x HAL), **69** (HMC x LAL) and **114** (HMC x HAL), and **110** (WC x LAL) and **103** (WC x HAL). One analyst spent roughly 10.5 hours cleaning data and writing analytic memos for each session. Participants took between five and eight minutes to complete both tasks. Video observations/memo writing consumed roughly 80% of the quoted time. Key observations from analytic memos include (1) instructors use different strategies to establish grounding with the learner and (2) instructors assert themselves into the learner's process at different points in time. At this point, we revisited our main research question, compared it to our key observations, and devised two sub-research questions to focus our analysis. (See RQ1a and RQ1b in the findings section below for details.)

Transcription, coding, interaction analysis, and framing key observations

We built upon analytic memos by including observations from eye tracking videos overlaid with SGVs of both participants in each dyad. We transcribed the learner-led tasks (tasks 1 and 2) and inductively coded transcripts for strategies (Bogdan & Biklen, 2007) used by instructors to establish grounding. (See Figure 1 below for details.) To understand the interactions between the learner and instructor, we did a second round of analysis using Jordan and Henderson's (1995) guideline for Interaction Analysis with "the structure of events" as our analytic focus. Events that included strategy for grounding became viable supportive evidence for RQ1a. These processes took roughly 15.5 hours to complete for one analyst.

In the following section, we share key findings for our research questions using intrinsic case studies as they offer an opportunity to understand particularities (Mills et al., 2010). We use Clark and Brennan's (1991) principles of least collaborative effort as a general guideline to structure key observations. We modify interpretations of media constraints for germaneness and to acknowledge technological advances in media since the publication of the article. Jordan and Henderson's (1995) "the structure of events" framework for Interaction Analysis scaffolds supportive evidence for key observations. Examples selected as supportive evidence describe common behaviors found in literature on grounding and JVA (e.g., pointing, holding up objects to screens, and asking a learner to verbalize their thought processes) or illuminate just-in-time support and other ways instructors assert themselves into a learner's cognitive process to provide support.

Findings

In this section, we share insights into our main research question and its two sub-questions. We start by sharing key observations for each sub-question and then follow-up with supportive evidence.

Key Observations for RQ1a: We observe in the SGV condition that participants were able to exert least collaborative effort to achieve grounding as they were able to use their gaze as a deictic gesture to establish grounding (Finding 1a). In non-SGV conditions, participants resort to forms of grounding that are more costly in collaborative effort than those observed in the SGV condition. In the HMC condition, the instructor requested the learner to verbalize all their thoughts as they assembled the circuit in task 1 (Finding 2a). In the WC condition, the instructor asks the learner to hold up the completed circuit to the camera (Finding 3a).

Key Observations for RQ1b: We observe in the SGV condition that the instructor asserts themselves into the learner's process just-in-time to support the learner with a challenge during task 2 (Finding 2a). The instructor is able to do this because they able to track the learner's cognitive state for the duration of the event. This translates to a low collaborative effort exerted to provide and receive help. In the HMC condition, the instructor asserts themselves into the learner's thought process sporadically in attempt to help the learner with a challenge during task 2 (Finding 2b). The instructor in the WC condition asserts themselves into the learner's process only after the learner has announced they have finished the circuit (Finding 3b). In the following subsections we provide brief examples of each of the findings, starting with RQ1a.

Supportive Evidence for RQ1a: What strategies do instructors use to establish/ maintain grounding and what do these interactions look like?

Finding 1a (SGV): Instructor uses gaze as a deictic gesture

The event we chose for this example is a stretch of interaction from session 12 (SGV x LAL) that occurs as a transition event between tasks 1 (i.e., circuit assembly) and 2 (visual search for errors in line-based code). The smaller unit of particular interest within the event is the preparatory action the instructor took to establish grounding before launching into a series of smaller events whose end goal was to help the learner understand how programs communicate with computers to produce desired outputs.

The learner begins the segment with a question ("Okay, so even though you're showing me the kind of behind the curtain view, I still don't understand how these machines know what to do—like how they speak these languages. This is wild!") and the instructor acknowledges the learner's confusion ("Oh, it knows what to do because people have written this code and already loaded it into the GoGo board") and then points with her gaze ("So, if it's not loaded, you'll have to write it here").

Before beginning their explanation, the instructor engages in a preparatory action: get the learner to look at a specific tab of the widget screen where the program for the prior task was written. In a physical setting, the instructor might have simply pointed to the specific tab; however, in this remote setting enhanced by shared gaze visualizations, the instructor simply used their gaze as a deictic gesture to achieve JVA. (See Figure 1 below.) By having access to this tool, the instructor was able to establish grounding effectively and easily. Ease of use allowed them to stay focused on the main purpose of their interaction until the learner indicates the end of the interaction by expressing understanding ("Ohh, I see. Oh wow, okay").

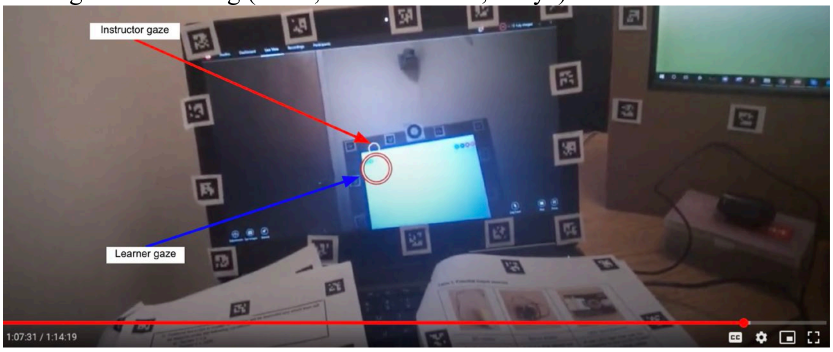
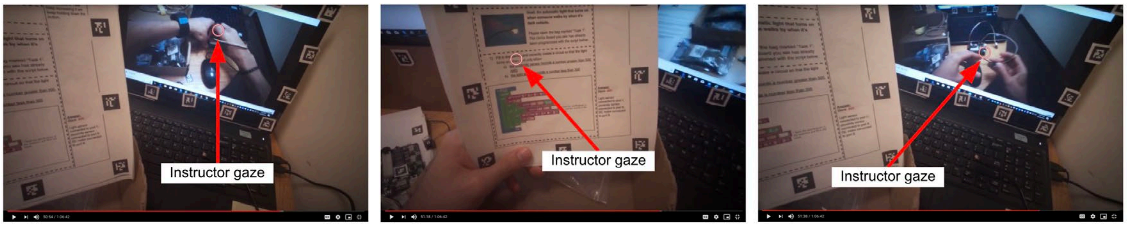


Figure 1. The instructor (white circle) points to a tab on the widget screen using their gaze to create grounding through JVA. The learner shifts their gaze (red circle) to the new object of interest.

Finding 2a (HMC): Instructor asks the learner to think aloud and attends to two external objects

The event we chose for this example is a stretch of interaction from session 114 (HMC x HAL) where the learner is assembling the input components to the GoGo board (i.e., task 1). While the learner is verbalizing each step in their thought process the instructor listens while panning back and forth between the laptop monitor (to attend to the learner's assembly process) and the instruction manual:



"The input in 1 is a light sensor." "Then the input in 8 is a proximity sensor." "Okay, so the inputs are in."

(a) (b) (c)

Figure 2. The instructor attends to two external objects as the learner verbalizes their thought process. The images show the instructor's gaze panning from the computer screen (a), to the manual (b), back to the computer screen (c). The quotes below the images show what the learner is saying during each screenshot.

The strategy the instructor uses—requesting the learner verbalize their thought process—helps establish and maintain grounding. To confirm the learner's actions, the instructor collates the learner's utterances (i.e., auditory information) by reading with each line of code in the manual (i.e., visual information). By asking the learner to verbalize their thought process, the instructor can see that the learner reads each line of code sequentially and assembles the components accordingly. The learner's behaviors are consistent, which allows the instructor to predict the learner's next actions. The think aloud strategy appears to be an effective way to create and maintain grounding, but not without much collaborative effort from both participants. Additionally, since the instructor must attend to two objects of interest (the laptop computer screen and the instructional manual), when they look toward one, they lose the benefit of the other. So, for instance, when the instructor monitors the learner's process on the laptop, they temporarily lose the power to predict the learner's next steps as they are not looking at the manual. While mostly effective, verbalizing thought processes does not appear to be a sustainable way to maintain grounding between social partners.

Finding 3a (WC): Instructor asks the learner to hold up the assembled circuit to the camera

The event we chose for this example is a stretch of interaction from session 103 (WC x HAL) where the learner has just completed task 1 (i.e., circuit assembly). Just prior to the completion of the task, the instructor tells the learner to show her the circuit once they are done ("You can hold up the circuit once you're done so that I can see it"). To show they are done, the learner must both demonstrate the input and output cables are plugged into the correct ports and demonstrate the correct input and output sources are connected to the cables, which dangle outside the field of view of the camera. A sequence of smaller units within the event unfolds, where the learner must continually calibrate the distance of the GoGo board from the camera (i.e., moving it closer or farther away) as well as move the whole apparatus up and down in slight movements to bring the light and proximity sensors and the LED output into the field of view of the camera. The instructor leans in and holds up their hand to communicate to the learner to hold still. Overall, holding an object up to a computer camera, particularly one that is as dynamic as a microcontroller with dangling cables and small components, seems to require high collaborative effort from both participants.

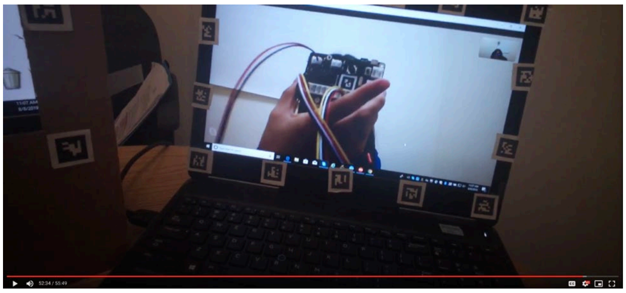


Figure 3. The learner holds up the completed circuit to the camera to let the instructor assess assembly.

Supportive Evidence for RQ1b: At what points in time do instructors assert themselves into the learner's cognitive process and what do these interactions look like?

Finding 1b (SGV): Instructor asserts themselves just in time to support the learner

The event we chose for this example is a stretch of interaction from session 67 (SGV x HAL) that shows how the SGV tool allowed the instructor to provide just-in time support to the learner during task 2 (i.e., visual search). Just-in-time support is enabled by the SGV and allows the instructor to stay engaged with the learner's thought process for the duration of the task. The goal is for the learner to identify five errors in the line-based code by comparing it with error-free block-based code.

In this event, the instructor prompted the learner to identify four remaining errors in the line-based code. The learner identified the first three errors ease. Throughout the task, the instructor tracks the learner's cognitive state. When the learner gets stuck finding the fourth error and makes a second attempt, the instructor is able to see the learner searching each line of code for an error. The instructor notices the learner's attempt is complete and that they are still stuck. So, they immediately give the learner a hint directly related to the problem, which helps the learner reach their goal quickly. This example suggests that SGVs predisposed the instructor to the series of events that led up to an issue (e.g., learner is stuck), so that when the issue surfaces the instructor is able to provide accurate, just-in-time support to the learner. Figure 4 shows the instructor tracking the cognitive state of the learner while they try to solve the problem.

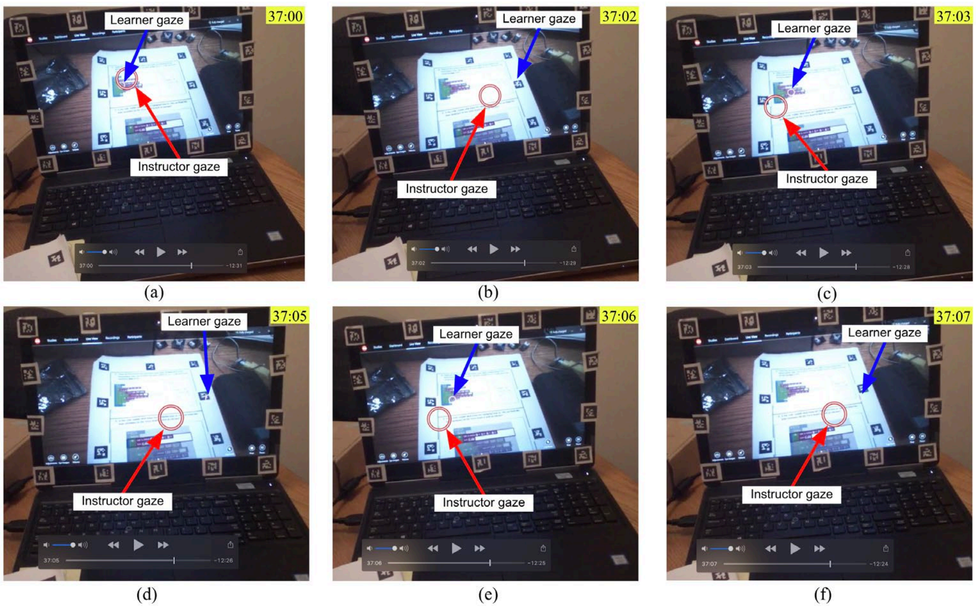


Figure 4. The instructor tracks the learner's cognitive state during a visual search task. It is difficult to see the white circle that represents the learner's gaze. See the middle image on the bottom row (e) for a clear example.

Finding 2b (HMC): Instructor asserts themselves sporadically into the learner's process

The event we chose for this example is a stretch of interaction from session 69 (HMC-LAL) where the learner is searching for the last error in the line-based code in task 2. The beginning of the event starts with the instructor quickly summarizing the goal of the task and then hands it off to the learner. We observed the SGV of the learner's gaze (imperceptible to the instructor) moving between the first line of code in the task and the instruction manual, suggesting the learner needed time to orient themselves with the task. The learner identifies all but one of the errors and get stuck, so they start the search again (gaze moves back to top code). As time passes without a response from the learner, the instructor begins sporadically interjecting with suggestions (e.g., "You can take a look at other code if you want to get a sense", "I'll give you a hint. It has to do with 'to main', "That was a big hint", "What's the last block say?"). All except the last hint are ignored, indicating the instructor's suggestions were not useful in solving the learner's challenge and that the instructor was unaware of the learner's needs.

Finding 3b (WC): Instructor asserts themselves at the end of the learner's process

The event we chose for this example is a stretch of interaction from session 110 (WC x LAL) where the learner works on task 1 (i.e., circuit assembly). Throughout the learner's process, the instructor took a hands-off and eyes-off approach, rendering all the learner's processes imperceptible. They attended to the widget screen (located on a separate external monitor) until the learner announced they completed the task ("I think I got it together"). Then the instructor pivots to engage with the learner ("Great, okay. Let's go ahead and hit your run button"). What follows is 5 minutes and 51 seconds of backtracking through the task and rewiring the circuit together as the instructor discovers multiple interrelated challenges the learner faced (i.e., (1) how to interpret nested functions, (2) how threshold values written into the program relate to the two inputs, (3) the necessity of plugging cables into their respective ports, and (4) how inputs relate to the outputs through the microcontroller). This is evident as the learner has completed the circuit incorrectly (i.e., proximity sensor is plugged into the wrong input port) and incompletely (i.e., output components are missing). This example illustrates the importance of maintaining grounding throughout the learning process so the instructor can identify appropriate opportunities to help the learner before complex problems arise.

Discussion and conclusion

In this study, we showed ways that the presence or absence of SGVs influenced the dyadic interactions of instructors and learners during a remote learner-led tangible computing activity. We found supportive evidence that SGVs were useful for helping instructors establish and maintain grounding by using gaze as a deictic gesture and tracking the cognitive state of learners. These findings support the established understanding that non-verbal communication cues, such as eye movements, play a key role in facilitating teaching and learning in that they communicate information about what the other person is presently attending to. In this study, being able to perceive the learner's thought processes in real time enabled the instructor to identify appropriate points in time to interject with support, suggesting potential other benefits or uses of synchronous shared gaze visualizations not anticipated prior to this study. SGVs can be used as a feedback mechanism that helps instructors self-regulate the timing of the support they provide learners by tracking the learner's cognitive state. In turn, learners are afforded the space and time needed to work on a problem independently, potentially contributing to meaning making, transference, and even flow. This paper was a first step in understanding learner-instructor interactions in a remote tangible computing activity and the benefits of SGVs therein. Future work aims to build upon the concept of tracking learner cognitive state and its perceived benefits, such as just-in-time support and instructor self-regulation as well as identifying limitations to SGVs in similar teaching and learning scenarios.

References

Allopenna, Paul D, James S Magnuson, and Michael K Tanenhaus. 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.

Bates, Elizabeth, & Hammel, E. A. (2014). The Emergence of Symbols: Cognition and Communication in Infancy (Language, thought, and culture). Academic Press.

Bates, E. (1976). Language and context: The acquisition of pragmatics. New York: Academic Press.

Bogdan, R.C. and Biklen, S.K. (2007). Qualitative Research for Education: An Introduction to Theory and Methods. 5th Edition, Allyn & Bacon, Boston.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), Perspectives on socially shared cognition (p. 127–149). American Psychological Association. https://doi.org/10.1037/10096-006

Clark, Herbert H. 1985. "Language use and language users." Handbook of social psychology .

Corbin, J., & Strauss, A. (2008). Basics of qualitative research (3rd ed.): Techniques and procedures for developing grounded theory. Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781452230153

D'Angelo, Sarah, and Darren Gergle. 2016. "Gazed and Confused: Understanding and Designing Shared Gaze for Remote Collaboration." In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2492–2496. ACM.

Dillenbourg, P., Baker, M. J., Blaye, A., & O'Malley, C. (1996). The Evolution of Research on Collaborative Learning. In E. Spada, & P. Reiman (Eds.), Learning in Humans and Machine: Towards an Interdisciplinary Learning Science (pp. 189-211). Oxford, UK; New York: Pergamon.

Engle, R. (2006). Framing Interactions to Foster Generative Learning: A Situative Explanation of Transfer in a Community of Learners Classroom. The Journal of the Learning Sciences, 15(4), 451-498. Retrieved November 15, 2020, from http://www.jstor.org/stable/25473531

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research.

Higuchi, K., Yonetani, R., & Sato, Y. (2016). Can Eye Help You?: Effects of Visualizing Eye Fixations on Remote Collaboration Scenarios for Physical Tasks. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.

Jarodzka, Halszka, Katharina Scheiter, Peter Gerjets, and Tamara Van Gog. 2010. "In the eyes of the beholder: How experts and novices interpret dynamic stimuli." Learning and Instruction 20 (2): 146–154.

Jarodzka, Halszka, Tamara van Gog, Michael Dorr, Katharina Scheiter, and Peter Gerjets. 2013. "Learning to see: Guiding students' attention via a Model's eye movements fosters learning." Learning and Instruction 25: 62–70.

Mason, Lucia, Patrik Pluchino, and Maria Caterina Tornatora. 2015. "Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning." Contemporary Educational Psychology 41: 172–187.

McClure, Elisabeth R, Chentsova-Dutton, Yulia E, Holochwost, Steven J, Parrott, W. G, & Barr, Rachel. (2018). Look At That! Video Chat and Joint Visual Attention Development Among Babies and Toddlers. Child Development, 89(1), 27-36.

Mills, A. J., Durepos, G., & Wiebe, E. (2010). Encyclopedia of case study research (Vols. 1-0). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412957397

Mundy, Peter, Sigman, Marian, & Kasari, Connie. (1990). A longitudinal study of joint attention and language development in autistic children. Journal of Autism and Developmental Disorders, 20(1), 115-128.

Roschelle, Jeremy & Teasley, Stephanie. (1995). The Construction of Shared Knowledge in Collaborative Problem Solving. Computer Supported Collaborative Learning. 10.1007/978-3-642-85098-1_5.

Scaife, M, & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. Nature (London), 253(5489), 265-266.

Schneider, Bertrand, and Roy Pea. 2013. "Real-time mutual gaze perception enhances collaborative learning and collaboration quality." International Journal of Computer-Supported Collaborative Learning 8 (4): 375–397.

Stahl, Gerry. 2007. "Meaning making in CSCL: Conditions and preconditions for cognitive processes by groups." In Proceedings of the 8th international conference on Computer supported collaborative learning, 652–661. International Society of the Learning Sciences.

Sung, G., Feng, T., & Schneider, B. (2021). Learners Learn More and Instructors Track Better with Real-time Gaze Sharing. International Conference on Computer-Supported Collaborative Work / Proceedings of the ACM: Human Computer Interaction (PACM HCI).

Tomasello, M. (1995). Joint Attention as Social Cognition. In Moore, C., & Dunham, P. (Eds.), Joint attention: its origins and role in development (pp. 103-130). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Challenging Joint Visual Attention as a Proxy for Collaborative Performance

Kshitij Sharma, Norwegian University of Science and Technology, kshitij.sharma@ntnu.no
Jennifer K. Olsen, University of San Diego, jenniferolsen@sandiego.edu
Himanshu Verma, Delft University of Technology h.verma@tudelft.nl
Daniela Caballero, University of Chile, daniela.caballero@ciae.uchile.cl
Patrick Jermann, École Polytechnique Fédérale de Lausanne, patrick.jermann@epfl.ch

Abstract: Researchers have used joint visual attention (JVA) as a proxy for collaborative quality and/or performance during the last decade due to its association with both measures. However, the notion of looking at the same object does not necessarily indicate that students are solving the problem together (or learning together). We propose a complementary approach to joint visual attention by augmenting it with joint mental effort (JME). JME is computed as a cross-recurrence of the cognitive load of the peers in a dyad. We use data from 41 dyads to show the synergy between JVA and JME and the insights that they can shed in the collaborative process. The results show that in certain episodes of collaboration (characterized by the dialogue and division of labor strategy of the dyad) combining these two dual-eye tracking measures provide deeper insights about the collaborative processes and performance than JVA alone.

Keywords: Dual eye-tracking, Collaborative Processes, Joint Visual Attention, Cognitive load

Introduction

It is not a new idea that collaborative learning can be beneficial for student learning outcomes nor within the computer-supported collaborative learning (CSCL) community that technology can be used to support the collaborative process (Johnson & Johnson, 1987; Lou, Abrami, & d'Apollonia, 2001). However, it continues to be an on-going process to understand what a productive collaboration process entails. As technology and computational methods continue to develop, our ability to measure the collaborative learning process changes by measuring it through different modalities and over time (Olsen, Sharma, Rummel, & Aleven, 2020; Starr, Reilly, & Schneider, 2018). A common motivator for finding new means of assessing the quality of the collaboration process has been the difficulty and time-consuming process of analysing student dialogue, especially in real time to be able to put interventions in place (Sharma et al., 2017). In these cases, the new measures are often a proxy for the dialogue content. When we consider collaboration measures in this way, we either intentionally or unintentionally assign moments of silence as less valuable. Rather, new measures of collaboration that can complement existing measures can fill in these gaps. In this paper, we aim to deepen our understanding of what effective collaboration looks like through the assessment of dual eye tracking measures. Furthermore, by assessing these eye tracking measures, we contribute to the understanding of how eye tracking can be used to analyse the collaborative learning process.

One collaborative learning theory is that students are able to assess and update their mental models of the domain by working with peers (Chi & Wylie, 2014). Through the process of co-construction, students can reflect on their own mental model to make repairs, incorporate their partner's ideas into their model, and construct new knowledge by building upon their partner's ideas (Hausmann, Chi, & Roy, 2004). In this case, the benefits of collaboration come from the joint construction of knowledge that occurs as students work together. To measure these processes, we can analyse episodes of interaction for indicators of students integrating their partner's ideas into their thought process. For example, one can measure collaboration through the use of transactivity (Joshi & Rosé, 2007) or through interactive dialogue as proposed in the ICAP framework (Chi & Wylie, 2014). Many of these coding schemes focus on the different ways in which students can construct knowledge. In a collaborative setting, this is just one aspect of the collaboration with students also needing to coordinate the work and in CSCL settings, coordinate with the technology (Rummel et al., 2011). Across all of these aspects, researchers have mainly used dialogue content as the gold star measure for collaboration but interactions with the activity are also common. When other collaborative learning measures are proposed, too often they are used as a proxy for analysing the dialogue content rather than complementing it (Sharma et al., 2017). The work in multi-modal learning analytics begins to address this gap by investigating how multiple modalities of data can be used together. However, before combining data streams, it is important to understand what each data stream can provide.

In this paper, we focus on the information that can be provided through dual eye tracking (DUET). In previous research, DUET has been used as a tool to explain the socio-cognitive mechanisms underlying collaborative learning (Jermann and Nuessli, 2012; Sharma et al, 2018 & 2020; Olsen et al, 2018). Information extracted from DUET data has been used to explain collaboration quality (Schneider et al, 2019), collaborative task-performance (Sangin et al., 2011; Jermann and Nuessli, 2012; Sharma et al, 2017), and collaborative learning gains (Olsen et al., 2020). Duet also has been used to explain certain processes related to collaborative learning, for example, mutual modelling (Lemaignan and Dillenbourg, 2015), repairs of misunderstanding (Cherubini et al., 2008), shared understanding (Richardson et al., 2007), and coordination (Brennan, et al., 2008). Additionally, researchers have used DUET as a method to provide collaborative awareness to the peers attempting to solve a given problem (Schneider and Pea, 2015, D'Angelo and Begel, 2017; D'Angelo and Gergle, 2016). In most of these studies, the basic outcome or the working hypothesis is that Joint Visual Attention (JVA) is a decent proxy of collaborative mechanisms. All of these studies emphasize a social extension of the eye-mind hypothesis, "what you see is what you process", to "looking together is processing together". However, this notion has not been verified in some studies over the past few years (e.g., Belenky et al, 2014). In this contribution, we revisit the concept of JVA and complement it with another DUET measurement, Joint Mental Effort (JME). This measurement is inspired by the Kirscher's view (Kirschner et al., 2018) of how transactive activities can exert cognitive loads on collaborating peers and that the absence of synchrony in the collaboration can be detrimental for collaborative performance (Popov et al., 2017). JME provides an attempt to create a proxy for the collaborative cognitive load synchrony.

Specifically, in this paper, we investigate what JVA and JME, both collected through eye tracking, indicate about the collaborative learning process. We analysed 82 master students working in pairs to construct a concept map related to the resting membrane potential. We were interested in how their collaborative process impacted the quality of their concept map. To measure the collaborative process, we collected student dialogues, eye tracking data and computer logs. In this paper, we aimed to answer two research questions through our analysis. First, how did our eye tracking measures (JVA and JME) relate to student performance? Second, how do JVA and JME relate to other indicators of collaboration, such as dialogue content and division of labor, and how do the interactions with student performance associate with JVA and JME? Based on previous studies, we hypothesize that JVA will be positively related to student performance (Richardson et al., 2007; Jermann and Nuessli, 2012) and that JME also will be positively associated with performance (Kirschner et al, 2018; Popov et al, 2017). Based on the results of these research questions, we discuss the benefits of using eye tracking measures to assess the collaborative learning process.

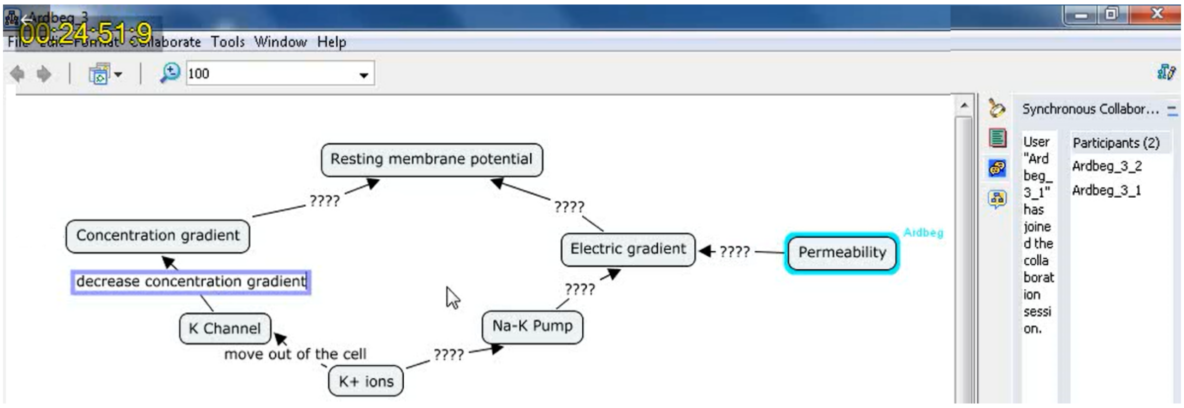
Methodology

Participants and procedure

We had 82 master students from École Polytechnique Fédérale de Lausanne participating in the present study in pairs. Of these students, 16 were female. Before beginning to create the concept map collaboratively, all participants individually watched two videos about "resting membrane potential", a topic about which the students did not know prior to participating in the task. Each video was 17.05 minutes long and provided the students with the information they would need for the development of the concept maps. While watching the videos, the participants had full control over the video player and no time constraints. After both partners completed the videos, they were asked to create a collaborative concept-map using IHMC CMap tools (Figure 1, top). The collaborative concept-map phase was 10-12 minutes long. Although each student remained working at their own computer, the participants could talk to each other while their screens were synchronized, i.e., the participants in the pair were able to see their partners' actions. There were 14 concepts preloaded on the Concept map tool and the main task for the pairs was to connect the given concepts with correct relationships. They could also add new concepts if they wanted.

Data collection

From the interaction of the dyad with the concept map tool, we collected the following data. 1) We collected eye-tracking data using two SMI remote eye-tracking devices (SMI RED 250) at the sampling rate of 250 Hz. For each participant, we use a 5-point calibration and a 5-point validation mechanism. The fixation and saccades were identified using the built-in algorithm of the BeGaze software. 2)We recorded the audio of the students' dialogues using the system audio from one of the computers. 3) We recorded all the actions done by the dyad on the concept map. The logs contained the timestamp of the action, peer ID, action type (add, delete, move, resize, text edit), conceptID and metadata (Figure 1, bottom). For example, if a student adds two concepts with a link. The system would log the time the action took place, the ID of the students, the action is logged as an "add" action, the object will be the "connection", there will be a new ID generated for this connection, and the metadata would log the two concepts it linked.



timestamp	SubjectID	Action	Object	ObjectID	Metadata
Mon Oct 28 19:11:30:168	Clynelish_1_1	Move	Concept	ge:1M5NG8B0L-1BJ5XH8-B9	'Resting membrane potential' x:134 y:43 w:188 h:28
Mon Oct 28 19:11:52:232	Clynelish_1_2	Add	Connection	ge:1M88V4FYD-3G09QV-PT	from:ge:1M88V4B39-1JK1H3M-KR to:ge:1M88V4FY8-G83G3M
Mon Oct 28 19:11:53:272	Clynelish_1_2	Move	Concept	ge:1M88V4FY8-G83G3M-PH	'????' x:516 y:280 w:44 h:26
Mon Oct 28 19:12:23:356	Clynelish_1_1	Add	Concept	ge:1M88V5H7G-MLDB4J-2TH	'????' x:488 y:271 w:44 h:26
Mon Oct 28 19:12:24:751	Clynelish_1_1	Delete	Concept	ge:1M88V5H7G-MLDB4J-2TH	'????' x:488 y:271 w:44 h:26
Mon Oct 28 19:12:31:756	Clynelish_1_1	Modify Text	Linking Phrase	ge:1M88V4B39-1JK1H3M-KR	'' x:512 y:282 w:4 h:18
Mon Oct 28 19:12:31:845	Clynelish_1_1	Move	Concept	ge:1M5NG9QK9-Y0RB8H-FX	'Cl channel' x:469 y:325 w:80 h:26

Figure 1. (Top) An example of the concept map under construction in the CMap tools. The two participants' names are on the top-right side, and their pointers have different colors. Whenever they perform an action, the relevant object (concept or link) is highlighted. (Bottom) Snapshot of the log file produced from CMap.

Measurements

All the measurements were computed at the dyad level, the time-unit for each computation was one utterance, and all the measurements were aggregated for the dyad.

Cognitive load similarity (CLS): Joint mental effort (JME)

From the eye tracking data, we calculated the students' JME, *a measure of the cognitive load similarity*. To calculate this measure, we first compute the individual cognitive load from the pupil dilation data using the method found in Duchowski et al. (2018). Next, we discretize the value to represent an integer value in the range zero to ten. Once we have the cognitive load for both peers in the dyad, we compute the cross-recurrence between the two time-series, using the method proposed by (Richardson et al., 2007).

Gaze similarity (GS): Joint visual attention (JVA)

JVA is a measure of *how similar two individual gaze patterns are.* In order to compute the similarity between the gaze patterns of two collaborating students, we computed the similarity between the two proportionality vectors discussed above by using the reverse function (1/(1+x)) of the correlation matrix of the two vectors (where x is the distance between the two proportionality vectors). A similarity value of 1 shows complete similarity between the two gaze patterns during a given time window. A lower value of similarity shows that the two participants spent less time looking at a similar set of objects on the screen during the same time window.

Dialogue codes

One of the authors transcribed the audio data and two authors coded the dialogues. The intercoder-consistency between the two coders was 0.86 (for 20% dyads). The dialogues were coded based on the fact whether the dyad is talking about the concept map tool and aesthetics (CMAP) or about the content of the concept map (KNWL). For example, *"Let's write something to remove the question marks"* would be coded as CMAP, and *"Resting membrane potential is the equilibrium between Na+ and Cl-" would* be coded as KNWL.

Division of labor (DoL)

Following a definition provided by Jermann (2004), we compute the division of labor using the number actions taken on a specific concept by one member of the dyad. Specifically, we compute the Sum of Differences (SD) and Sum of absolute differences (SAD) between members of a dyad using the formulae (1) and (2). Using thresholds on SD and SAD, we define three DoL levels, role, task and concurrent, which we outline in more detail below.

$$SD = \frac{\sum_i (S_1 C_i - S_2 C_i)}{S_1 C + S_2 C} \ (1) \quad SAD = \frac{\sum_i |S_1 C_i - S_2 C_i|}{S_1 C + S_2 C} \ (2)$$

In formulae (1) and (2), S1 and S2 are the peers in a dyad. C is the concept. S1C and S2C are the total number of actions done by peers S1 and S2, respectively. S1Ci and S2Ci are the actions done on concept Ci by S1 and S2, respectively. SD has a range of [-1, +1] with -1 indicating that S2 did all the actions, +1 indicating that S1 did all the actions and 0 depicting equal participation. SAD has a range of [0, 1] with 0 indicating equal participation and 1 indicating that all the actions were done by one peer.

We defined the DoL strategies – role, task and concurrent – based on SD and SAD values. The DoL strategy is classified as *role* if SAD is in the range [0.5, 1] and SD in either [0.33, 1] or [-1, -0.33] indicating that one student did all of the actions within a certain time window - implying the other student was either a free-rider or acting as a navigator. The DoL strategy is classified as *concurrent* if SAD is in the range [0, 0.5] and SD in range [-0.33, 0.33] during the time window, indicating that the students had equal participation on the same concepts. Finally, the DoL strategy is classified as *task* if SAD is in the range [0.5, 1] and SD in either [-0.33, 0.33] during the time window indicating that the students were each participating in taking actions on the concept map, but on different concepts.

Learning performance: Correctness of the concept map

The learning performance for this activity is the correctness of the concept map. We asked two domain experts to create a map using the same 14 concepts. All the participant maps were compared against this expert map. We followed the following map-evaluation scheme: 1) 2 points for correct link and correct label; 2) 1 point for correct link and no label; and 3) 0.5 point for correct link and incorrect correct label. We added the points for each link between all the concepts and that was the dyad's performance score. Finally, we applied a median split to divide the dyads into high and low performance groups.

Data analysis

To examine the direct relationship between eye tracking measures and performance, dialogue codes, and DoL, we used a set of ANOVAs. We tested for the normality and homoscedasticity conditions using Shapiro-Wilk and Bausch-Pegan tests, respectively. In the case where the normality was violated, we normalized the data, and in cases where the homoscedasticity was violated, we used a Welch correction. We also tested for the pairwise interaction for all the variables using ANOVA. For the post hoc pairwise tests we applied Bonferroni corrections. We also computed the Cohen's d as the effect size for each ANOVA calculation. According to Cohen effect sizes can be low (below 0.2) medium (0.2 - 0.8) and high (above 0.8).

Results

JVA and JME relation with performance levels

To answer our first research question, we analysed the relationship between JVA and JME and the performance level of our dyads on their concept map. For both measures, we observed significant associations with performance level (see Figure 2, left column). The JVA for high performing dyads is significantly higher than the JVA for the low performing dyads ($F(1,38) = 18.67$, $p < .0001$, $d = 0.65$). Similarly, the JME for high performing dyads is significantly higher than the JME for the low performing dyads ($F(1,35.27) = 23.91$, $p < .0001$, $d = 0.81$).

JVA and JME relation with other process variables

To answer the first part of our second research question, we investigated how JVA and JME relate to the other collaborative process measures (i.e., dialogue content and division of labor). We observed significant associations between the DoL strategies with both JVA ($F(2,37) = 25.21$, $p < .0001$, $d = 0.87$) and JME ($F(2,32.45) = 8.29$, $p < 0.01$, $d = 0.24$). As seen in the middle column of Figure 2, JVA is highest when the students are engaged in role division compared to concurrent ($F(1,38) = 16.89$, $p < .01$, $d = 0.58$) or task ($F(1,38) = 27.49$, $p < .0001$, $d = 0.92$). Additionally, JVA is higher for a concurrent division than a task ($F(1,38) = 24.35$, $p < .001$, $d = 0.84$). Similarly, JME is highest for a role division compared with concurrent ($F(1,35.59) = 17.01$, $p < .01$, $d = 0.62$) or task ($F(1,31.35) = 28.33$, $p < .001$, d = 0.95), and task is also lower than concurrent ($F(1,33.24) = 22.43$, $p < .001$, $d = 0.78$).

Further, we found a significant relationship between both JVA and JME and the dialogue codes. In the right column of Figure 2, we see that JVA is significantly higher during concept-map dialogues than knowledge dialogues ($F(1,38) = 11.17$, $p < .001$, d = 0.38). We found the opposite for JME with JME being significantly higher during knowledge dialogues than concept-map dialogues ($F(1,36.83) = 31.29$, $p < .0001$, d = 1.03).

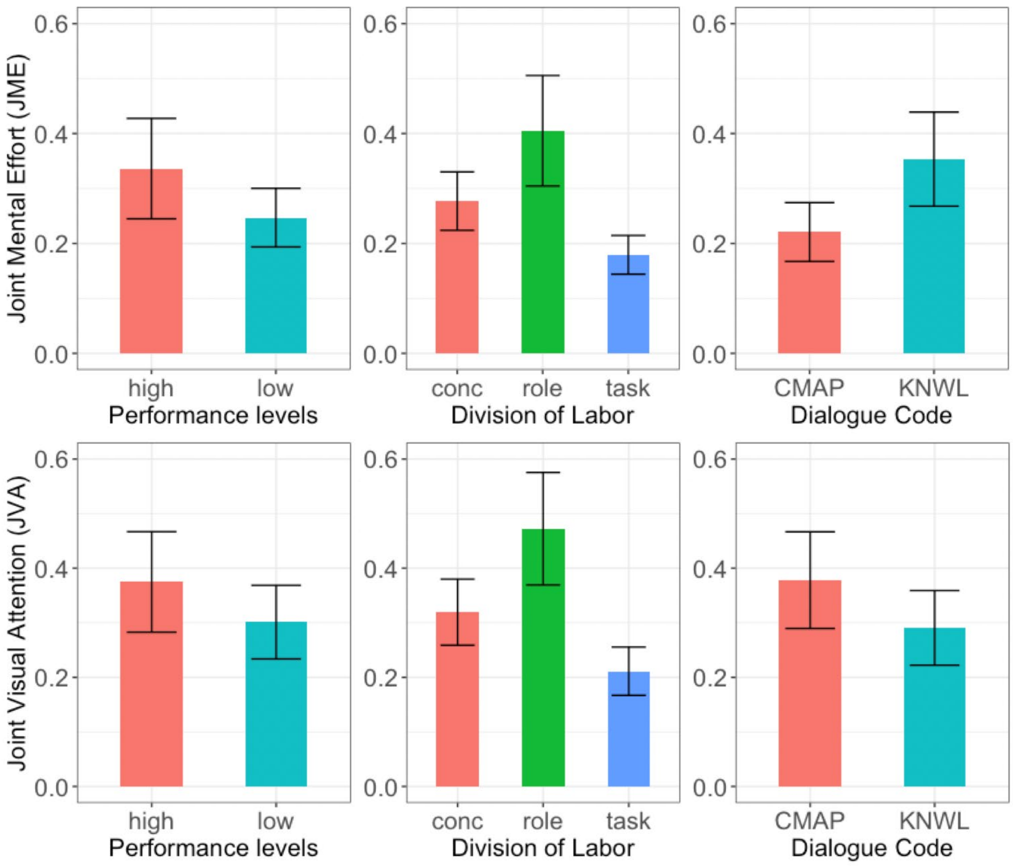


Figure 2. Comparing Joint Mental Effort (top panels) and Joint Visual Attention (bottom panels) across the different categories of Performance (left, high/low), division of labor (middle, concurrent/role/task), and dialogue codes (right, CMAP/KNWL).

Interaction effect between process variables and performance on eye-tracking

Finally, to answer the second part of our second research question, we analysed the impact of the interactions between the different process variables and their relation with the performance outcome. Concerning the interaction effects on JVA, we found a significant interaction between the performance levels and DoL strategies on JVA ($F(1,38) = 20.79$, $p < .0001$, $d = 0.72$). As we can see in Figure 3, the low performing students have a relatively stable JVA across the three DoL strategies. On the other hand, the high performing students fluctuated to have no significant difference with concurrent, higher JVA for role ($F(1,35.56) = 13.11$, $p < .0001$, $d = 0.47$), and lower JVA for task (d = $F(1,32.72) = 9.38$, $p < .0001$, $d = 0.38$) compared to the lower performing students. In terms of dialogue, we did not observe any interaction effect of performance and the dialogue category (CMAP/KNWL) on the JVA.

As with the JVA, we found an interaction between performance levels and DOL strategies for JME ($F(1,38) = 9.56$, $p < .001$, $d = 0.29$). There is not a significant difference between high and low performing dyads during concurrent and task divisions, but the difference is significant between high and low performing students during role divisions ($F(1,22.23) = 11.23$, $p < .0001$, $d = 0.39$) as seen in Figure 3. We did not observe a significant interaction between performance levels and the dialogue codes (CMAP/KNWL) on JME.

Discussion and conclusion

In this paper, we aimed to explore alternate measures of the collaboration processes, namely, ones gathered through dual eye tracking. As with other multi-modal studies, we are interested in the additive property that analysing the collaborative process from multiple perspectives can provide. With that in mind, in this discussion, we will present different interpretations of the data and how these interpretations are narrowed down as we add new measures, indicating the additive property of the measures rather than providing a set of proxies.

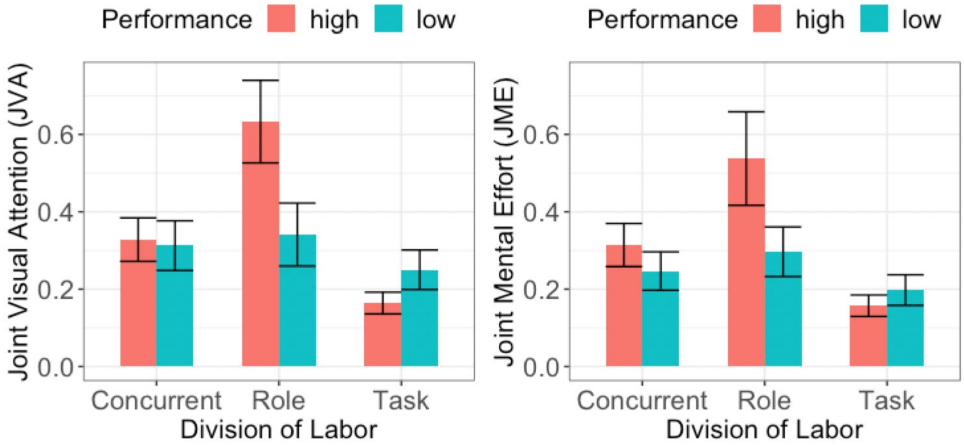


Figure 3. Interaction effect of performance levels and division of labor strategies on joint visual attention (left) and joint mental effort (right).

In terms of our first research question, how did the eye tracking measures relate to student performance, we confirmed our hypothesis that both higher JVA and JME were positively related with higher performance. This finding provides more of a confirmation of previous findings (Sangin et al., 2011; Jermann and Nuessli, 2012) than necessarily providing a new insight on its own. However, we would like to highlight that although researchers have studied the impact of individual cognitive load extensively (Amadieu et al, 2009, Kalyuga, 2011), using a joint measure to assess the collaborative process is relatively new. What this means is that students that are putting in the same amount of mental effort at the same time are more likely to perform well, and it is not just about the amount of mental effort of an individual student. The positive relation between JME and performance in a task where the transactive actions are taking place (division of labor, communication, and coordination of activities, Popov et al., 2017), gives an indication towards JME being a decent proxy for collaborative cognitive load. Nonetheless, it is an early indication and further studies are required for generalizability.

In terms of our second research question, how do JVA and JME relate to other process variables and what is the interaction when we include performance, it is interesting to discuss the results in terms of the division of labor measure. First, let's look at the concurrent division. Recall from our description of this measure that the concurrent division occurs when each member of the dyad is working on the same concepts in the same time window (although not necessarily trying to take the same actions). From our analysis, we found that during concurrent division, students had both JVA and JME measures somewhere between those of the role and task divisions. This might be expected as the students are looking in the same general area but, as they are doing separate actions, this overlap does not mean that they are necessarily working together. However, because the JME is also in between, it is unlikely that one student is just doing aesthetic changes while the other is enhancing the concept diagram. Although at the surface level, we may want to classify concurrent division as weak collaboration, as the students are doing separate actions, the eye tracking measures indicate that this is not always the case and warrants further investigation as to what occurred in the collaboration process prior that led to this division of labor before determining that an intervention is needed.

In contrast, during the role division, one student is doing all of the actions during a time frame. This pattern may be due to one partner free-riding (Le et al., 2018) or due to a productive driver/navigator collaboration (Bryant et al., 2006). If the students are focused on the same thing (high JVA) and are putting in the same mental effort (high JME), it is more likely to be a productive collaboration than a student free-riding. Although we found high JVA and JME in general during the role division, this was not the case when we took into account student performance. In this case, low performers had significantly lower JVA and JME, indicating that this may be instances of free-riding. Unlike with the concurrent collaboration, the use of roles is often considered as a productive collaboration script (King, 1999). Nonetheless, roles alone do not guarantee interdependence, and these may be moments for a clear intervention.

During the task division, students are working on different concepts during the same time window. As a first interpretation, this may mean that the students have divided the work evenly and are each working on a different part. However, it may also indicate that one student is doing the majority of the work while the other is making aesthetic changes – like when one participant writes a paper and the other corrects typos. From the JVA, we cannot differentiate these actions, as the JVA is low, as expected, due to the students working on different parts of the map. The JME can provide more insight though. We might expect the JME to be high if the students have an equal division of labor. In our case, we found the JME to be low, most likely indicating that there was not an equal divide. Perhaps more surprisingly, the JME was not different for high and low performers, but the JVA was. This may indicate that although in both cases the division of labor may not have been even, the high performers may have had more confidence carrying out their tasks independently. Future work would be needed to assess the exact task division and how these actions fit into the students' broader collaboration processes.

Finally, we observe that the JVA and JME have an opposite relationship with the dialogue, i.e., whether the dyad is talking about the interface or the domain knowledge. JVA is higher for the interface-based dialogues while the JME is higher for the knowledge dialogues. This indicates the complementary nature of two gaze measurements. The JVA is higher when there is strong visual support to ground the verbal references and JME is higher when the discussion is focused on domain knowledge. If we were to ignore one of these measurements, we would have received only half the picture (either in terms of attention management or in terms of effort management). Moreover, the high performing dyads have both JVA and JME that are higher than the low performing dyads, showing that, in terms of both their attention and effort management, the high performing dyads have more of an equal participation than the low performing dyads. This could possibly lead to better task-performance at the end of the collaborative session.

Although this paper presents a first analysis of how different types of eye tracking measures can shed light on the collaborative process, there are still several limitations. First, we only have an end performance measure. It is not clear if the collaborative process patterns we see are due to the knowledge of the students when they begin or if these actions led to better learning thereby leading to a better performance. Second, there are many combinations of process variables and temporal aspects that we did not explore in this paper that would provide further insights into the collaborative process and how it impacts JVA and JME. Due to space, we could not include them all.

In this paper, we aimed to deepen our understanding of what effective collaboration looks like through the assessment of dual eye tracking measures. We found that an effective collaboration is not necessarily a one-size-fits-all where a single metric can be used to judge the quality of the collaboration. Further, we found that joint mental effort can provide additional information than joint visual attention alone to better assess the collaborative process, contributing to the use of dual eye tracking methodology.

References

Amadieu, F., Van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction, 19*(5), 376-386.

Belenky, D.M., Ringenberg, M.,Olsen, J.K., Aleven, V., & Rummel, N. (2014). Using dual eye-tracking to evaluate students collaboration with an intelligent tutoring system for elementary-level fractions. In the *Procs. of the 36th Annual Conf. of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition, 106*(3), 1465-1477.

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.

Bryant, S., Romero, P., & du Boulay, B. (2006). The collaborative nature of pair programming. In *International Conference on Extreme Programming and Agile Processes in Software Engineering* (pp. 53-64). Berlin, Heidelberg: Springer.

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 the 2008 symposium on Eye tracking research & applications* (pp. 173-180).

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist, 49*(4), 219-243.

D'Angelo, S., & Begel, A. (2017). Improving communication between pair programmers using shared gaze awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 6245-6290).

D'Angelo, S., & Gergle, D. (2016). Gazed and confused: Understanding and designing shared gaze for remote collaboration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2492-2496).

Duchowski, A. T., Krejtz, K., Krejtz, I., Biele, C., Niedzielska, A., Kiefer, P., ... & Giannopoulos, I. (2018). The index of pupillary activity: Measuring cognitive load vis-à-vis task difficulty with pupil oscillation. In *Procs. of the 2018 CHI Conf. on Human Factors in Computing Systems* (pp. 1-13).

Hausmann, R. G., Chi, M. T., & Roy, M. (2004). Learning from collaborative problem solving: An analysis of three hypothesized mechanisms. Proceedings for *26nd Annual Conference of the Cognitive Science society* (pp. 547-552). Mahwah, NJ: Erlbaum.

Jermann, P. (2004). *Computer Support for Interaction Regulation in Collaborative Problem-Solving* [Unpublished doctoral dissertation]. l'Université de Genève, Switzerland.

Jermann, P., & Nüssli, M. A. (2012). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. In *Procs. of the ACM 2012 Conf. on Computer Supported Cooperative Work* (pp. 1125-1134).

Johnson, D. W., & Johnson, R. T. (1987). Research shows the benefits of adult cooperation. *Educational leadership, 45*(3), 27-30.

Joshi, M., & Rosé, C. P. (2007). Using transactivity in conversation for summarization of educational dialogue. In *Workshop on Speech and Language Technology in Education*.

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review, 23*(1), 1–23.

King, A. (1999). Discourse patterns for mediating peer learning. In A.M. O'Donnell, A. King (Eds.), *Cognitive Perspectives on Peer Learning* (pp. 87–117). Lawrence Erlbaum Associates, Mahwah, NJ.

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior, 25*, 306–314.

Le, H., Janssen, J., & Wubbels, T. (2018). Collaborative learning practices: teacher and student perceived obstacles to effective student collaboration. *Cambridge Journal of Education, 48*(1), 103-122

Lemaignan, S., & Dillenbourg, P. (2015). Mutual modelling in robotics: Inspirations for the next steps. In *10th ACM/IEEE Intl. Conf. on Human-Robot Interaction (HRI)* (pp. 303-310). IEEE.

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of educational research, 71*(3), 449-521.

Mercer, N. (2008). The seeds of time: Why classroom dialogue needs a temporal analysis. *The Journal of the Learning Sciences, 17*(1), 33–59.

Olsen, J., Sharma, K., Aleven, V., & Rummel, N. (2018). Combining gaze, dialogue, and action from a collaborative intelligent tutoring system to inform student learning processes. Proceedings of *the 13th International Conference of the Learning Sciences* (pp. 689–696). London, UK: ISLS.

Olsen, J. K., Sharma, K., Rummel, N., & Aleven, V. (2020). Temporal analysis of multimodal data to predict collaborative learning outcomes. *British Journal of Educational Technology, 51*(5), 1527-1547.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination. *Psychological science, 18*(5), 407-413.

Rummel, N., Deiglmayr, A., Spada, H., Kahrimanis, G., & Avouris, N. (2011). Analyzing collaborative interactions across domains and settings: An adaptable rating scheme. In *Analyzing interactions in CSCL* (pp. 367-390). Springer, Boston, MA.

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*(3), 1059-1067.

Sangin, M. (2009). *Peer Knowledge Modeling in Computer Supported Collaborative Learning* [Unpublished doctoral dissertation]. École polytechnique fédérale de Lausanne, Switzerland.

Schneider, B. (2019). Unpacking collaborative learning processes during hands-on activities using mobile eye-trackers. Proceedings of *the 13th International Conference Computer-Supported Collaborative learning* (pp. 41-48). Lyon, France: ISLS.

Schneider, B., & Pea, R. (2015). Does seeing one another's gaze affect group dialogue? A computational approach. *Journal of Learning Analytics, 2*(2), 107-133.

Sharma, K., Jermann, P., Dillenbourg, P., Prieto, L. P., D'Angelo, S., Gergle, D., ... & Rummel, N. (2017). Proceedings of *the 12th International Conference Computer-Supported Collaborative learning*. Philadelphia, PA: ISLS.

Sharma, K., Chavez-Demoulin, V., & Dillenbourg, P. (2018). Nonstationary modelling of tail dependence of two subjects' concentration. *The Annals of Applied Statistics, 12*(2), 1293-1311.

Sharma, K., Leftheriotis, I., & Giannakos, M. (2020). Utilizing Interactive Surfaces to Enhance Learning, Collaboration and Engagement: Insights from Learners' Gaze and Speech. *Sensors, 20*(7), 1964.

Starr, E. L., Reilly, J. M., & Schneider, B. (2018). Toward Using Multi-Modal Learning Analytics to Support and Measure Collaboration in Co-Located Dyads. Proceedings of *the 13th International Conference of the Learning Sciences* (pp. 448–455). London, UK: ISLS.

“Houston, We Have a Problem!” Homogeneous Problem Perception, and Immediacy and Intensity of Strategy Use in Online Collaborative Learning

Martin Greisel, University of Augsburg, Germany, martin.greisel@uni-a.de
Laura Spang, University of Augsburg, Germany, laura.claudia.spang@uni-a.de
Kerstin Fett, University of Augsburg, Germany, kerstin.fett@uni-a.de
Nadine Melzner, Technical University of Munich, Germany, nadine.melzner@tum.de
Markus Dresel, University of Augsburg, Germany, markus.dresel@uni-a.de
Ingo Kollar, University of Augsburg, Germany, ingo.kollar@uni-a.de

Abstract: Under socially distant circumstances, university students frequently self-organize to collectively prepare for exams online through video chat. To learn effectively, emerging challenges need to be regulated successfully. This regulation is supposed to work best when problems are perceived homogeneously in the group, and when regulation strategies which immediately solve the problem are chosen and executed with sufficient intensity. We investigated what problems occur during collaborative online learning and how these are regulated by N=222 university students in 106 groups. We found that overall problem prevalence was low. Multilevel-modeling indicated that homogeneous problem perception—contrary to immediate and intensive strategy use—predicted subjective learning success, while objective learning success was not associated. Thus, in well-structured learning contexts, knowing what the problem is seems to be more important than knowing the best possible reaction to the problem. Students might be trained in problem perception in order to increase regulation competency.

Problem statement

Many students deliberately join together in self-organized small groups, e.g. to prepare for exams together. Taking positive effects of collaborative learning on knowledge acquisition found in the literature into account (e.g., Springer et al., 1999), this is a sensible decision. However, collaborative learning unfortunately is not always as effective (Weinberger et al., 2012). In fact, students may be confronted with a variety of problems during collaborative learning that are obstacles to effective learning (Järvenoja et al., 2013). This is also true for online collaborative learning, where learners are often frustrated due to various problems such as an imbalance in commitment, unshared goals or communication difficulties (Capdeferro & Romero, 2012). Only if the group is able to regulate these problems successfully, collaborative learning is effective (Järvelä & Hadwin, 2013).

The ability to regulate occurring problems independently of any instructional support is very important for regulation success especially for students outside formal instructional contexts, who form learning groups on their own initiative. Thus, acquiring necessary regulation skills beforehand is crucial for regulation success during periods of self-organized collaborative learning. To foster these skills, scientific knowledge is needed on how problems are regulated best in such situations. Further, the context how the meeting takes place might be relevant, too: When self-organized study groups cannot meet in person (e.g., at institutions for distance learning, in areas with large physical distances between students, or during times of a pandemic), collaborative learning typically happens online through video conference tools such as Zoom or Skype. Yet, not much is known about how this virtual context influences processes associated with specifically the regulation of problems during self-organized study group meetings. Therefore, this study focuses on how problems are regulated in virtual collaborative learning through video conferencing.

Regulation of problems in collaborative learning

Based on previous research (e.g., Järvenoja et al., 2019), problems in self-organized collaborative learning can be divided into at least the following categories: (a) comprehension problems (e.g., learners may have difficulty understanding the task), (b) coordination problems (e.g., learners may have different objectives for learning together), (c) motivation problems (e.g., the learning material may be perceived as irrelevant) and (d) resource-related problems (e.g., necessary learning material may not be available). For self-organized collaborative learning to be successful, groups must be able to cope with such problems successfully.

To conceptualize the processes involved in this problem regulation, we (Melzner et al., 2020) developed a heuristic process model (see Fig. 1). Following process models of self-regulated learning (e.g., Zimmermann &

Moylan, 2009), metacognitive processes are crucial for the successful regulation of problems in self-organized collaborative learning, with the help of which students (1) perceive and classify these problems. Based on the assessment of a problem, a reaction is initiated to ensure that the goal is achieved despite the problem at hand. For this purpose, students (2) select a strategy to address the problem and (3) execute this strategy with a certain intensity. Once the problem is solved, the learning process can be continued. Along with Melzner et al. (2020), we assume that these three processes (problem perception, choice of regulation strategy, intensity of strategy execution) should predict success in the regulation of problems that occur during collaborative learning.

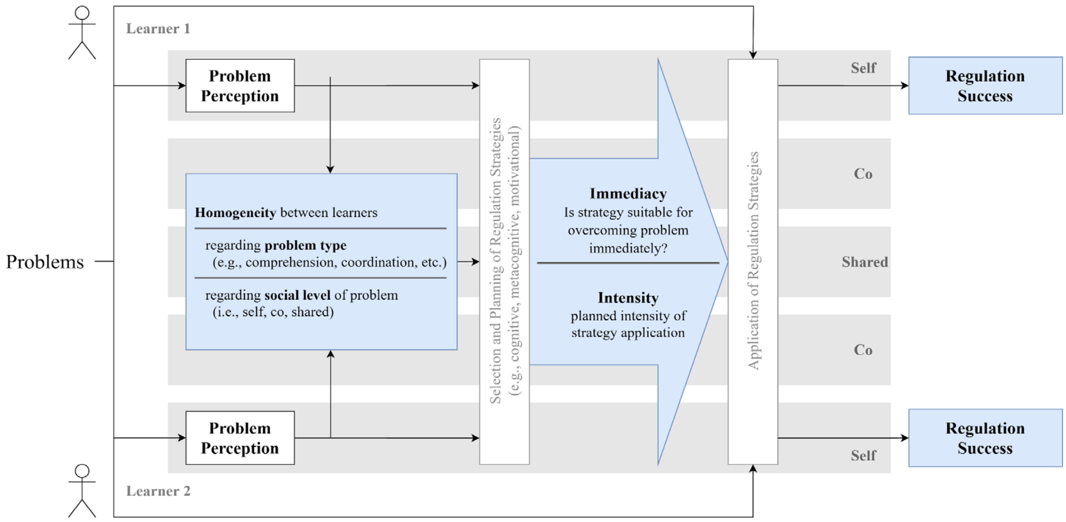


Figure 1. Theoretical model of the regulation of problems during collaborative learning (visualization inspired by Wecker and Fischer, 2014). Concepts in boldface are measured in the present study. Adapted by permission from Springer Nature: IJCSCL. Regulating self-organized collaborative learning: The importance of homogeneous problem perception, immediacy and intensity of strategy use. Melzner, N., Greisel, M., Dresel, M., & Kollar, I. (2020). https://doi.org/10.1007/s11412-020-09323-5

Homogeneity of problem perception

At the beginning of the regulation process, learners perceive and classify a given problem (see Fig. 1). Different group members may arrive at different problem assessments. Divergences can basically be based on two dimensions: First, the type of problem (see e.g., Järvenoja et al., 2013) that is perceived may vary. For example, while one learner may perceive a comprehension problem to be present, another learner may categorize this problem as motivational. On the other hand, there may also be disagreement about the social level at which the problem is located. Using the classification of Järvelä and Hadwin (2013), it can be distinguished whether a learner is affected himself (self-level), whether the problem affects individual other group members (co-level), or whether the whole group is affected (socially shared level). The homogeneity of the problem perception is thus to be understood in terms of (a) the type of problem and (b) the question who is affected by the problem. We suspect that diverging perceptions of the problem within the group make collaborative learning more difficult, since the individual group members are then more likely not to coordinate their regulation efforts. Findings of Melzner et al. (2020) corroborate this.

Immediacy of regulation strategy use

Next, learners select a strategy for the regulation of the previously perceived problem (see Fig. 1). Models of self-regulated learning (e.g., Zimmermann & Moylan, 2009) assume that at this point, the choice of a strategy that fits the learning goal is crucial. Not every strategy is supposed to be equally well suited to achieve a particular goal (e.g., Engelschalk et al., 2016). In our view, a similar assumption may be made regarding the fit between an emerging problem and the chosen strategy for its regulation (e.g., Engelschalk et al., 2016). However, previous research has hardly made statements about what is meant by *fit*. In order to operationalize fit, we have proposed the concept of *immediacy* (Melzner et al., 2020): A strategy can be considered to be appropriate for a problem if it is in principle possible to actually solve the problem when the respective strategy is executed optimally. An example of an immediate strategy would be to switch off cell phones when the group is distracted by incoming messages during learning. An example of a non-immediate strategy, on the other hand, would be if learners make themselves aware of the importance of the exam they are preparing for in order to motivate them to continue learning despite the incoming messages. This strategy would not eliminate the source of distraction and thus would not immediately make the problem disappear, but would only allow learners to continue learning despite the presence of the problem. Thus, for the operationalization of fit, a theoretical assignment of strategies to problems as immediate or non-immediate was proposed by Melzner et al. (2020) and was found to predict satisfaction with the group learning experience in completely self-organized, offline groups.

Intensity of the execution of the regulation strategy

To be effective, the selected strategy must be applied in the next step (see Fig. 1). Depending on the severity of the problem, however, a single application of the strategy may not be sufficient to achieve the desired effect. For example, if learners bored by the learning materials think only briefly about their goals for the future, this may have little effect on their motivation to devote effort towards understanding the material. However, if they work intensively on how the material will help them to achieve their own goals, this should increase their motivation. We therefore assume that the intensity of strategy use is positively related to regulation success. However, not only the intensity of immediate strategies should be relevant, since non-immediate strategies might also increase regulation success, even if the specific problem is not solved that way. Findings on the effect of regulation intensity are mixed (Eckerlein et al., 2019; Melzner et al., 2020; Schoor & Bannert, 2012). Thus, more research is needed to clarify its influence on regulation success.

Operationalizing regulation success in collaborative learning

Once the regulation process is executed in accordance with Fig. 1, it should be successful. Yet, regulation success may be conceptualized and measured in various ways (e.g., Melzner et al., 2020; Noroozi et al., 2019; Zimmermann & Moylan, 2009). In this paper, we focus on three different conceptualizations: (1) success in applying a regulatory strategy (i.e., the extent to which the problem is overcome after the strategy is applied), (2) satisfaction with the group learning experience, and (3) the subjective and objective learning success resulting from the group learning session. So far, only satisfaction was empirically investigated in this context (e.g., Melzner et al., 2020; Bellhäuser et al., 2019). Yet, not much is known about how problem perception, immediacy and intensity of strategy use contribute to further measures of regulation success.

Research questions and hypotheses

The present study addresses two research gaps: First, it is an open question to what extent the three processes (homogeneity of problem perceptions, immediacy of strategy use, and intensity of strategy use) would be predictive of successful regulation in collaborative online settings. Second, little is known about whether the three processes are differentially predictive of the three conceptualizations of regulation success described above. Therefore, we established the following hypotheses:

1. The more homogeneous learners perceive problems within their groups, the more positive the results on different measures of regulation success are.
2. Learners who use immediate strategies to regulate their problems achieve more positive results on different measures of regulation success than learners who use only non-immediate strategies.
3. The more intensively learners apply regulation strategies, the more positive the results on different measures of regulation success are.

Method

Sample

University students (N = 222) from two basic psychological lectures within the majors educational sciences (29%) and teacher training (70%) answered an online questionnaire. They had an average age of 22 years (M = 21.84, SD = 4.39, 83% female) and were on average in the third semester of their current study subject (M = 2.78, SD = 1.50) and also in their third university semester overall (M = 3.34, SD = 2.57). Participants self-selected into 106 small groups of three persons on average, but not all members of each group participated in the study. Thus, data from 25 groups which were represented in our data by a single person only had to be excluded from regression analysis because a calculation of homogeneity of problem perception only is possible for groups with data of two or more learners.

Procedure

The study was embedded in two large lectures which mainly consisted of weekly uploaded recordings of PowerPoint-presentations provided for individual, asynchronous studying. One session of collaborative learning replaced the regular lecture in the respective week. Learners were instructed to meet online at a time suitable for all group members using a video conference software of their choice to study the lecture content on their own. As learning material, the regular slide deck for this session was provided alongside two excerpts from a textbook, each about one page long. Topics were the ICAP-Model of learning activities (Chi & Wylie, 2014) and the multi-store model of memory (Atkinson & Shiffrin, 1968). We did not structure or scaffold the collaborative learning with additional instructions except the following tasks: “The goal of the group work is to work out the slide contents as well as possible together with your group members. You are welcome to use the additional texts provided.” In addition, students were told to record the results of their group work in a shared concept map. Yet, besides this, learners were free to decide in which way, with which activities or tools, they wanted to work on the topic. For learners who were not familiar with an online tool suitable to produce a concept map, we recommended www.mindmeister.com and provided a short tutorial video explaining all functions necessary for accomplishing the task.

After the study meeting, participants were asked to individually answer an online questionnaire. The questionnaire was advertised as containing a knowledge test for which students would receive immediate feedback regarding right and wrong answers. The questions were comparable to the ones in the final exam in the corresponding lectures, so taking the test would be a good chance to practice for the “real” exam.

Measures

To measure the *prevalence of problems during collaborative learning*, we developed a questionnaire with 32 different problems represented by three items each. Each item had to be rated on a Likert-scale (from 0 = *did not occur/no problem* to 4 = *big problem*). Based on problem typologies or theoretical classifications in the literature (e.g., Järvenoja et al. 2013; Koivuniemi et al., 2017), our questionnaire covered four broad categories of problems: comprehension, coordination, motivation, and resources (see Fig. 2 for a complete list of individual problems). For example, for the problem of “low value of learning method”, a sample item was “Single/multiple group members did not find group work as a learning method useful in the given situation.” An extensive series of confirmatory factor analyses comparing the theoretical factor structure to other theoretical plausible clusterings of items indicated that the theoretical factors with three items per factor were distinguishable from each other, and that the theoretical solution has the best fit to the data. Cronbach’s alpha was .79 on average. After rating each problem, participants selected one of them as the biggest problem they encountered during the learning session.

To determine the *homogeneity regarding the type of problem within each group*, we calculated the variance within each group for each rated problem separately, and then determined the average variance per group over all problems. To transform the variance into a measure of homogeneity, we multiplied it by −1 and centered it. To determine the *homogeneity regarding the social level*, we used three items measuring the extent to which the biggest problem affected the self-, co-, or shared-level on a five-point Likert-scale (from *not at all true* to *completely true*). A sample item representing the self-level was: “The mentioned problem had effects on my personal learning process.” The ratings for each item were dichotomized by median split, resulting in a zero-one-coding. Then, groups were coded as being homogeneous regarding the social level of problem perception when the social level at which they located the biggest problem matched the respective ratings of each other group member. For example, a group was considered to be homogeneous when one person located the problem only at the self-level, while the two other group members located the problem only at the co-level.

To measure *immediacy and intensity of strategy use*, we asked participants to name the strategies they used to regulate the problem they marked as the biggest one at the self-, co- and shared level in an open answer format (e.g., at the self-level: “What did you personally think, do, or say to ensure high quality of your own learning in this situation?”; at the shared level: “What did you as a group think, do, or say to ensure high quality of the learning of the whole group in this situation”). These answers were segmented into single regulation strategies (interrater-agreement 90-91%). Then, each strategy was classified as one out of 27 possible types of strategies (for a list, see Melzner et al., 2020). Interrater-reliability was sufficient (Gwet’s AC1 = .73). Next, each strategy was automatically coded as being either immediate for the selected biggest problem or not, using a theoretical determined mapping of strategies to problems (previous version published in Melzner et al., 2020). In the end, a person was dichotomously classified as reporting an immediate strategy when at least one strategy could be considered as immediately solving their biggest problem. To determine the intensity of strategy use, we added up the number of valid regulation strategies reported at all social levels.

To measure *successful problem regulation*, we adapted three items from Engelschalk et al. (2016) (e.g., “During group learning, we got the biggest problem under control.”). Each item had to be rated on a Likert-scale (from 1 = *not at all true* to 5 = *completely true*). Cronbach's alpha was .96.

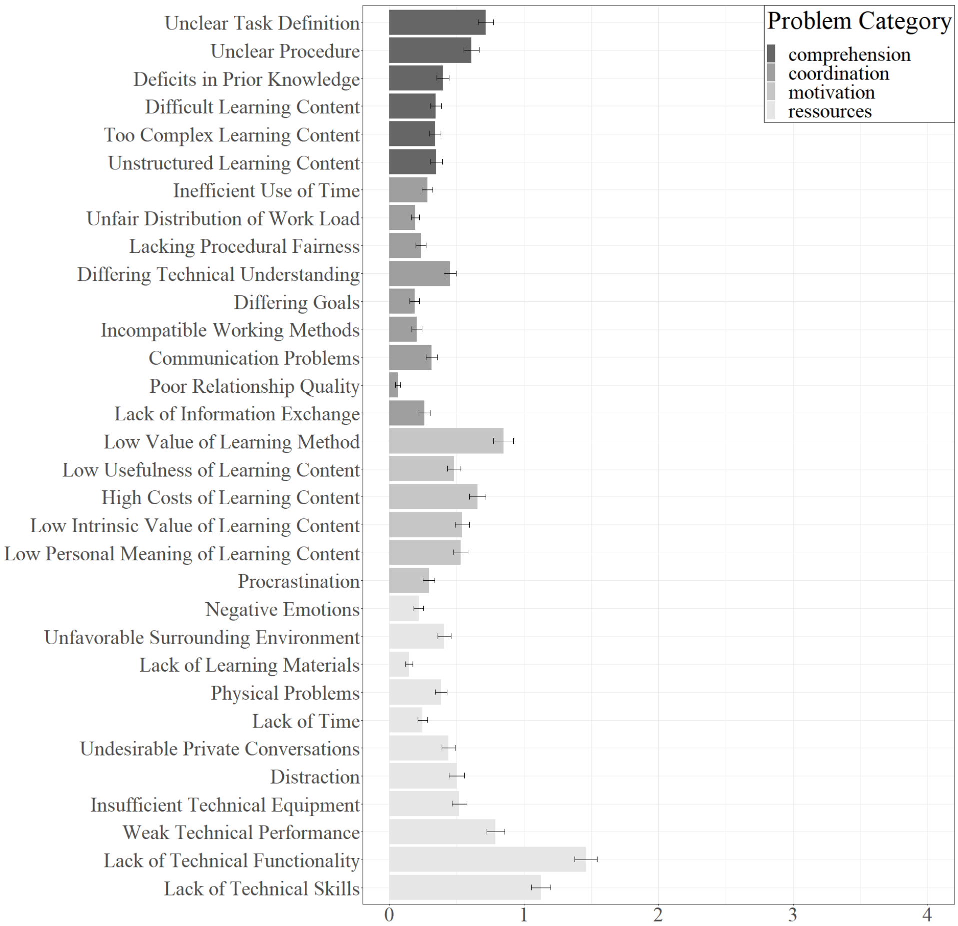
Satisfaction with the group learning experience was measured by five items from the German version of the Satisfaction with Life Scale (SWLS; Glaesmer et al., 2011) adapted to the group learning context (e.g., “Our group work was excellent."). Each item employed a 5-point Likert scale ranging from 1 (*not at all true*) to 5 (*completely true*). Cronbach's alpha was .92.

We assessed *subjective learning success* by using six adapted items from the Training Evaluation Inventory (TEI; Ritzmann et al., 2014). Learning success with regard to the ICAP-Model (Chi & Wylie, 2014) and learning success with regard to the multi-store model of memory (Atkinson & Shiffrin, 1968) were measured separately by three items each (e.g., "I have the impression that my knowledge on the ICAP-Model/the multi-store model of memory has expanded on a long-term basis") on a 5-point Likert-scale (from 1 = *not at all true* to 5 = *completely true*). Cronbach's alpha was .92.

As an *objective measure of learning success*, we mimicked a typical standardized psychology exam: We constructed eight multiple choice questions with four dichotomous answer alternatives each (four questions for each theory). As a total test score, we used the percentage of right answers (= mean).

Results

First, we investigated the descriptive distribution of different problems (see Fig. 2). Overall, the magnitude of problems was low. Even the most pronounced problems seemed to be not severely problematic. The most frequent were technical problems (mostly centered around the recommended mind mapping-software), followed by motivational and comprehension problems regarding the collaboration method, followed by low motivation to study the learning content. Comprehension and coordination problems were very low to almost non-existent.



<u>Figure 2.</u> Size of problems during collaborative learning (means and standard errors).

Second, we inspected descriptive statistics of predictor and criterion variables (see Tab. 1). Twenty-one percent of participants located the biggest problem at the same social level within their groups. Regarding immediacy, 71% of the participants applied at least one immediate regulation strategy to remedy the biggest problem. Regardless of the type, about four strategies were reported on average. Successful problem regulation and satisfaction with the group learning experience were estimated to be rather high, while subjective learning success was appraised a bit lower. Of all test questions measuring objective learning success, 75% were solved correctly on average. Predictor variables were not significantly associated with each other, except for immediacy and intensity. The subjective measures for regulation success were associated with each other, but only content-related homogeneity of problem perception was associated with these outcomes. The objective measure of learning success was not related to any of the other variables.

Table 1. Means, standard deviations, and correlations.

Variable	*M*	*SD*	1	2	3	4	5	6	7
1. Homogeneity problem type	0.00	0.30							
2. Homogeneity social level	0.21	0.41	.06						
3. Immediacy	0.71	0.45	.07	−.06					
4. Intensity	3.99	2.39	−.00	.10	.34**				
5. Successful problem regulation	4.12	1.07	.21**	.11	.09	.08			
6. Satisfaction with group learning	4.12	0.84	.42**	.06	.11	.09	.53**		
7. Subjective learning success	3.76	0.89	.29**	−.01	.04	.04	.33**	.33**	
8. Objective learning success	0.75	0.10	.10	−.08	.06	.11	.02	−.05	.06

Note. ***p* < .01.

Third, we conducted multilevel regression analyses to account for the two-level structure (students in groups) and covariations between predictor variables (see Tab. 2, all variables standardized before analysis). However, the pattern of findings remained the same as with the bivariate correlations reported above. To check if the results would remain stable when covariations between dependent variables were considered as well, we also conducted a structural equation model with all eight predictor and dependent variables in one model and group as a cluster variable, which led to an identical pattern of effects.

Table 2. Multilevel modeling of four different measures of regulation success.

	Satisfaction with learning		**Successful problem regulation**		**Subjective learning success**		**Objective learning success**	
Predictors	β	(*SE*)	β	(*SE*)	β	(*SE*)	β	(*SE*)
(Intercept)	.00	(0.07)	.00	(0.08)	.00	(0.08)	.02	(0.08)
Homogeneity problem type	.42 ***	(0.07)	.19 *	(0.08)	.30 ***	(0.08)	.08	(0.08)
Homogeneity social level	.04	(0.07)	.10	(0.08)	–.03	(0.08)	–.10	(0.08)
Immediacy	.03	(0.07)	.03	(0.08)	–.03	(0.07)	–.01	(0.08)
Intensity	.09	(0.02)	.03	(0.07)	.03	(0.07)	.03	(0.07)
Random Effects								
σ^2	0.70		0.89		0.82		0.84	
τ^00	0.12 (GrNr)		0.08 (GrNr)		0.12 (GrNr)		0.17 (GrNr)	
ICC	0.15		0.08		0.13		0.17	
N	74 (GrNr)		74 (GrNr)		74 (GrNr)		74 (GrNr)	
Observations	193		193		193		193	
Marginal R^2 / Conditional R^2	.187 /.307		.048 /.127		.086 /.204		.014 /.181	

Note. **p* < .05 ****p* < .001

Discussion

This study investigated which problems occurred during one session of (relatively) self-organized online collaborative learning and how groups regulated these problems. Descriptive analyses of problem ratings and means of regulation success variables draw a picture of a rather successful learning experience: All problems were reported as being small or very small, and at the same time, subjective measures of regulation success indicated successful regulation of these problems, high satisfaction and solid subjective learning success. This is good news for university teachers who are forced to move their regular classrooms into the online domain: In general, students seem to be prepared to successfully collaborate in this realm. This finding is in contrast to Capdeferro and Romero (2012), for example, who found students to report frustrations about online collaborative learning more frequently. The main question of this study was how homogeneity of problem perceptions within study groups and immediacy and intensity of regulation strategy use would be associated with different measures of regulation success. In sum, homogeneity of problem perception was the only significant predictor of subjective measures of regulation success. This might mean that groups who have a commonly shared perspective on what their problems are were more successful in regulating their problems. This finding replicates the same finding of Melzner et al. (2020). Contrary to Melzner et al. (2020), we did however not find immediacy and intensity of strategy use to be associated with regulation success. This also contrasts with Engelschalk et al. (2016), who found strategies to be selectively used for different kinds of problems, but is in line with Schoor and Bannert (2012), who also did not find an effect of intensity of regulation strategy use on regulation success. To better interpret this finding, it is informative to take the difference between the two studies into account: Melzner et al. (2020) investigated completely self-organized groups preparing for important exams for an extended period of time, while the present study explored a single session of collaborative learning during a regular lecture. Thus, we compare an extensive, high stakes setting to a less extensive, lower stakes setting. In addition, the level of autonomy and instructional support differed: In Melzner et al. (2020), the learning content, materials, and method were completely self-selected, while in the present study, all this was fixed. In other words, in the present study, the instructional context might have helped to pave the road for collaborative learning enough, so that the specific strategy choice and intensity of its application did not matter for regulation success as much, because just any regulation strategy (applied with random intensity) might have been good enough to overcome a (rather) insignificant problem. We conclude that the full model of problem regulation shown in Fig. 1 might only apply to truly self-organized learning contexts with sufficient prevalence of problems, while problem regulation might follow a simpler process only relying on a shared problem perception when problems are low due to effective instructional support. The fact that the instructional support in the present study seemed to be sufficient is slightly surprising: When taking recommendations for instructional design of instances of collaborative learning (Strauß & Rummel, 2020) into account, only few principles were realized here. The same is true for the technical realization: Only three out of seven affordances for computer supported collaborative learning (Jeong & Hmelo-Silver, 2016) were observed here (video chat as communication means, concept map as representational tool, and facilitation of group formation). And when considering the concrete actions of students themselves, it remains unclear if students applied more than two strategies out of 10 (MacMahon et al., 2020), namely scheduling uninterrupted work and creating a shared concept map. This may mean that a low-level instructional support already makes a big difference and helps to simplify the dynamics of self-organized collaborative learning in a way that students cope successfully with upcoming problems.

When interpreting the results, we have to take the following limitations into account. First, neither the predictor variables nor the subjective measures of regulation success were associated with the results of the objective knowledge test. There are several explanations for this: It might be that the actual knowledge is influenced by many other variables not in the scope of this study which might increase unsystematic error variance making it difficult to find small effects. Alternatively, the lack of a significant association might be due to the low prevalence of problems which might have created a ceiling effect, therefore reducing variance and possible covariation. Second, all measures (except the knowledge test) were based on self-report, though regulation strategies were measured by open-ended questions at least in order to reduce social desirability bias. True associations might be different.

The interpretation of the different findings in the previous study by Melzner et al. (2020) and the present study has important implications for theory building: A new theoretical model of problem regulation during collaborative learning has to be developed that includes problem intensity and variety as moderator of the relations between problems, their regulation, and learning outcome. For teaching practice, the study might imply that recommendations of good instructional design for collaborative learning (see above) also apply to relatively self-organized online collaborative learning and that simple and few scaffolding aids might already help to reach satisfying knowledge gain.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 2, pp. 89–197). Academic Press.

Bellhäuser, H., Müller, A., Konert, J., & Röpke, R. (2019). Birds of a feather learn well together? An experimental study on the effect of homogeneous and heterogeneous learning group composition on satisfaction and performance. In K. Lund, G. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A*

Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) (Vol. 2, pp. 721–722). International Society of the Learning Sciences.

Capdeferro, N., & Romero, M. (2012). Are online learners frustrated with collaborative learning experiences? *International Review of Research in Open and Distributed Learning*, *13*(2), 26–44.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Eckerlein, N., Roth, A., Engelschalk, T., Steuer, G., Schmitz, B., & Dresel, M. (2019). The role of motivational regulation in exam preparation: Results from a standardized diary study. *Frontiers in Psychology*, *10*, 81. https://doi.org/10.3389/fpsyg.2019.00081

Engelschalk, T., Steuer, G., & Dresel, M. (2016). Effectiveness of motivational regulation: Dependence on specific motivational problems. *Learning and Individual Differences*, *52*, 72–78.

Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the Satisfaction with Life Scale (SWLS). Psychometric properties, validity, and population-based norms. *European Journal of Psychological Assessment*, *27*(2), 127–132. https://doi.org/10.1027/1015-5759/a000058

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist*, *48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006

Järvenoja, H., Näykki, P., & Törmänen, T. (2019). Emotional regulation in collaborative learning: When do higher education students activate group level regulation in the face of challenges? *Studies in Higher Education*, *44*(10), 1747–1757. https://doi.org/10.1080/03075079.2019.1665318

Järvenoja, H., Volet, S., & Järvelä, S. (2013). Regulation of emotions in socially challenging learning situations: An instrument to measure the adaptive and social nature of the regulation process. *Educational Psychology*, *33*(1), 31–58. https://doi.org/10.1080/01443410.2012.742334

Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported collaborative learning: How to support collaborative learning? How can technologies help? *Educational Psychologist*, *51*(2), 247–265. https://doi.org/10.1080/00461520.2016.1158654

Koivuniemi, M., Panadero, E., Malmberg, J., & Järvelä, S. (2017). Higher education students' learning challenges and regulatory skills in different learning situations. *Journal for the Study of Education and Development / Infancia y Aprendizaje*, *40*(1), 19–55.

MacMahon, S., Leggett, J., & Carroll, A. (2020). Promoting individual and group regulation through social connection: Strategies for remote learning. *Information and Learning Sciences*, *121*(5/6), 353–363.

Melzner, N., Greisel, M., Dresel, M., & Kollar, I. (2020). Regulating self-organized collaborative learning: The importance of homogeneous problem perception, immediacy and intensity of strategy use. *International Journal of Computer-Supported Collaborative Learning*, *15*(2), 149–177.

Noroozi, O., Bayat, A., & Hatami, J. (2019). Effects of a digital guided peer feedback system on student learning and satisfaction. In K. Lund, G. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL)* (Vol. 2, pp. 809–810). International Society of the Learning Sciences.

Ritzmann, S., Hagemann, V., & Kluge, A. (2014). The Training Evaluation Inventory (TEI)—Evaluation of Training Design and Measurement of Training Outcomes for Predicting Training Success. *Vocations and Learning*, *7*(1), 41–73. https://doi.org/10.1007/s12186-013-9106-4

Schoor, C., & Bannert, M. (2012). Exploring regulatory processes during a computer-supported collaborative learning task using process mining. *Computers in Human Behavior*, *28*(4), 1321–1331.

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. https://doi.org/10.3102/00346543069001021

Strauß, S., & Rummel, N. (2020). Promoting interaction in online distance education: Designing, implementing and supporting collaborative learning. *Information and Learning Sciences*, *121*(5/6), 251–260.

Wecker, C., & Fischer, F. (2014). Lernen in Gruppen [Learning in groups]. In T. Seidel & A. Krapp (Eds.), *Pädagogische Psychologie* (6th ed., pp. 277–296). Beltz.

Weinberger, A., Stegmann, K., & Fischer, F. (2012). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior*, *26*(4), 506–515.

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 299–315). Routledge. https://doi.org/10.4324/9780203876428

Who I Am, What I Know, and What I Want:
An Epistemic Network Analysis of Student Identity Exploration

Amanda Barany, Drexel University, amb595@drexel.edu
Yeyu Wang, ywang2466@wisc.edu, University of Wisconsin-Madison
David Williamson Shaffer dws@education.wisc.edu, University of Wisconsin-Madison
Aroutis Foster, Drexel University, anf37@drexel.edu

Abstract: This paper reports outcomes of 57 students' exploration of urban planning and environmental science identities through Virtual City Planning, a course implemented in a science museum that leveraged a virtual learning environment supported by in-class play-based experiences. Identity exploration trajectories were assessed using the Projective Reflection framework, which consists of constructs that capture cognitive, affective, and behavioral features of the self in addition to learners' self-perceptions and definitions. Researchers constructed a parsimonious epistemic network that was supported by in-depth qualitative interpretations to a) visualize students' general trends of student self-reflection across the course experience and b) highlight which Projective Reflection constructs were highly nascent to participants as they engaged in identity exploration. Results further theoretical understandings of how courses designed to support identity exploration influence the sophistication and content of learners' reflections on the self and illustrate the utility of epistemic networks for visualizing identity exploration trajectories over time.

Introduction

Education research has examined ways to encourage learners to engage in identity exploration, or "the deliberate internal or external action of seeking and processing information in relation to the self" (Kaplan, Sinai & Flum, 2014, p. 250). Identity exploration as a form of situated, intentional, and self-directed learning can encourage identity shifts in targeted directions over time, such as a steps toward a career in science, technology, engineering, or mathematics (STEM) (i.e. Foster, 2014). Interventions that support identity exploration may therefore be of value in the 21st century for fostering adaptive skill development and career preparation in emerging and under-accessed STEM careers (Callahan, Ito, Campbell, Wortman & Wortman, 2019).

Virtual learning environments such as games and simulations have been highlighted as useful tools for promoting shifts in domain or career-specific knowledge (cognitive), motivation (affective), and relevant behaviors (Qian & Clark, 2016). Enactment of such cognitive, affective, and behavioral shifts often centers around identification with specific roles (self-definitions) that players may not have access to in real-world settings (Turkle, 1996). While meta-reviews of game-based learning suggest that game design and implementation are increasingly influenced by education theory (i.e., Clark, Tanner-Smith & Killingsworth, 2016), research is needed in the context of identity to inform how theoretically informed interventions might shape identity exploration outcomes for students across diverse contexts. Emerging research on games for identity exploration will also benefit from the use of methodological approaches that can illustrate the nuances of student identity exploration as it unfolds across a designed game-based learning experience.

To address this gap, this work leverages the Projective Reflection (PR) framework to operationalize learning as identity exploration that can result in identity changes over time, as facilitated by games and game-based learning environments (Foster, 2014). PR was used to design three iterations of Virtual City Planning (VCP), a course that leveraged a virtual learning environment (Philadelphia Land Science) and supportive in-class curriculum to promote exploration of urban planning and environmental science career identities. VCP was implemented in a museum classroom context with a diverse sample of high school students (N=57). Identity exploration is conceptualized using PR as shifts in reflection on 12 constructs that relate to cognitive, affective, behavioral, and self-definitional aspects of the self. Student trajectories of identity exploration were visualized using Epistemic Network Analysis, a quantitative ethnographic technique for modeling connections among key concepts to represent underlying longitudinal phenomena. The Parsimonious Removal with Interpretive Alignment approach (Wang, Swiecki, Ruis & Shaffer, 2021) was then used to optimize the twelve-construct model as an eight-construct network that maintains interpretive power. Epistemic networks were supplemented by qualitative case findings from the student cohorts. Results (a) illustrate how VCP supported statistically significant shifts in student conceptualizations of self over time as defined by PR, and (b) illustrated which identity concepts were more or less nascent and discussed in students' written and spoken reflections on the self. The work concludes with implications for games and education practitioners, designers, and researchers.

The research question asks: *How did learners characterize their processes of identity exploration (cognitive, affective, behavioral, self-definitional) through participation in Virtual City Planning?*

Review of literature

Virtual learning environments and identity

Identity exploration is conceptualized by Kaplan and Garner (2017) as not only the self-perceptions and self-definitions a participant iteratively applies during a learning experience, but also the beliefs, values, goals, emotions, and actions that are central to a specified role as it emerges. This process is role-specific in the sense that a learner exploring a career in urban planning, for example, will immerse herself in a different semiotic and social system than that of an art historian. Participation in virtual learning environments can support more explicit awareness of perceptions and definitions of self due to the capacity of such spaces for offering authentic simulations of professional praxis (Shaffer, 2006). Games implemented in learning contexts have also been lauded as valuable for the expression of nested identities (i.e., student and player) as design constraints of the game space intersect with real-world roles and contexts (Gaydos & Devane, 2019).

Reviews of the growing body of research on games for learning have affirmed the potential of virtual learning environments for supporting a variety of cognitive, affective, and behavioral processes that contribute to identity exploration. Most prominently featured in games research is work that points to the efficacy of virtual media for supporting knowledge acquisition and content understanding in contexts such as primary education (Hainey, Connolly, Boyle, Wilson & Razak, 2016), informal learning settings such as museums (Koutromanosa & Avraamidou, 2014), and with computer-based and serious games and simulations (Boyle et al., 2016). Game-based learning has also been lauded as valuable given emerging theoretical conceptualizations of learning, which shifted from passive knowledge acquisition to more *collaborative* and *interest*-driven negotiation of domain-specific content (Orr & McGuinness, 2018). Virtual learning environments excel in their capacity to support student engagement and motivation around specific content (Wouters, Van Nimwegen, Van Oostendorp & Van Der Spek, 2013), and can also serve as spaces in which players repeatedly practice goal setting, self-monitoring, and self-regulation behaviors (Gabbiadini & Greitemeyer, 2018). Finally, the communities of practice (Lave & Wenger, 1991) that develop in and around games (What Gee (2003) defines as affinity spaces) offer opportunities for socially mediated regulation of learner goals and activities as players collectively negotiate aspects of their identities and learn from the expertise of others. This aligns with research on coregulation (McCaslin, 2009) and socially shared regulatory processes (Hadwin & Oshige, 2011) as a part of identity work. Virtual tools, particularly those that promote active discussion and collaboration, show promise for promoting the externalization of learning processes and reflection on individual progress (Zheng, Li & Huang, 2017).

Assessing learner identity exploration

While seminal identity research has characterized identity as a developmental process that emerges over time (Erikson, 1959) as mediated by external sociocultural features (Vygotsky, 1978), contemporary researchers have further characterized such role exploration as a complex and dynamic system (Kaplan et al., 2014). This complexity represents a methodological challenge for educational researchers and practitioners looking to examine learners' identity exploration processes as they manifest over time in play-based experiences. Fortunately, reviews of game-based learning literature have highlighted the emergence of increasingly sophisticated methods for understanding learner processes (de Freitas, 2018), such as data modelling (e.g., Westera, 2017) and individual analytics (e.g., Drachen, El-Nasr & Canossa, 2013). de Freitas also argues for the use of combined quantitative and qualitative measures in forthcoming game-based learning research.

Quantitative Ethnography (QE; Shaffer, 2017) offers a method for exploring learning as a form of complex thinking by offering analytic techniques that can visualize constructs (such as facets of identity) as dynamic network models. Epistemic Network Analysis (ENA) is a QE technique in which qualitative data is quantified so that patterns of association may be visualized between a learners' developing "knowledge, skills, values, habits of mind, and other elements" (Shaffer, Collier & Ruis, 2016, p.10). ENA is validated by examining alignment between qualitative constructs and quantitative representations, defined as interpretive alignment. In addition, model parsimony, as another key concept in QE research, concerns about capturing the "right" amount of detail to explain the phenomenon from both qualitative analysis and quantitative representation. Existing studies of identity exploration that leverage ENA have only applied a priori model simplification (i.e., Barany & Foster, 2020). Generating parsimonious models of identity exploration, based on prior research, is an important next step in assessments of student outcomes that was implemented in this work.

Theoretical framework

This study leveraged Projective Reflection (Foster, 2014) as a research-informed theoretical framework to structure the design of the course (Virtual City Planning) and the virtual learning environment (Philadelphia Land Science). The model was also used as an analytical tool for conceptualizing how learners engage in identity exploration in play-based and virtual learning environments. Identity exploration is captured through individuals' reflections on the self in one moment, which is meaningfully connected to how they conceptualized themselves across prior moments longitudinally. This way, identity change can be assessed over time as participants project forward and reflect back on (a) their current knowledge of a topic, (b) what aspects of the topic they care about, (c) how they think and the processes they use to make choices and take actions, (d) what they want and expect to be in the future, and (e) how they see themselves in the present (Foster, Shah, Barany & Talafian, 2019). PR leverages twelve theoretical constructs to conceptualize identity in game-based learning contexts (see Table 1) under four features of identity exploration: (1) knowledge (i.e., Kereluik, Mishra, Fahnoe & Terry, 2013), (2) interest and valuing (i.e., Eccles, 2009; Hidi & Renninger, 2006), (3) self-organization and self-control (i.e., Hadwin & Oshige, 2011), and (4) self-perceptions and self-definitions (i.e., Kaplan et al., 2014). Constructs were developed and refined through in-depth review of literature on identity, learning, motivation, and individual and socially mediated change to capture the role-specific cognitive, affective, behavioral, and self-definitional features of self that shift over time through participation in identity exploration processes.

Table 1: Projective Reflection theoretical constructs

	Constructs	Construct manifestations
1. Knowledge (awareness of *cognitive* capabilities)	1.1 Foundational knowledge	• Describing knowledge of a domain-specific topic • Defining domain-specific terms or concepts
	1.2 Meta knowledge	• Describing awareness of how to use or apply foundational knowledge in context. • Enacting domain-specific processes/applying concepts
	1.3 Humanistic knowledge	• Knowledge of the self and its location in a broader social, global, and professional context
2. Interest and valuing (awareness of *affect*)	2.1 Interest	• A predisposition to re-engage with a domain or topic over time, psychological state of engagement. • Describing a domain or concept as interesting
	2.2 Subjective task valuing	• Values attached to a domain, topic or concept that motivate the choice to engage. • Describing a domain, topic, or concept as valuable
	2.3 Relevance	• Awareness of a domain, topic or concept's importance for the self, a learner's community, or society broadly
3. Self-organization and self-control (awareness of *behaviors*)	3.1 Self-regulation	• Describing one's strategic and metacognitive behaviors aimed at achieving a goal (i.e., goal setting, self-monitoring, outcome assessment)
	3.2 Coregulation	• Regulatory behaviors that are supported by a more knowledgeable peer or mentor
	3.3 Socially shared regulation	• Regulatory behaviors that are negotiated and enacted collectively by a group
4. Self-perceptions and self-definitions	4.1 Self-efficacy	• Confidence in one's ability to achieve goals/results. • Engaging in self-monitoring and self-evaluation
	4.2 Current self-concept	• Descriptions or labels applied to the self in the present. • Careers/roles a learner is enacting currently
	4.3 Possible selves explored	• Future role/career a learner wants or expects to have. • Roles a learner has tried, but may not wish to pursue

Methods

Study context

This research was conducted as part of a CAREER project awarded to support the study, design, and implementation of virtual learning environments and curricula for promoting Projective Reflection (Foster, 2014). To enact this process in a meaningful real-world learning context, the primary investigator and his team of researchers partnered with a local science museum in Philadelphia. The museum offers weekly science-related learning opportunities to middle school students from a local school. The research team also partnered with the Epistemic Analytics Group at the University of Wisconsin-Madison to redesign the existing virtual internship Land Science (Barany et al., 2017) to support identity exploration and match the needs of the science museum context. Philadelphia Land Science built on the strengths of Land Science as an immersive environment but was informed by Projective Reflection to position learners collaborating in-person during VCP as interns at a fictitious urban planning firm. The virtual learning environment and in-class mentors roleplaying as urban planners guided participants through the process of creating zoning plans for downtown Philadelphia, an area with which students were familiar. Students worked in groups of five to (1) learn about the process of creating a city rezoning proposal, (2) research specific environmental and economic needs of city stakeholders, and (3) rezone a virtual map of downtown Philadelphia to enact desired changes (e.g., decrease air pollution). Students concluded by writing a final rezoning plan outlining their city's needs and the nature of their zoning changes.

The play-based course titled Virtual City Planning was developed and implemented across three consecutive courses held at the museum between the academic year 2016 and 2017 with 57 racially diverse middle school participants. Virtual City Planning involved weekly use of the virtual learning environment supported by in-class opportunities for role-play, self-reflection, and discussion with peers. Examples of in-class activities included supplementary materials (e.g., a documentary video), group discussions on activities and processes of identity exploration, and analogous paper activities designed to support students with less technical literacy (i.e., rezoning the city by drawing on paper maps). Design of each weekly session included virtual and in-class opportunities for individual reflection and collaborative discussion on each facet of students' identity exploration processes (the 12 constructs), in addition to periods uninterrupted play and group engagement in activities. For example, in one class, students rezoned areas of Philadelphia in small groups, negotiated with other design groups to create a map that met everyone's needs, then collaboratively discussed what it felt like to act as an urban planner.

Data collection

Qualitative and quantitative data was obtained through in-game (e.g., written reflections as urban planning interns) and classroom artifacts (e.g., survey responses). Text data was organized chronologically for each student to track changes in identity exploration processes from beginning to end of VCP. After each class, researchers collaborated to write detailed memos of interactions with students; memos were segmented by discussion of student and organized chronologically in each student's data file. Player data was collected from the following sources:

- A pre and post survey consisting of (a) 5-point Likert-style questions (ranging from Strongly Agree to Strongly Disagree on questions such as "I can see myself in an urban planning career in the future"), and (b) short answer questions (e.g., "describe your interests in learning about cities and the environment").
- Responses to writing prompts in Philadelphia Land Science, framed as emails to the design firm.
- Written posts made on an online forum website as a curricular activity.
- Digitized copies of handwritten reflections from paper handouts and notebook annotations, etc.
- Written researcher memos on student interactions, discussions, and activities.
- Screenshots and images of student map designs, from the virtual internship tool and from in-class design activities using paper maps. Images were examined for qualitative analyses but not ENA.

Data analysis

Once data collection and organization were completed, researchers then engaged in a deductive or directed coding process for each case (Krippendorff, 2004) in which each line of data was coded for self-reflection on/demonstration of one or more aspect of identity exploration, with agreement reached by two coders. Lines were coded for the occurrence (1) or non-occurrence (0) of the Projective Reflection constructs to prepare for visualization of identity exploration patterns using Epistemic Network Analysis (ENA). For example, a student's reflection reading, "the big ones [issue] I care about is pollution," was coded (1) for the construct '2.1 Interest.'

We applied ENA (Shaffer, 2017) to our data using ENA1.5.2 Web Tool. ENA assumes that a single piece of student data (written, observed) may be representative of individual change in one or more codes (the PR identity constructs), but also that the data has local structure and that an important feature of the data is the way codes are connected. Based on this assumption, ENA generates network visualizations of the co-occurrence of codes within a moving stanza window, which means that all codes applied to a single line of student data are connected to each other and to codes applied to the previous 3 lines of chronological student data (as recommended by Siebert-Evenstone et al., 2017). This process is appropriate given the conceptualization of identity exploration as a developmental process of change. Epistemic networks for code relationships were generated for the first half (Time 1) and second half (Time 2) of class for sessions 1-3 to explore how student identity exploration shifted over time as supported by each iteration. ENA also analyzes all chronological networks simultaneously so that they can be compared visually and statistically. To achieve this, ENA models normalize the networks for all units of analysis before they are subjected to a dimensional reduction, which accounts for the fact that different units of analysis may have different amounts of coded lines in the data (see Shaffer et al., 2016). Epistemic networks were generated for Time 1 and Time 2 for each session to compare within and across them over time. In addition, two sample t-tests were completed to test whether changes from Time 1-2 in each session were statistically significant along the top two dimensions explaining the most variances. The results also reference themes identified from qualitative studies of the data (i.e., Foster et al., 2019) to close the interpretive loop and provide deeper understanding of the modeled phenomena. After the 12-construct epistemic network was developed, we applied Parsimonious Removal with Interpretive Alignment (PRIA), which reduced the network of students' identity exploration to an eight-construct model without losing interpretive alignment (Wang et al., 2021). PRIA takes an existing ENA model and finds a model with the fewest codes that retains high goodness of fit, correlations of ENA scores and correlations of node positions between the simpler model and the original.

Results and discussion

To answer the question "How did learners characterize their processes of identity exploration (cognitive, affective, behavioral, self-definitional) through participation in Virtual City Planning?" The parsimonious epistemic network model of eight Projective Reflection constructs is presented in a difference model (see Figure 1). Two cognitive constructs (foundational knowledge and meta knowledge), three affective constructs (interest, subjective task valuing and relevance), no behavioral constructs, and three self-definitional constructs (self-efficacy, current self-concept, and possible selves explored) were identified as highly nascent to students' reflections on the self as they engaged in any of the three sessions of VCP. In the difference model, red lines are associations between constructs that were more prevalent in the first half of each course (Time 1), and blue lines represent associations that were more prevalent in the second half (Time 2). Nonparametric Mann-Whitney test showed that Time 1 associations were statistically significant from Time 2 associations along the X-axis at the alpha=0.05 level (Mdn=0.21, N=214 U=10863.50, p=0.00, r=0.48). This suggests that students' processes of identity exploration were enacted in meaningfully different ways over time, to be discussed further below.

Though associations between student reflections on *foundational* and *meta knowledge* were strongest compared to other associations for Time 1 and Time 2, students were slightly more likely to make connections between these two cognitive features in Time 1. Qualitative examinations of the data revealed they were often able to describe or discuss perceived knowledge of urban planning and its relevant topics throughout the course. Discussions of knowledge typically began with more binary judgments on perquisite knowledge and expertise early in the course; for example, Zola (pseudonym) wrote "I have experience with urban planning through my mom" while Megan wrote "I can definitely tell you I don't know much about urban planning." Towards the end of the course, discussions of knowledge trended toward more sophisticated use of situated definitions, terms and processes as students gained experience with the urban planning role; for example, Jake initially did not know what a stakeholder did, but was later able to explain that "A stakeholder is a rich person who has interest in what urban planners do. Stakeholders can make change in our neighborhoods that positively or negatively."

All associations between PR constructs and students' current *self-concepts* were stronger in Time 1 than in Time 2. This result is at first counterintuitive, but an in-depth examination of qualitative reflections reveals that students were initially more likely to affirm concrete descriptions of self, but that these self-definitions were often distanced from environmental science and urban planning. For example, Kevin wanted to be a "professional dancer because I'm a great dancer" and then shared that he had never considered urban planning as a future career because he didn't "really know how good urban planning is." As the VCP course progressed, however, student reflections on the self became less concrete, but also more connected to urban planning roles they had explored. For example, Ellen wrote "Urban planning could contribute to the job I hoped to have because I learned a lot about the importance of high density housing and compacting space for where people live." These more nuanced reflections on the self meant that learners at the end of the course were more often unsure whether or not they might consider urban planning as a future role, where before they may have dismissed the role outright.

While the location of the eight PR constructs on the three-dimensional plane are positioned to allow for model comparison, positionality of the overall means for Time 1 and Time 2 (the red and blue squares) can be interpreted in relation to the constructs and in relation to each other. For example, the overall mean of associations for Time 2 in the model is skewed to the right compared to Time 1, toward the affective and motivational factors of identity. Prior research on this data found trends toward interests and valuing more broadly (Barany, Talafian & Foster, 2020), but this parsimonious eight-code model reveals the specific processes driving this shift: (1) increased learner associations between *foundational knowledge* and *relevance*, and (2) *self-efficacy* and *interest*. These trends play out qualitatively, as learners were more likely to explain why urban planning and environmental science were relevant for themselves or their communities over time. Consider Emil, for example, who connected new knowledge of urban planning concepts to his awareness of climate change issues: "I'm very scared for the health of not only our city, but our planet. We destroy natural ecosystems to create businesses and heat up the Earth just to run our cars. I'm hoping by adding more green open spaces, we will create a better Philadelphia." As students gained confidence in their abilities as urban planners, they were more likely to affirm interest in the topic as well. Ali reflected that he enjoyed taking on the role of an urban planner: "[it felt] good because its my responsibility to actually take part of helping my community out by planning things or seeing what things look like in the modern world". He then described confidence (self-efficacy) in his ability to enact urban planning changes in his community: "i can see myself being a construction worker, on the urban planning things that i know that i can change, it would be easy for myself to create the open space for the people in my neighborhood."

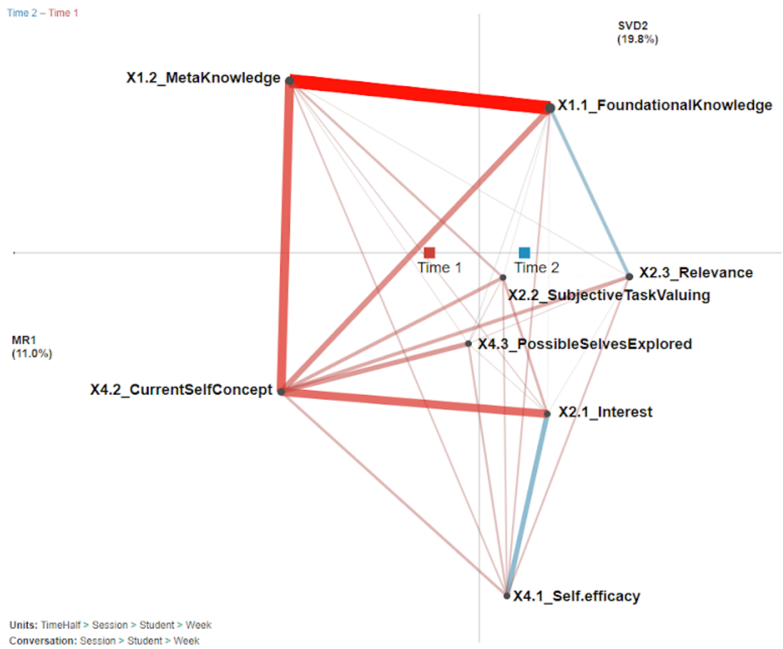


Figure 1. A parsimonious difference model of student identity exploration in which strength of construct associations in Time 2 (blue) were subtracted from the strength of construct associations in Time 1 (red).

In addition to the removal of humanistic knowledge as a construct in the networks, the parsimonious model with the best goodness of fit advocated for the removal of all constructs related to behavioral features of identity (self-regulation, coregulation, and socially shared regulation). While regulated activities remain a highly important and relevant feature of learners' identity exploration processes, students in VCP were more likely to connect their emerging perceptions and definitions of self (who I am), to their developing interest and perceived relevance of the topic (what I want), to their increasingly specific knowledge of the topic (what I know). While the behavioral 'what I do' piece is important from a theoretical perspective, students were less likely to meaningfully integrate discussions of their behaviors into reflections on their identity exploration processes. This could result from students' newness to identity exploration processes, a lack of self-awareness of their own behaviors or a lack of intentionality when enacting them or could be a result of design choices in curricular design that limited student reflection on behavioral features. Further research is needed to understand why behavioral features emerged as less nascent to students' reflections on the self in VCP.

Conclusions and implications

Results illustrate the potential of educational experiences designed to facilitate Projective Reflection (Author, 2014) as a way to develop learners' skill in enacting situated, targeted and intentional identity exploration related to STEM domains (i.e., environmental science and urban planning). Though characteristics of the designed experience (VCP) and features of the student cohort may have influenced how students reflected on their identity exploration processes, trajectories of identity exploration over time shifted from an emphasis on more concrete and simplistic discussions of initial knowledge and self-concepts to a more nuanced conceptualization of self that was grounded in emerging interests and perceived relevance of the topic. These findings align with summary reports on the acquisition of STEM careers, which suggests that identity exploration may be closely linked to students' developing interest and motivation around a topic, resulting in closer and deeper engagement with the topic over time (CAISE, 2018). Given these findings, designed virtual learning environments such as Virtual City Planning have potential to serve as particularly valuable avenues for promoting the exploration (and potential future acquisition) of STEM identities. Further work is needed to address limitations of this study design, such as (1) examinations of student change across longer time periods and with a more diverse group of students, (2) assessments of the influence of specific curricular design features (see Author, in press for preliminary work on this topic), and (3) the application of the PRIA model to identity exploration processes across more contexts.

Quantitative Ethnographic (QE) (Shaffer, 2017) techniques such as Epistemic Network Analysis (ENA) served as a valuable and innovative approach for understanding whole-group trajectories of identity exploration as operationalized by Projective Reflection. The parsimonious epistemic network not only allowed researchers to examine large quantities of student data related to identity exploration by providing a nuanced view of the relationships between the most nascent identity constructs, but also supported comparison of group characteristics over time (Time 1 to Time 2). Future studies will test and refine new virtual learning environments that can facilitate Projective Reflection in different contexts, and also incorporate methods such as Social-Epistemic Network Analysis (See Gašević, Joksimović, Eagan & Shaffer, 2019) to examine identity exploration as both an individual/developmental and collective/situational process of change over time.

References

Barany, A., & Foster, A. (2020). Context, Community, and the Individual: Modeling Identity in a Game Affinity Space. *The Journal of Experimental Education*, 1-18.

Barany, A., Shah, M., Cellitti, J., Duka, M., Swiecki, Z., Evenstone, A., Kinley, H., Quigley, P., Shaffer, D. W., & Foster, A. (2017, October). *Designing Philadelphia Land Science as a game to promote identity exploration.* In D. Sampson, J. M. Spector, D. Ifenthaler & P. Isaias (Eds.) proceedings of the 14th international conference on the Cognition and Exploratory Learning in Digital Age (CELDA). Vilamoura, Algarve, Portugal, October 18-20, 2017.

Barany, A., Talafian, H. & Foster, A. (2020). *An epistemic network analysis of STEM identity exploration in virtual learning environments.* In M. Gresalfi & I. S. Horn (Eds.) Proceedings of the 14th International Conference on Learning Sciences (ICLS) 2020, Vol 1. (pp. 191-198). Nashville, TN: International Society of the Learning Sciences.

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., ... & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education, 94*, 178-192.

Callahan, J., Ito, M., Campbell R., Wortman, S. and Wortman, A. (2019). *Influences on occupational identity in adolescence: A review of research and programs*, Connected Learning Alliance, Irvine, CA.

Center for Advancement of Formal Science Education (CAISE). (2018). *Identity in science and STEM: Reflections on interviews with the field.* Washington, DC.

Clark, D. B., Tanner-smith, E. E., & Killingsworth, S. (2016). *Digital Games for Learning: A Systematic Review and Meta-Analysis Preliminary Meta-Analysis Results*.

de Freitas, S. (2018). Are games effective learning tools? A review of educational games. Journal of Educational *Technology & Society, 21*(2), 74-84.

Drachen, A., El-Nasr, M. S., & Canossa, A. (2013). Game analytics–the basics. In A. Drachen, A. Canossa, and M. S. El-Nasr (Eds.) *Game analytics* (pp. 13-40). Springer, London.

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78-89.

Erikson, E. H. (1959). *Identity and the life cycle: Selected papers.* New York, NY: International Universities Press.

Foster, A (2014). CAREER: Projective reflection: Learning as identity exploration within games for science. Drexel University: National Science Foundation.

Foster, A., Shah, M., Barany, A., & Talafian, H. (2019). High school students' role-playing for identity exploration: findings from virtual city planning. *Information and Learning Sciences*, *120*(9/10), 640–662. doi: 10.1108/ils-03-2019-0026.

Gabbiadini, A., & Greitemeyer, T. (2017). Uncovering the association between strategy video games and self-regulation: A correlational study. *Personality and Individual Differences*, 104, 129-136.

 © ISLS

Gašević, D., Joksimović, S., Eagan, B. R., & Shaffer, D. W. (2019). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior*, *92*, 562-577.
Gaydos, M. J., & Devane, B. M. (2019). Designing for identity in game-based learning. *Mind, Culture, and Activity, 26*(1), 61-74.
Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE), 1*(1), 20-20.
Hadwin, A. and Oshige, M., 2011. Self-regulation, coregulation, and socially shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record, 113*(2), pp.240-264.
Hainey, T., Connolly, T. M., Boyle, E. A., Wilson, A., & Razak, A. (2016). A systematic literature review of games-based learning empirical evidence in primary education. *Computers & Education*, 102, 202-223.
Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational psychologist, 41*(2), 111-127.
Kaplan, A., & Garner, J. K. (2017). A complex dynamic systems perspective on identity and its development: The dynamic systems model of role identity. *Developmental Psychology, 53*(11), 2036.
Kaplan, A., Sinai, M. and Flum, H., 2014. Design-based interventions for promoting students' identity exploration within the school curriculum. In *Motivational interventions* (pp. 243-291). Emerald Group Publishing.
Kereluik, K., Mishra, P., Fahnoe, C. and Terry, L., 2013. What knowledge is of most worth: Teacher knowledge for 21st century learning. *Journal of Digital Learning in Teacher Education, 29*(4), pp.127-140.
Koutromanosa, G., & Avraamidou, L. (2014). The use of mobile games in formal and informal learning environments: A review of the literature. *Educational Media International, 51*(1), 49–65.
Krippendorff, K. (2004). Content analysis: An introduction to its methodology. Sage Publications.
Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge, UK: Cambridge University Press.
Orr, K., & McGuinness, C. (2018). What is the "Learning" in Games-Based Learning? In Management Association (Ed.), *Gamification in Education: Breakthroughs in Research and Practice* (pp. 611-634). IGI Global. http://doi:10.4018/978-1-5225-5198-0.ch031
Qian, M., & Clark, K. R. (2016). Game-based Learning and 21st century skills: A review of recent research. *Computers in Human Behavior*, 63, 50-58.
Shaffer, D. W. (2006). *How computer games help children learn.* New York, NY: Macmillan Publishers.
Shaffer, D.W. (2017). *Quantitative ethnography*. Madison, WI: Cathcart Press.
Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, *3*(3), 9-45.
Siebert-Evenstone, A.L., Irgens, G.A., Collier, W., Swiecki, Z., Ruis, A.R. & Shaffer, D.W. (2017), In search of conversational grain size: Modeling semantic structure using moving stanza windows. *Journal of Learning Analytics, 4*(3),123-139.
Turkle, S. (1996). Working on Identity in Virtual Space. Constructing the self in a mediated world, 156.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Wang, Y., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2021, February). Simplification of Epistemic Networks Using Parsimonious Removal with Interpretive Alignment. In *International Conference on Quantitative Ethnography* (pp. 137-151). Springer, Cham.
Westera, W. (2017). How people learn while playing serious games: A computational modelling approach. *Journal of Computational Science*, 18, 32-45.
Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A Meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*(2), 249-265.
Zheng, L., Li, X., & Huang, R. (2017). The effect of socially shared regulation approach on learning performance in computer-supported collaborative learning. *Journal of Educational Technology & Society, 20*(4), 35-46.

Acknowledgments

This material is based upon work supported by the National Science Foundation (DRL-1350707). We extend thanks to the Epistemic Analytics Group at the University of Wisconsin-Madison for access to Land Science (DRL-1661036, DRL-1713110). The opinions, findings, and conclusions expressed are those of the authors, and do not reflect the views of funding agencies, cooperating institutions, or other individuals.

Using Idea Thread Mapper to Support Cross-Classroom "Super Talk" among Four Grade 5 Knowledge Building Communities

Guangji Yuan, Jianwei Zhang, Mei-Hwa Chen
yuanguangji@gmail.com, jzhang1@albany.edu, mhc@cs.albany.edu
University at Albany, State University of New York, USA

Abstract: Enabling collaborative interaction across social levels over longer timescales represents a key research challenge in computer-supported collaborative learning (CSCL). This study investigates a multi-layer interaction approach to cross-community knowledge building supported by the design of Idea Thread Mapper (ITM). A design-based research study was conducted in four Grade 5 classrooms that studied human body systems over six months. ITM was used to support student knowledge building discourse in each classroom and cross-classroom interaction in a shared meta-space focused on a cross-cutting challenging problem: How do people grow? Multi-level discourse analysis traced students' collective idea development in the cross-classroom discourse that built on the diverse lines of inquiry about the different body systems within each classroom. The findings contribute new understanding and designs for expanding CSCL practices across networks of classrooms, enabling a larger creative context for students' ever-deepening and expansive work with ideas.

Introduction

As our societies enter a new era facing extraordinary challenges, rapid changes and hyper-connectedness, researchers call for critical efforts to make computer-supported collaborative learning (CSCL) relevant and contributive, addressing potential tensions and blind spots and enabling educational transformation (Cress, Oshima, Rosé, & Wise, in press). While existing research has produced deep insights into collaborative learning interaction in small groups and individual classrooms, new research is needed to expand the collaborative interaction to higher social levels and over longer timescales to enable transformative classroom change (Stahl, 2013; Wise & Schwarz, 2017). Building on our prior work (Yuan & Zhang, 2019; Zhang, Yuan, & Bogouslavsky, 2020), the current study analyzes a design for enabling student collaboration across classrooms, which work as interconnected communities to build knowledge and address complex problems. The higher-level interaction enabled by new technology design allows students to connect with and build on an expanded pool of ideas across the boundaries of different classrooms. Valuable ideas developed in each classroom community have the opportunity to travel up to a cross-community meta-space for high-level discourse.

Realworld knowledge creation takes place in a multi-level social system, in which individuals and teams create knowledge in various domain areas while working with peers from the larger field (Csikszentmihalyi, 1999). The social dialogues and interactions extend across different social levels: individuals collaborate in groups/teams within each organization/community, which is further part of an intellectual field that advances the collective knowledge of a domain (Csikszentmihalyi, 1999; Sawyer, 2007). The larger discourse in a field creates a macro and dynamic context that shapes and sustains the knowledge work in each local community over time and across generations.

CSCL researchers need to tackle the challenge of how to extend collaborative knowledge building to the higher social levels (Stahl, 2013) across classroom communities. The prominence of this challenge is heightened in contemporary times when we face increasingly complex and connected problems. In this context, it is more critical than ever for students to learn to listen, converse, and collaborate across boundaries to solve complex problems and build shared understanding. Aligned with this need, researchers have made initial explorations of cross-classroom collaboration (Laferriere, Law, & Montané, 2012; Lai & Law, 2006). Through the direct sharing of online discussion spaces between different classrooms, students read the online posts of their partner classrooms and respond. As a challenge arising in this context, students often find it difficult to understand other classrooms' distributed discourse and engage in meaningful dialogues across different communities.

In light of the above multi-level social system view of knowledge creation, we have been testing technology-enabled support to sustain collaborative knowledge building across a network of science classrooms (Yuan & Zhang, 2019; Zhang, Yuan & Bogouslavsky, 2020). Our design uses a multi-layer emergent interaction approach, which integrates the local knowledge space of each knowledge building community and a meta-space shared across communities. While members of each classroom work in their community's local discourse space to investigate various problems and deepen their understanding, they selectively contribute their major knowledge progress and challenges to the meta-space for cross-classroom sharing and discourse. We conceptualize such multi-layer emergent interaction in light of the related theories, including social emergence (Sawyer, 2015), expansive learning that integrates horizontal moves across borders and vertical moves across levels (Engeström, 2014), and expansive framing of unfolding learning trajectories across contexts (Engle et al., 2012). A critical design challenge is to facilitate “the micro-macro link” across levels, which is essential to the function of emergent complex systems (Sawyer, 2015). The micro-macro link involves the bottom-up emergence of ideas from each group and community to the larger discourse space and the downward influence of the cross-community discourse on the future unfolding of inquiry and discourse in each community. Valuable ideas and problems developed in each community can travel up to the cross-community space for extended sharing and higher-level discourse. At the same time, knowledge advances and practices developed in the cross-community space are brought back to each individual community to stimulate further inquiry and discourse and develop integrated understanding in light of the knowledge and perspectives from the different communities. This process may leverage expansive cycles (Engeström, 2014) of inquiry through the dynamic contact and re-orchestration of different viewpoints, expertise, and inquiry practices of the various participants.

To support cross-community knowledge building, this study used a multi-layer collaboration system—Idea Thread Mapper (ITM, http://idea-thread.net), which integrates support for student-driven knowledge building in each community and boundary-crossing interaction across different communities and school years (Zhang & Chen, 2019). ITM inter-operates with Knowledge Forum (Scardamalia & Bereiter, 2006). The ITM supports for knowledge building in each classroom and encourages emergent “reflective structuration” (Tao & Zhang, 2018) by which students co-organize unfolding lines (threads) of inquiry as their collective work proceeds. A multi-layer framework is further used to organize collaborative discourse across different levels. The discourse spaces include the local collaborative space of each classroom where students conduct collaborative discourse and inquiry to advance their understanding of various problems; and a cross-classroom meta-space where students view the inquiry directions of their partner classrooms, post/share Super Notes (syntheses), and engage in cross-classroom Super Talk focusing on challenging issues of common interests. While collaborating on inquiry within their home classroom, students have ongoing access to the cross-classroom meta-space, where they can interact with peers and ideas from their buddy classrooms (including those from the prior school years). Students can see the visual inquiry organizers of the buddy classrooms that show their “Wondering Areas” (inquiry questions) and idea threads, read their Super Notes co-authored using the Journey of Thinking tool to synthesize major progress of inquiry, and, if interested, access their original online discourse (in a read-only mode). A set of analytics is integrated to feedback on emerging idea connection and progress. Students can also propose challenging issues as potential topics for cross-classroom joint discussion, which is called “Super Talk.” The Super Talk topic, once approved by their teacher, becomes a shared idea thread for cross-community discourse. Figure 1 shows an example topic about how people grow shared by a set of Grade 5 classrooms studying human body systems. There is a function for flexible note importing, so students can import notes (ideas) from their local discourse threads to the Super Talk for the larger discourse, and vice versa. While Knowledge Forum already has a Rise-Above tool for writing synthesis notes, the ITM features for super note sharing and Super Talk further turn reflective rise-above into a meta-space for cross-community discourse, which reorchestrates the different insights, problems, and expertise developed in each community to work on complex challenges and ideas.

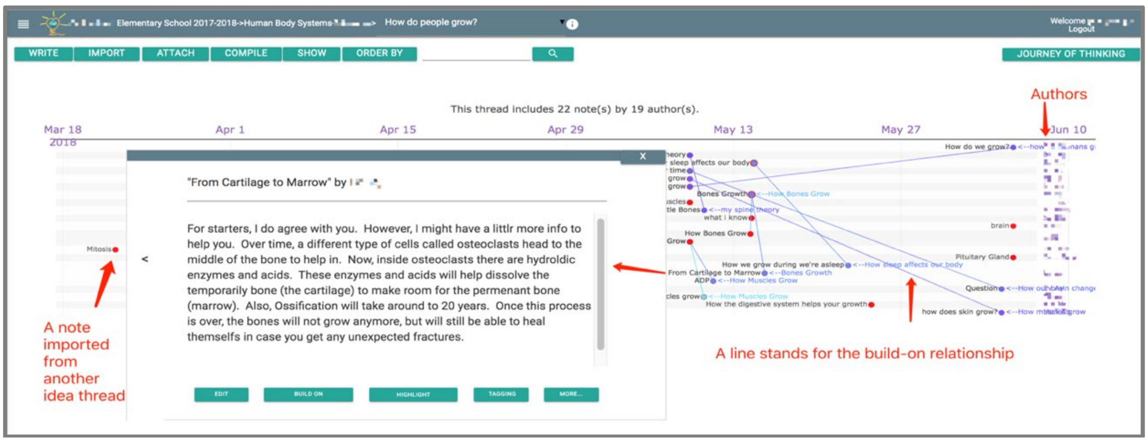


Figure 1. Super Talk about “How do people grow?” among 19 students from four classrooms.

To test and elaborate on the multi-layer emergent interaction design of cross-community knowledge building, we conducted multi-year design-based research in a network of upper elementary science classrooms. A set of specific studies was embedded in this project, addressing unique design challenges and research questions. The first two iterations (school years) in the design-based research tested cross-classroom collaboration support using Knowledge Forum, beginning with two Grade 5/6 classrooms in the first iteration and expanding to a set of four parallel classrooms in the second iteration (Yuan & Zhang, 2019; Zhang et al., 2020). Motivated by the goal of producing knowledge advances for cross-community sharing, students engaged in intentional and collaborative efforts to improve their understanding toward higher epistemic levels. They generated Super Notes to consolidate their knowledge advances, capturing sophisticated scientific explanations and questions developed in productive areas of inquiry. Social network analysis of who had read whose Super Notes revealed intensive connections formed among the students within each classroom, between different classrooms, and across school years (student cohorts). The findings further suggest potential opportunities for such cross-community sharing to stimulate deeper inquiry within each classroom and collaborative dialogue across the partner classrooms. However, the above studies only explored this potential in a preliminary manner due to a lack of technology support and systematic data collection tracing ideas across social spaces and classroom settings. The current study was part of the third iteration of our design-based research implemented with the new technological support of ITM. Analysis reported in an earlier paper has examined epistemic quality and complexity of students' Super Notes shared in the meta-space (Yuan, Zhang, & Chen, 2019). The data analysis reported in the current paper investigates the cross-classroom discourse among four Grade 5 science classrooms supported by the Super Talk function of ITM. Our analysis attends to the dynamic movement of ideas from each classroom to the Super Talk for collective knowledge building as well as the travel (incorporation) of ideas from the Super Talk to the discourse in each community. Our specific research questions ask: What knowledge advances were achieved in the cross-classroom Super Talk and how did the collective advances emerge from—and rise above—the works and ideas developed within each home classroom?

Method

Classroom Contexts

This study was part of a design-based research conducted in four Grade 5 classrooms at a public school located in Northeastern U.S. The participants included 89 students who studied human body systems as part of their science curriculum over a period of six months. Their inquiry of human body systems was implemented using a knowledge building pedagogy supported by ITM. The four classrooms were taught by two experienced teachers, each teaching science in two classrooms.

Knowledge Building Design and Implementation

At the beginning of the semester, students participated in a set of activities (e.g. apple tasting, high kicks, etc.) that triggered their interests and wonderings about the human body. Students then generated initial questions and clustered the questions based on the body systems involved. Students used multiple resources to support the inquiry of their questions, such as books, websites, online videos, and models. They shared ideas through face-to-face metacognitive meetings, where students built on each others' ideas to explore problems of understanding, reflected on idea progress, and identified problems and knowledge gaps for further study. This knowledge building discourse continued in ITM in their online space; teachers created each Wondering Area based on student-identified research topics. As students made progress in understanding how each body system functions, they started to create a reflective super note using ITM's Journey of Thinking (JoT) to synthesize the "big ideas" learned and questions for deeper research, leading to further activities in their home classes to advance their inquiry.

As the inquiry about the different body systems progressed in each classroom, at the beginning of May, students in Class 1 suggested a challenging question for the whole fifth grade to discuss using ITM's Super Talk function: "How do people grow?" Students from the four classrooms participated in this Super Talk over the next month and contributed ideas to solve the challenging problem. A total of 19 students from the four classrooms posted 22 notes in total in the cross-classroom discussion. At the beginning of June, a whole class metacognitive meeting was held in each room to share and integrate the knowledge that they had gained from the Super Talk and build connections with their own inquiries.

Data Sources and Analyses

The data resources included students' ITM notes posted in their home class space and the "Super Talk," researchers' detailed field notes, classroom video recordings, and students' notebooks. All four classrooms' science lessons were video-recorded and selectively transcribed. As Lemke (2000) suggested, understanding an ecosocial system needs to describe the interdependent processes which occur on a certain timescale. We adopted his suggestion to integrate multiple levels and units of analysis, with each unit interpreted in the context of the larger unit of analysis and elaborated using the more specific episodes involved. Specifically, our analysis traced how the ideas emerged from individual and small-group research interests in each home class and traveled to the cross-community Super Talk. The researchers applied temporal analysis to trace the core ideas developed in the Super Talk to explain how people grow. Based on the conceptual elements and their contributors, we further traced back to identify the related inquiry work in the contributors' home classrooms, as video recordings, field notes, and ITM online posts. Classroom videos and ITM online posts were further analyzed to identify when and how the ideas were generated, and by whom, with the major contributions of inquiry mapped out on a timeline. Students' notebooks and field observation notes were further used for data triangulation.

Results

What knowledge advances were achieved in the cross-classroom Super Talk?

To understand students' ideas generated in the Super Talk to explain how people grow, we analyzed the content of students' Super Talk posts and identified ten key conceptual elements. Each conceptual element explained the process of human growth from a specific angle, ranging from the growth of muscles and bones to digestion, brain control, growth hormones, and so forth. Conceptual connections were further identified based on students' discourse responses (e.g. build-on). The conceptual elements and connections are shown in the top layer of Figure 2, displayed based on the sequence of time. As noted above, based on the concepts and their contributors, we further traced backward to identify the related inquiry work in the contributors' home classrooms. The lower section of Figure 2 illustrates the first time each concept is shown in the home class, a dotted line connecting between the lower section and the top layer illustrates the information from the focal home classroom that fed student contribution to the Super Talk.

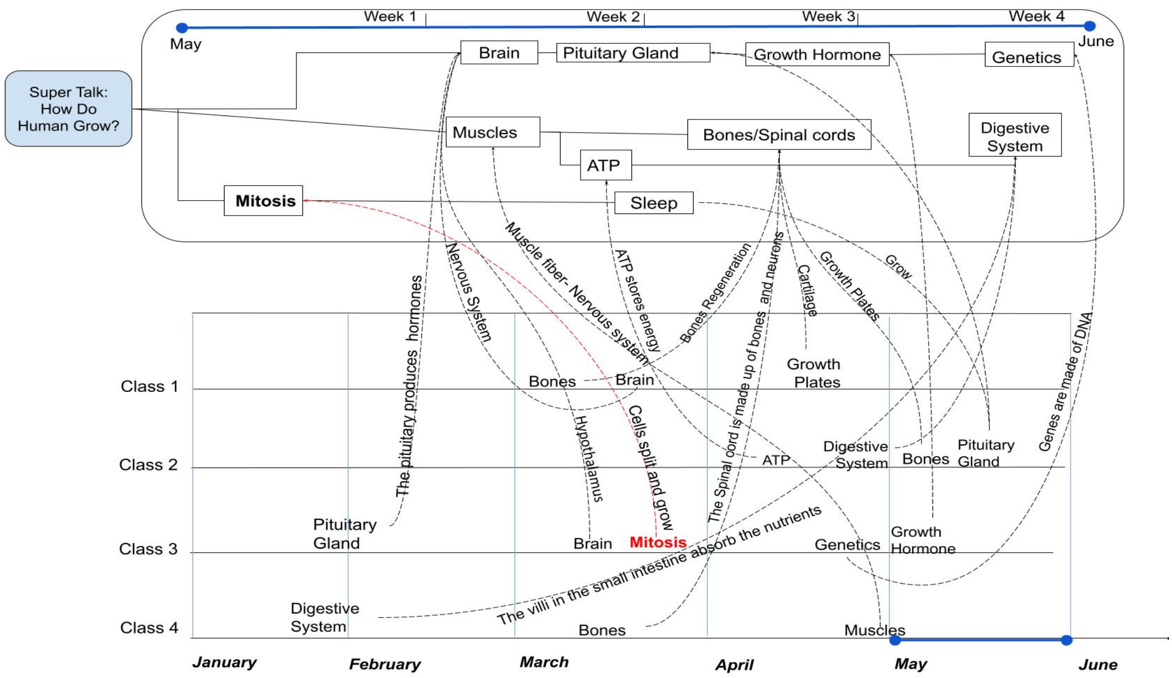


Figure 2. Tracing idea development in the Super Talk (upper area, from May to June) in connection with the related knowledge building work and discourse in each home class.

Students from Classroom 1 contributed to explaining “how do people grow” from the perspective of the brain, muscles and bones, with the insight that the pituitary gland controls the growth hormones and sends messages to the muscles and the joints, and that muscles grow by repairing rips. The key concept of ATP and bones were further built on by Class 2 with a key idea that bones grow through ossification and ATP is formed when muscles repair rips. A new perspective regarding the role of sleeping was added by students from Classroom 3: During NREM sleep, the body is repairing damaged tissues and growing. A new key concept of mitosis was incorporated to explain how cells grow in four stages. Ideas about genetics were later added by members of Classroom 3 to further explain what determines height. Members of Classroom 4 further highlighted the role of digestion: it breaks down and delivers nutrients throughout the body to help it grow. Extensive discourse occurred focusing the major concepts and systems (bones, muscles and brain) that were closely related to the main inquiry questions of each classroom. Other related ideas were further incorporated and built on in a reflection of students' special interests and expertise (genetics, mitosis, and ATP).

How did the collective advances emerge from—and rise above—the works and ideas developed within each home classroom?

Based on the above graph, we use the conceptual elements of muscle growth as related to mitosis as examples to analyze how student contributions to the Super Talk emerged from and rose above the inquiry work in each home classroom.

Classroom 1

Eight students from Classroom 1 participated in the Super Talk discussion from the perspective of bones, muscles, growth hormones, and sleeping; of those, six students mentioned how growth relates to muscles and bones. The topic of muscles originally branched out from the topic of the heart. At the beginning of January, a group of learners interested in the heart (Hugo, Jane, Maxwell, Nevan, and Otis) first investigated how the heart functions and problems caused by heart holes. As they accumulated enough knowledge, on March 5, the heart group held a metacognitive meeting with the whole class, during which they shared key information about how blood travels through the circulatory system and made a new connection between heart and bones (that ribs protect your heart). On March 15, the teacher talked to this group to see whether they had new research questions. Jane, who had focused on the skeleton, was inspired by the connection between the heart and bones and proposed new inquiry questions: “how did your bones heal?” and “how can bones make blood?”. The teacher created an idea thread in ITM for students' inquiry of the new research questions. Later, Maxwell, Nevan, and Otis, who were core members of the heart group, joined Jane to explore these issues. Their thinking about bones and muscles was deepened and elaborated over time to understand the various categories of bones (axial bones and appendicular bones), joints, bone fracture, and the treatment of snapped bones (put on a cast). Conceptual connections were built among the different human body systems such as by understanding how the bone marrow creates red blood cells and brain control of joint movement through sending nerve signals.

In the above context, in early May, students in Classroom 1 initiated the Super Talk topic of how people grow. The students working on bones and muscles were very motivated to share their knowledge in the Super Talk discussion space because it was closely related to their research topics. On May 9th, Nevan and Otis co-authored a note in the Super Talk to explain how the brain connects to the bones: “*Humans grow by the brain: the pituitary gland controls the growth hormones and sends messages to the muscles and the joints. The brain helps the body grow. The pituitary gland controls growth.*” Nevan also played an important role as a boundary broker to bring the concept of *mitosis* from Classroom 3 back to his home class.

On June 5, Classroom 1 held the last metacognitive meeting with a theme of how people grow to summarize their learning. Students participated in the discussion from their focused areas. At the same time, they integrated what they've learned from the Super Talk back into their conversation. When students were sharing the content that related to brain cells, Nevan brought back the information about cell mitosis that he read from the Super Talk and leveraged students' understanding about this cutting-edge concept and made connections with all other human body organs as cell mitosis is how each body organ grow at the base level.

Classroom 2

Classroom 2 contributed to the Super Talk discussion about how bones and muscles grow through building connections with digestion and cells. Tracing back to Classroom 2's inquiry journey, we observed that students in Classroom 2 first investigated issues related to the digestive system, brain, heart and lungs and blood in the first two months. One of the cross-cutting themes connecting these topics looked at how humans get and use energy from food. Students in the energy research group advanced their understanding by elaborating on the process of digestion; the digestive system breaks down food and further delivers nutrients through the bloodstream. On

March 4, a new connection was made between the digestive system and muscles by Frank, who posted in ITM:"*…ATP is what 'charges" your body… when you eat, ATP is made which then powers up your body… if your body is low on ATP, it will be stored in your muscle cells… ATP is your body's main energy source* ". On April 26th, students who researched muscles made a cross-system connection and started to examine "what is protein?" and "what is a cell?"

After Classroom 1 initiated the Super Talk topic of how people grow, on May 11, Teacher Mrs. Harris held a whole class metacognitive meeting in Class 2 to advertise the Super Talk topic. Students first read the notes that were already posted by several peers from the other classrooms, discussed and analyzed how Class 2 can learn from the cross-classroom discussion and further add to it. After reading the existing notes, students found that although the existing notes talked about how muscles grow (by fixing the rips that were caused by extra force) and shared basic information (e.g. bones grow as you grow), the information posted had not fully answered the question of how bones and muscles grow. The teacher highlighted the importance of posting non-redundant information to advance the collective understanding and explaining HOW people grow. After this meeting, a few students worked on explanations of how bones grow, drawing upon the above-noted inquiries about bones, muscles, digestion, and cells. Henry, who first worked with a few peers on energy research and later joined the bones and muscles group, posted a new note in the Super Talk thread that built on an existing note about bones. He wrote: "*Babies are born with 100 more bones than adults, the bones fuse together to make longer bones as we grow. What babies have is not really bones, it is cartilage. With the help of calcium, the cartilage gets turned into bones through the process of Ossification.*" His classmate Frank read Henry's note and further built on it by saying: "*I might have a little more info to help you. Over time, a different type of cell called osteoclasts head to the middle of the bone to help in. Now, inside osteoclasts, there are hydrolytic enzymes and acids. These enzymes and acids will help dissolve the temporal bone (the cartilage) to make room for the permanent bone (marrow). Also, Ossification will take around 20 years. Once this process is over, the bones will not grow anymore, but will still be able to heal themselves in case you get any unexpected fractures."*

Classroom 3

In Classroom 3, the topic of muscles and bones emerged relatively late in mid-March involving only two students. The two students did not post in the Super Talk discussion. However, students who studied cells made active and unique contributions to the Super Talk discussion, highlighting the role and process of mitosis. Below, we trace how their ideas were developed within their group and classroom and contributed to the cross-classroom discussion.

In Classroom 3's human body inquiry, one of the most productive lines of inquiry investigated the function and structure of the brain. As a specific insight, students found that the pituitary glands in the brain release hormones. This topic was further connected to the inquiry about lungs. Students from the lungs group found that the brain and lungs work closely together, noting that oxygen gets to the tissues (including those in the brain) through red blood cells (Week 5), and tissues in the body need oxygen (Week 6). Blake, a key member of the heart and lungs group, contributed his knowledge about cells during a metacognitive meeting: "*The cells contain sugar except they need the oxygen to turn it into energy*." In week 7, the concept of the cell was expanded to consider white blood cells, such as through Blake's build-on: "*Neutrophils look for things that shouldn't be in your body, and macrophages look for and digest dead germs…Amino acids are what make proteins.*" In a whole class metacognitive meeting, the teacher asked: "*What tissue of our body needs oxygen?"* Students said: "*Everywhere, because we need our oxygen to survive*." The understanding of tissues and cells was further deepened on March 15th when Blake introduced a key concept related to human growth: "*Mitosis is the process of one cell splitting into two new cells as it is a complex process with many steps*". In the same week, Blake suggested that the teacher create a new thread of discussion called "How do we grow?" This thread was created in Class 3's own discussion space. However, this topic did not get much attention from Blake's peers in Class 3.

Blake's idea about mitosis did not catch others' interest until May when Classroom 1 initiated the Super Talk topic asking exactly the same question. Blake was thus able to connect with other peers from the whole Grade 5 who were interested in exploring how people grow. He joined in the collaboration, with his early note about mitosis copied to the cross-classroom Super Talk thread in ITM. This idea further caused Nevan's (a student from Class 1) attention. After reading Blake's note, Nevan brought the knowledge about mitosis to his home class discussion and extended his peers' understanding and conversation during their last metacognitive meeting.

Classroom 4

In Classroom 4, the topic of muscles and bones sprouted from the inquiry about the immune system. Tim and the other two members first investigated the topic of the immune system at the early stage with a guiding question: "What happens with blood cells in the immune system?" This idea was first explained by Tim in the first month, who wrote: "*Your immune system is a process of white blood cells that kill bacteria, the white blood cells in the immune system are Leukocytes*." This idea was further connected with the inquiry about bones. Tim posted in the fourth week: "*Bone marrow, a tissue inside of your bones, makes white blood cells which enter a system called the lymphatic system, which helps your body from getting diseases… There are 2 different types of blood cells, they are phagocytes and lymphocytes. When a phagocyte sees a virus, it immediately sends a signal to lymphocytes to make the correct antibody for a virus. …cells and antibodies sort of have a mind of its own when the immune system gets a virus.*" From the second month, the inquiry of the immune system was expanded to include HIV and the lymphoid. On May 3, during a metacognitive meeting, the teacher emphasized that May is the "Month of Connection." One of the learning activities was finding connections among human body systems. Tim pointed out a connection by saying: "*Muscles are a huge part of your body. Without muscles, you couldn't blink, jump, smile or have your heartbeat. There are 3 types of muscles: skeletal, cardiac and smooth muscles.*"

After the teacher introduced the Super Talk topic of how humans grow to Classroom 4, Tim first read the notes already posted in the Super Talk, making connections with his understanding about the immune system and muscles. He then contributed to the Super Talk by adding a detailed explanation about how muscles grow: "*Muscles grow by when you stress muscle fibers, by lifting heavy weights or doing motions that you're not used to. They rip which lets out a chemical called cytokines, which activates your immune system and repairs it bigger than it was earlier, thereby making your muscles grow. Hypertrophy is how your muscles say you need to work more to make your muscles grow. If you stop exercising, your muscles will go through a process called muscular atrophy which makes your muscles shrink.*" This detailed answer advanced the understanding of the overarching question one step further.

Discussion and conclusion

This research explored students' collaborative interaction unfolding across emergent social levels, which included the local knowledge space of each classroom community and a meta-space (macro space) shared across communities. As students in each home classroom pursued progressive inquiries to deepen their understanding of the various human body systems, they shared knowledge advances with the partner classrooms using reflective Super Notes (see analysis in Yuan et al., 2019) and further pursued cross-classroom Super Talk to address a challenging problem. The Super Talk problem was not predetermined but emerged based on student interests at the intersection of the different lines of inquiry about the various body systems. The analysis of the cross-classroom Super Talk in connection with the knowledge building work in each classroom provided a detailed account of how students worked across the social levels to continually advance their knowledge. The multi-layer design enabled students from multiple communities to collaboratively solve the challenging problem, building on the interests and knowledge developed in each community. The "Super Talk" served as the cross-boundary meta-space where students formed extensive social connections and integrated distributed expertise to develop higher-level understanding. Students' multiple points of view (e.g. bones, brains, hearts) and diverse inquiry strategies came into contact in the dynamic interactions as they contributed their special knowledge about the different body systems and processes to explain the holistic problem of how people grow. With their teachers' facilitation, students read and learned from their peers' notes in the Super Talk, identified gaps and missing links, and further contributed their knowledge and perspectives. Some of the new knowledge gained from the Super Talk was further shared and discussed in the individual classrooms to complement and expand their own inquiry.

The above findings further enrich what we have learned through prior studies (Yuan & Zhang, 2019; Zhang et al., 2020), showcasing students' dynamic idea interactions for ever-unfolding inquriry as expanded and transformed through cross-community collaboration. The classroom processes and findings shed light on opportunities and strategies to design a larger creative socio-technological context that is critically needed for scaling collaborative learning across classrooms. Designs of cross-community knowledge building among students should capitalize on the power of different levels of discourse and create a synergy between the social extension and epistemic rise-above of ideas. With the interactive discourse within each group and community supporting continual idea improvement and diverse expertise, cross-community discourse provides a larger and higher-level space for students to further share and integrate their knowledge advances to tackle cross-cutting challenges and develop more sophisticated understanding, which may further leverage students' inquiry and thinking in each community.

References

Cress, U., Oshima, J., Rosé, C., & Wise, A. (in press). Foundations, processes, technologies, and methods: An overview of CSCL through its handbook. In U. Cress, J. Oshima, C. Rosé, & A. Wise (Eds.), *International handbook of computer-supported collaborative learning*. Berlin: Springer.

Csikszentmihalyi, M. (1999). Implications of a systems perspective for the study of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 313-335). Cambridge, UK: Cambridge University Press.

Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist*, *47*(3), 215-231.

Engeström, Y. (2014). *Learning by Expanding* (2nd edition). Cambridge, England: Cambridge University Press.

Laferriere, T., Law, N., & Montané, M. (2012). An international knowledge building network for sustainable curriculum and pedagogical innovation. *International Education Studies*, *5,* 148-160.

Lai, M, & Law, N. (2006). Peer scaffolding of knowledge building through collaborative groups with differential learning experiences. *Journal of Educational Computing Research*, *35* (2), 123-144.

Lemke, J. (2000). Across the scales of time: Artifacts, activities, and meanings in ecosocial systems. *Mind, Culture, and Activity, 7*, 273-290.

Sawyer, R. K. (2007). *Group genius: The creative power of collaboration*. New York: Basic Books.

Sawyer, R. K. (2015). *Social emergence: Societies as complex systems*. New York, NY: Cambridge University Press.

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.

Stahl, G. (2013). Learning across levels. *International Journal of Computer-Supported Collaborative Learning, 8*(1), 1-12.

Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, 12(4), 423-467.

Yuan, G., & Zhang, J. (2019). Connecting Knowledge Spaces: Enabling Cross-Community Knowledge Building through Boundary Objects. *British Journal of Educational Technology*, 50 (5), 2144–2161.

Yuan, G., Zhang, J., & Chen, M. (2019). Cross-boundary interaction for sustaining idea development and knowledge building with Idea Thread Mapper. In Lund, K., Niccolai, G. P., Lavoué, E., Hmelo-Silver, C., Gweon, G., & Baker, M. (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, Volume 1* (pp. 456-463). Lyon, France: International Society of the Learning Sciences.

Zhang, J., & Chen, M.-H. (2019). Idea Thread Mapper: Designs for sustaining student-driven knowledge building across classrooms. In C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), Proceedings *of the International Conference of Computer-Supported Collaborative Learning* (CSCL 2019). Lyon, France: International Society of the Learning Sciences.

Zhang, J., Tao, D., Chen, M. Sun, Y., Judson, D., & Naqvi, S. (2018). Co-organizing the collective journey of inquiry with Idea Thread Mapper. *Journal of the Learning Sciences*, 27(3), 390-430.

Zhang, J., Yuan, G. & Bogouslavsky, M. (2020). Give student ideas a larger stage: support cross-community interaction for knowledge building. *International Journal of Computer-Supported Collaborative Learning, 15*, 389–410.

Acknowledgments

This research was sponsored by National Science Foundation (#1441479). We thank the teachers and students for their creative work enabling this research. We also extend our thanks to our research team members for their contributions to the technology development and classroom data collection.

Scripting Small Group Processes within a Learning Community

Yanhong Li, LMU Munich, yanhong.li@ifi.lmu.de
Daniel Sommerhoff, IPN - Leibniz Institute for Science and Mathematics Education, sommerhoff@leibniz-ipn.de
Stefan Ufer, LMU Munich, ufer@math.lmu.de
James D. Slotta, University of Toronto, jim.slotta@utoronto.ca

Abstract: Building on research concerned with scripting and learning communities, this study explored how to script small group processes within a larger community-wide script. Small group scripts, Peer Instruction (PI), Community Supported Worksheets (CSW), and Community Knowledge Construction (CKC), were designed and implemented in an online preparatory mathematics course for 181 freshmen. The completion rate and completion quality of group activities were analyzed. Except for CKC activities, PI and CSW had a satisfactory completion quality. We analyzed the impact of group activities on students' epistemological beliefs about learning communities, and also performed content analyses of students' ideas and artifacts, to show the reciprocal influence between the community and small groups. Results show students had a significant agreement that the whole community is an important source for learning. Meanwhile, after taking this course, they had a more profound conceptual understanding of the context, purpose, means, and challenges of the learning community.

Introduction

In recent years, there has been an increasing interest in scripting for instructional design. As new technologies enter the wider practices of teaching and learning, we are seeing a surge of interest in phenomena like "flipped classrooms" (Akçayır & Akçayır, 2018) and "active learning" (Beichner, 2012), in which students are engaged in dynamic interactions with peers, leveraging collaboration and Computer-Supported Collaborative Learning (CSCL) techniques and technologies (Slotta, Tissenbaum, & Lui, 2013). There has also been some research in the learning sciences about the structure and discourse patterns that occur within such learning designs, which includes ideas about collaborative groups (Dillenbourg & Jermann, 2007; Weinberger, Kollar, Dimitriadis, Mäkitalo-Siegl, & Fischer, 2009), design teams (Kozlowski, 2018), and learning communities (Bielaczyc & Collins, 2009; Slotta, Quintana, & Moher, 2018).

The present study builds on a body of research concerned with scripting (Dillenbourg & Jermain, 2007; Kollar, Fischer, & Slotta, 2005; Weinberger et al., 2009), with a particular interest in prior studies of the role of external collaboration scripts in relation to participants' internal scripts or knowledge (Kollar, Fischer & Hesse, 2006; Kollar et al., 2005). This research also builds on prior work concerned with learning communities or collective inquiry (Slotta et al., 2018), which argues for the importance of scripted interactions that allow community knowledge to take from and serve as a resource for subsequent (also scripted) inquiry within the community. In particular, we examine whether small group scripts can gain structure and definition, as well as valuable inputs, from being situated within a larger community-wide script. Jigsaw designs (Aronson, 1978) are a common example of such, where the specific scripts that guide several small specialist groups are designed to fit within a larger script to recombine those groups such that knowledge and products developed by various small group specialists become available across the community. The current paper builds on specific principles of learning communities (e.g. *Sharing Principle and Structural-Dependence Principle*) articulated by Bielaczyc and Collins (2009), to interconnect small group scripts within a broader community, in a math course for freshmen, focusing on logic and mathematical proofs. We examine the impact of such connections on students' epistemological beliefs about the value of community in learning and also perform content analyses of student ideas and artifacts, to show the reciprocal influence between the community and small groups. We close with a discussion of considerations that are important to the CSCL research community.

Literature Review

Learning community and group process

The term *learning community* refers to advancing the collective knowledge to support the growth of individual knowledge (Scardamalia & Bereiter, 1994), where everyone is involved in a collective effort of understanding (Bielaczyc & Collins, 2009). In a learning community environment, individuals benefit from: (1) learning in a social constructivist environment to construct knowledge (Palincsar, 1998); (2) multi-cultural communication, where diverse cultural backgrounds are valued (Cifuentes & Murphy, 2000); and (3) extending individuals' Zone of Proximal Development with the collective knowledge of learning community (Hung & Chen, 2001). However, learning communities have complex social, cultural, and cognitive situations (Hung & Chen, 2001), which make it difficult to build a learning process with a vibrant and sustaining sense of community. Guiding a large number of students through a CSCL environment including facilitation of specific activities and providing feedback is a challenging task (Weinberger et al., 2009). The distribution of a global whole community process over different individuals or groups is a mechanism commonly exploited in CSCL scripts (Dillenbourg & Jermann, 2007). Small groups are like microelements, which interact and consist of the whole community. The interconnections of small group processes create opportunities for knowledge building and leveraging the collective resources of the community (Slotta & Peters, 2008). By focusing on small group processes, we aim to make learning communities more feasible and effective. A related area of work from the practitioner community is concerned with active learning (Beichner, 2012), where many different forms of interaction have been explored. Especially these three small group scripts: (1) Peer Instruction script (Fagen, Crouch, & Mazur, 2002; Mazur, 1997): students are engaged individually, in small groups, and as a whole class in reflecting on patterns of responses to carefully crafted multiple-choice items; (2) Community Supported Worksheet script (Li, Dai, Wang, & Slotta, 2020): students work on a difficult problem in a small group to find a correct solution. Groups are asked to provide solution hints to help other groups who have difficulties; (3) Community Knowledge Construction script (Slotta & Peters, 2008): students contribute to a shared knowledge base collectively to reflect and consolidate their understanding.

Personal epistemological development and epistemological beliefs have attracted researchers' interest since the late 1980s. Epistemological beliefs refer to learners' beliefs about the nature of knowledge and the process of its acquisition (Magolda, 1992). They can shape students' engagement in learning communities by influencing their cognitive thinking and reasoning (Peer & Lourdusamy, 2005) and active involvement in the learning process (Magolda, 1992). Previous research has found that a change of epistemological beliefs could help students understand the meaning and effects of learning science and learning communities (Slotta & Peters, 2008). However, epistemological beliefs are not easy to change (Peer & Lourdusamy, 2005). To some extent, a person's epistemological belief is a context of how knowledge is accessed, which comes from an accumulation of previous learning experience. As we know, learning communities are a culture to seek a collective effort of understanding (Bielaczyc & Collins, 2009). An effective learning community approach will influence students' epistemological beliefs in a productive way (Acosta et al., 2014). Thus, the change of student epistemological beliefs can be evidence for having a good learning community approach (Li et al., 2020).

Collaboration scripts

CSCL allows a wealth of new affordances for learning within the groups. However, learners find it hard to engage in productive collaboration processes without guidance (Weinberger et al., 2009). Kollar et al. (2005) introduced the construct of collaboration scripts as one means of providing such guidance. They used carefully constructed scaffolds to support pairs of students who created structured arguments concerning scientific debates. This study found that a highly structured external collaboration script supported the acquisition of *domain-general knowledge* of all learners regardless of their internal scripts. Ensuing work, conducted by Vogel, Kollar, Ufer, Reiss, and Fischer (2016), examined scripting in the context of a higher education mathematics course. It found that a highly structured domain-general collaboration script for argumentation was more effective than a less structured one to acquire disposition to use argumentation skills.

While collaboration scripts offer an interesting form of scaffolding for small group processes, there remains a wider question concerning the scripting of an entire class community, as it progresses through topics, activities, and assessments. Whole class scripts have been described by Dillenbourg, Nussbaum, Dimitriadis, and Roschelle (2013) as a way of offering higher-level guidance and structure to support the classroom community. For example, in the Concept Grid script (Dillenbourg & Jermann, 2007), the class is presented with a two-dimensional grid of concepts that must be addressed collectively, such that students must choose open squares to ultimately complete the grid. Such scripts are often described in close conjunction with the notion of orchestration (Dillenbourg et al., 2013; Slotta et al., 2013), such that the interaction of individuals, small groups, and the class as a whole is scaffolded jointly by the instructor and supportive CSCL technologies. Slotta and his colleagues (e.g., Slotta & Peters, 2008; Slotta et al., 2018) have advanced a model of scripting for learning communities called Knowledge Community and Inquiry (KCI). Dillenbourg et al. (2015) introduced the notion of an orchestration graph, to describe the shifting patterns of discourse and activity across social planes (e.g., individual, small group, whole class) that support smooth orchestration of activities within such designs.

However, while CSCL researchers have made advances in the forms of scripting and orchestration for whole-class inquiry, there remains a gap between the fine-grained studies of scaffolded collaboration (e.g., Kollar et al., 2007; Vogel et al., 2016), and the community level scripts such as those of Dillenbourg and Jermann (2007) or Slotta et al. (2018). Given the emphasis placed by CSCL on the importance of social practices within a community of learners (Kollar et al., 2006), well-designed collaboration scripts should serve to support group processes, enhance individual learning, but also reinforce exchanges amongst the wider community of learners (i.e., between a group and other groups or with the community as a whole). Scripts addressing both small groups and the community level are also supposed to support knowledge construction within the community, and the use of that knowledge as a resource for inquiry (Slotta et al., 2018). An important question for further research is concerned with how to define collaboration scripts such that they promote effective individual (and small group) learning as well as productive exchange amongst peers within a classroom community (Dillenbourg et al., 2013; Kollar et al., 2007).

There has been some research about how small group scripts help the learning communities. For example, "jigsaw" designs establish small groups that specialize in one aspect of the topic, then recombine into new small groups (each of which include at least one member who specialized in each of the previous topics) which serves to support the wider learning community (Dillenbourg & Jermann, 2007). Slotta et al. (2013) report on the use of scripted small groups within a KCI curriculum, where small groups were responsible for different parts of the inquiry, contributing to the progress of the community as a whole. Because small group activities were situated within the context of the broader community inquiry, this allowed new affordances for epistemic and pedagogical designs (Slotta et al., 2018). However, these studies did not explicitly address the specific guidance and scripting of small groups within the context of the broader scripts for the learning community. While there were small groups present within the designs, and these were instrumental to the collective progress, the specific nature of the scripting for these groups was not a formal matter of study.

Research questions

The present study seeks to define specific small group interactions in the context of a broader community of inquiry. The prior findings of this scripting research were helpful to guide our designs for small groups but neglected to include the interface with a learning community. This work will build on previous studies of Kollar et al. (2007) and Vogel et al. (2016), engaging students in the same higher education mathematics context, but with an additional level of scripting across the group and whole class contexts. The group processes were designed to explicitly engage the community context, making beneficial knowledge contributions, and gaining important community inputs. In order to explore what group processes may bring to a learning community and how small group collaboration scripts facilitate learning interaction and collective knowledge sharing, two research questions are addressed: **RQ1:** In what ways can we design scripts for small group processes to support and benefit from a learning community? **RQ2:** What changes in students' learning behavior and epistemological beliefs of a learning community's role in individual learning can be identified over a course using such scripts?

Methods

Context and Participants. The study was conducted within a two-week preparatory course for prospective mathematics university students in Germany. The course was offered before the beginning of their first semester to support them in the transition from secondary school mathematics to university mathematics. The class was held in German and contained twelve asynchronous online lectures and ten tutorial exercises on elementary number theory and other mathematical topics (e.g., basic propositional and predicate logic, proof techniques, induction, and recursion). Participation in the course was voluntary. Overall, 181 students registered on the learning platform, who were distributed in seven different tutors' classes. Finally, 129 (71.27%) students were included in the analyses, because they (1) agreed to participate in this study, (2) completed the course, and (3) took part in all learning activities and test sessions. As shown in Table 1, the gender distribution is nearly equal with 65 females and 64 males. The mean of their ages is 19.11, which ranges from 17 to 24.

Table 1. Number, gender, and age of participants

	Tutor 1	Tutor 2	Tutor 3	Tutor 4	Tutor 5	Tutor 6	Tutor 7	All
Registered students	31	27	26	25	25	26	21	181
Participants	27	26	6	14	18	21	17	**129 (71.27%)**
Female	10	19	3	6	7	9	11	65 (50.39%)
Male	17	7	3	8	11	12	6	64 (49.61%)
Age	19.00 [17, 24]	18.63 [17, 21]	19.33 [18, 24]	19.09 [17, 21]	19.47 [17, 24]	19.35 [17, 24]	19.33 [18, 23]	19.11 [17, 24]

Material and activity design. The course had two parts: (1) Watch lecture videos asynchronously and autonomously; (2) Participate in synchronous Zoom tutorial meetings (90 mins per one), which were conducted by seven mathematics tutors. Eight mathematical topics, such as logic, quantifiers, and divisibility, were addressed within the course. For each topic, three or four small group activities were designed for the tutorials. Materials used in the activities were designed by one mathematical lecturer, an experienced instructor for the subject matter. Meanwhile, both lecturers of this course improved and confirmed the use of these materials. Thus, the activity materials used were suitable for study purposes. All students were assigned to these seven tutors randomly and equally. A learning platform named SCORE (SCripting and ORchestration Environment) was used to implement the learning activities. The student epistemology belief survey (Acosta et al., 2014; Madhok et al., 2010) was adapted for pre-post tests, which had two multiple-choice questions (*1. What are your main learning methods?* and *2. What will you do when you have a learning problem?*), two five-point-Likert questions (*1. Discussing with my classmates helps me learn better; 2. The class community (all students in the class, considered together) is an important resource for my learning*) and one open question (*What do you think is a "learning community"?*). The Likert scale was from 1 (*strongly disagree*) to 5 (*strongly agree*).

Small group scripts. Based on the principles for the design of effective learning communities (Bielaczyc & Collins, 2009), we designed activities that help students expand the community's knowledge (*Community-Growth Principle*), and advance the overall quality of knowledge (*Quality-of-Products Principle*). In order to connect small group participants with the whole class community, three scripting patterns were designed and implemented (*Multiple-Ways-to Participant Principle*): Peer Instruction (PI; Mazur, 1997), Community Supported Worksheet (CSW; Li et al., 2020), and Community Knowledge Construction (CKC; Slotta & Peters, 2008). These scripts have been used in the authors' previous studies and applied in this study, for purposes of addressing the research of small groups within a learning community.

PI (*Sharing Principle*): Ten multiple-choice question tasks were implemented. The first two tutorials had two PI tasks each time, the other six tutorials had one per time. All tutors used the same tasks. The PI scripts had three stages: (1) Individual students submitted their answer; (2) Students were shown the combined answers from all members of their tutorial group, as well as the wider classroom community; The answer distribution charts changed when more students submitted their answers, allowing students to see dynamic community responses; (3) Students were asked to reflect: "*What is the difference of the answer distribution between your group and the whole class? What do you think is the correct answer? Is there anything that surprises you?*". Their answers were recorded in the learning systems as the discussion data.

CSW (*Structural-Dependence Principle*): Nineteen CSW activities were designed and implemented as well. Each tutorial had 2-3 activities and all tutors implemented the same activities. Each CSW included four steps: (1) Students were assigned to collaborate in small groups with 3-4 students in the Zoom breakout rooms; (2) A math worksheet was given to them to solve together; (3) If the group had completed the worksheet, they created a hint and provided it to other groups; If the group had difficulties, they could go see the hints made by others; (4) Students were asked to give feedback about the usefulness of hints.

CKC (*Quality-of-Products Principle*): Knowledge base templates for the course were created to invite students to contribute their understanding. There were eight lecture topics in all. Students were given the knowledge base document link after finishing the corresponding tutorial. CKC had two steps: (1) Knowledge base templates were created, which had the modules "*Key ideas we learned*", "*Why this topic is important in math*", "*Help request*", and "*Suggestions*"; (2) After finishing the learning of each topic, students were invited to co-write in a shared document to reflect on their learning.

The above scripts are seen to interact with the whole community on different levels. In integrating the PI script, individuals are engaged in thinking about the problems independently, then have opportunities to identify one's own position within the group and the group's situation within the whole community regarding the tasks. This is a micro script to help individuals benefit from the collective knowledge of the whole community. Unlike PI script, CSW aims to improve communication among small groups. This script engages individuals in "face-to-face" small group activities within Zoom breakout rooms. Connections to the community are of the form seeking help (*benefit from the community*) and giving help (*contribute to the community*). Moreover, "hints" (not "answers") can push small groups who provide help to think deeper because they need to diagnose possible difficulties. CKC is a critical script to connect the whole community. It has three roles: (1) collecting the inputs from PI and CSW; (2) collective knowledge contribution for summarizing and organizing what they have learned; (3) shared space for communication to sense the presence and benefits of the whole community. In all, PI, CSW, and CKC were designed to elaborate as small group process scripts to support the whole community.

Data sources and analyses. Data in this study came from (1) pre- and post- questionnaires, (2) data from learning platforms (i.e. SCORE and shared knowledge document). Data analyses were conducted based on the following steps: (1) Data preparation: data were put together from the sources mentioned above, anonymized, and prepared for analysis. (2) The qualities of each small group activity enactment were mainly evaluated by the research assistant and first author from 1 (*very bad*) to 5 (*very good*). The score was given for each student in each activity. More specifically, PI quality was determined by students' reflection (step 3 for PI) on answer distribution charts. CSW quality was decided by solution hint posts and help-request replies, CKC quality was the number of entries entered by students. (3) Qualitative content (e.g. open-ended survey questions, participants' responses to math problems, and knowledge base data) analysis was conducted by the research assistant and first author; (4) All results were translated from German into English by the research assistant (native German speaker) who was good at English.

Results

Here we begin by reviewing the outcomes of our small group activities, followed by an analysis of the impact on students' perception of the wider community's role, and finally an evaluation of the impact of the community on students' epistemic beliefs (i.e., about the importance of learning from peers and the learning community).

Participation in small group activities

Because the individual tutors varied in their priorities and approaches, they adopted the designed activities to a different extent. As shown in Table 2, PI activities had the highest completion (70%, varying from 20% to 100%). CSW had the least completion (51.9%, varying from 20.1% to 79.0%). Regarding the quality of the activity completion, PI scored the highest (3.53) and CKC the worst (2.39). As we can see, higher activity completion tended to have better quality. Overall, more than half of the activities were implemented by students from these 7 tutors, although some tutors' participants (e.g., # 3 and 4) completed less and with lower quality. A post-test question was designed to ask students which was their most favorite script. One hundred and four (80.62%) participants submitted their answers: 39.42% of them chose PI, 10.58% chose CSW, 6.70% chose CKC. The other 43.27% of participants had no strong preference. This might imply students would not engage in the learning community just because of a strong preference for only a specific script.

Table 2. Completion rate and quality of small group activities

Tutor #	Peer Instruction		Community Support Worksheet		Community Knowledge Construction	
	Completion	$M_{quality}$(SD)	Completion	$M_{quality}$(SD)	Completion	$M_{quality}$(SD)
1	100%	4.0 (1.03)	73.7%	3.2 (1.21)	100%	3.6 (1.32)
2	100%	4.2 (1.23)	79.0%	3.6 (1.01)	87.5%	3.3 (1.04)
3	20%	2.4 (0.93)	20.1%	2.3 (0.73)	0	0
4	30%	2.7 (0.88)	26.3%	2.5 (0.84)	12.5%	1.4 (0.83)
5	70%	3.5 (1.32)	52.5%	3.1 (1.03)	62.5%	2.7 (0.94)
6	90%	3.8 (0.96)	63.2%	3.2 (0.83)	75.0%	2.9 (1.21)
7	80%	4.1 (1.21)	47.4%	2.8 (1.12)	62.5%	2.8 (0.94)
All	**70%**	**3.53 (1.08)**	**51.90%**	**3.00 (0.97)**	**57.10%**	**2.39 (1.05)**

Influences of learning community approach

Three findings of influences of learning community approach (i.e. after the practical experience of the small group processes script activities in this study) were obtained: First, participants were asked about "what was your main learning method before" in the pre-test. The response to "attend the class" was 43.4%, "study alone" was 47.3%, and "learning with friends" was 9.3%. As we can see, more than 90% of participants didn't have a "learning community" approach as their main learning method before. Second, pre-post tests on students' preference for help-seeking showed participants had an increased preference at the end of the class to "ask peers" for help when they have a problem, which rose from 32.8% answers to 63.3%. In correspondence, the response to "I prefer to search the answer by myself" had decreased obviously from 38.0% to 12.4%. The choice "I prefer to ask the course teacher or tutor" had a minor decrease from 29.5% to 24.8%. Overall, students had an increased preference for the "peer learning" method. Finally, as shown in Table 3, there is a significant improvement in students' perception of the whole class community as an important source. Their perception of peers' help did not change significantly with means of 3.91 and 3.82 respectively.

Table 3. Paired *t*-test analysis of student perception of the learning community

Questions	Test	Mean	*SD*	*t*
Discussing with my peers helps me learn better	Pre-test	3.91	0.96	0.89
	Post-test	3.82	0.91	
The whole class community is an important source for my learning	Pre-test	3.40	1.06	-1.93*
	Post-test	3.61	0.92	

Note: $^*p < .05$

Student responses to the pre- and post- survey about the nature of communities revealed (in an open coding) four key dimensions, along which their ideas were seen to shift. The first is *Context,* which refers to the setting of the learning community. The second is *Purpose,* which has to do with why we need a learning community. The third dimension is *Means,* which is concerned with how to learn within a learning community. Finally, *Challenge* refers to the difficulties of working in a learning community. Table 4 provides some examples of how students' ideas shifted across these dimensions. As we can see, these were subtle conceptual understanding changes but implied students had a substantial epistemological change. For example, the same student in the pre-test mentioned the "Purpose" of the learning community was to "benefit from the strengths of others". This answer pointed out the advantage of learning community pedagogy. However, it didn't show an understanding of the "strengths" meant, which became clearer in the post-test response as "different perspectives".

Table 4. Examples of students' understanding difference of learning community

Themes	Pre-test	Post-test
Context	"spend time outside of the university"	"do something in a friendly atmosphere"
	"meet at agreed times to deal with a topic together"	"Giving and taking knowledge"
Purpose	"benefit from the strengths of others"	"better understand them through different perspectives"
	"have a higher chance of success"	"more effectively and, above all, more pleasantly in a group"
Means	"learn efficiently and support and help one another"	"talk to each other about the different solutions and thus find the best solution together"
	"collaboration and gathering of students in the same subject area in order to enable more successful and efficient learning"	"coming together and working together on the same topic (mathematics for us) in order to gain the greatest success from learning"
	"come together to learn and help one another"	"work together on tasks, develop possible solutions, help each other and fill in gaps in knowledge"
Challenge	"a group of students who try to work on topics together and support each other to better understand learning content"	"On the whole, I was able to work very well with the other participants in the tutorial, but it was sometimes difficult because one or the other was sometimes very quiet and you didn't really work together."

Discussion and conclusion

This study demonstrated that small group processes scripts could help individuals become better connected to the wider class community, where individuals contribute knowledge and resources to the community and gain helpful hints and information. According to principles for the design of effective learning communities, three small group scripts were designed and implemented. Three main findings can be highlighted in this study: First, simple small group scripts, such as PI, have a higher activity completion rate and completion quality than more complex scripts (i.e. CSW and CKC). At the same time, PI is the most favorite script for participants, which was adapted from Mazur's (1997) original F2F script. The advantage of an online PI script is to make students have more opportunities to do a deep self-reflection of the solution of problems because writing down thoughts needs more mental engagement, especially for mathematics (Peer & Lourdusamy, 2005). Second, student engagement in the learning community depends on many factors. From the data analysis of the most favorite scripts, most students didn't show a strong preference for a specific small group script. We can interpret this result from two perspectives: (1) small group scripts are better to be designed more diverse (i.e. *Multiple-Ways-to Participant Principle*; Bielaczyc & Collins, 2009); (2) in order to form a productive learning community, group processes should be considered from a more flexible and dynamic perspective. Third, students' learning preferences and epistemological beliefs can be changed by participating in small group script activities. The analysis results showed students had an improved preference for the "peer learning" method and a more profound conceptual understanding about context, purpose, means, and challenges of learning community pedagogy. As we mentioned, the participants were freshmen. It is easy to think the above changes might result from the transition of a more lecture-based high school teaching methods to a rather new, open and collaborative university lecture environment. However, it seems that this is not common for university mathematics students to have a good collaborative learning experience.

In addition, from our findings, we see two main issues. The first issue is the limited group activity time. To address this problem, we see the need to design more elaborate scripts, especially for supporting CSW and CKC. Basically, these two scripts are suitable for activities that have enough time or performed asynchronously. However, if external collaboration scripts are designed properly, students might work more effectively or benefit from the support from learning communities within a limited time. In other words, the scripts should provide clear instructions on how to work on the activities and how to interact with each other in different situations. This will be our next step to iterate this study. The second issue is how to connect support levels of individuals, small groups, and the community successfully. In this study, interactions from different perspectives were designed to make the three small group process scripts more interconnected. The interconnective design is very helpful because it provides broader opportunities for students to access the whole community. It means group processes should be understood from a more global perspective, where the community is in the center but with flexible and various interactions between groups. In the future study, we plan to make the small group processes more connected with the whole community and think about how to measure this interconnection. In this study, the evidence of interaction among the group process scripts is not sufficient, but this is a critical problem related to the effectiveness of small group processes script design.

With the group process scripts, we aimed to help individuals benefit from a broader community. Indeed, students reported a higher agreement that the whole class community was an important source for their studies. It means that they realized the value of the learning community. However, they didn't show an improved perception that discussing with their peers helped them learn better. This also reflects their low preference for CSW script. When analyzing students' answers on "what is a learning community", the question arises whether students truly understand the concept of the learning community and whether the group process scripts adequately change their comprehension. The results show that from the perspectives of learning communities' context, purpose, means, and challenges, students show an improved conceptual understanding of learning community pedagogy. Epistemological beliefs are a foundation for learners' engagement in the communities. In all, this study delivers first ideas and insights from a specific context on how to design small group process scripts to construct a productive learning community. The results showed that the scripts received satisfied feedback from participants. Meanwhile, there are some good recommendations for future studies: First, the small group process scripts in this study need to be more elaborate. It describes how to scaffold the activity flow and interaction (i.e. external collaboration scripts) but also needs to consider how internal group interactions occur. Second, we could look at the small group interactions and adjust the scripts to encourage all (or more) participants to be more active. Finally, in order to increase the value of the community knowledge, it needs better integration or dependency of small group process scripts with the community.

Reference

Acosta, A., Lui, M. & Slotta, J. D. (2014). Exploring group-level epistemic cognitions within a knowledge community and inquiry curriculum for secondary science. *Proceedings of the Eleventh International Conference of the Learning Sciences*, *2*, 673–680.

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education*, *126*, 334-345.

Aronson, E. (1978). *The Jigsaw Classroom*. Beverly Hills, CA: Sage.

Beichner, R. J. (2014). History and evolution of active learning spaces. *New Directions for Teaching and Learning*, *2014*(137), 9-16.

Bielaczyc, K., & Collins, A. (2009). Learning communities in classrooms: A reconceptualization of educational practice. *Instructional Design Theories and Models*, *2*, 269–291.

Cassidy R., Charles E. S. & Slotta J. D (2019) Editorial: Active Learning: Theoretical Perspectives, Empirical Studies, and Design Profiles. *Frontiers in ICT*, *6*, 3.

Cifuentes, L., & Murphy, K. L. (2000). Promoting multicultural understanding and positive self-concept through a distance learning community: Cultural connections. *Educational Technology Research and Development*, *48*(1), 69-83.

Dillenbourg, P., & Jermann, P. (2007). Designing Integrative Scripts. In F. Fischer, I. Kollar, H. Mandl & J. M. Haake (Eds.), *Scripting Computer-Supported Collaborative Learning* (275–301). Boston: Springer.

Dillenbourg, P., Nussbaum, M., Dimitriadis, Y., & Roschelle, J. (2013). Design for classroom orchestration. *Computer & Education*, *69*, 485-492.

Driver, M. (2002). Exploring student perceptions of group interaction and class satisfaction in the web-enhanced classroom. *The Internet and Higher Education*, *5*(1), 35-45.

Fagen, A. P., Crouch, C. H., & Mazur, E. (2002). Peer instruction: Results from a range of classrooms. *The Physics Teacher*, *40*(4), 206-209.

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.

Hung, D. W., & Chen, D. T. (2001). Situated cognition, Vygotskian thought and learning from the communities of practice perspective: Implications for the design of web-based e-learning. *Educational Media International*, *38*(1), 3-12.

Kollar, I., Fischer, F., & Slotta, J. D. (2005). Internal and external collaboration scripts in web-based science learning at schools. In Koschmann, T., Suthers, D., and Chan, T. -W. (eds.), *Computer Supported Collaborative Learning 2005: The Next 10 Years*, Lawrence Erlbaum, Mahwah, NJ, pp. 331-340.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts–a conceptual analysis. *Educational Psychology Review*, *18*(2), 159-185.

Kollar, I., Fischer, F., & Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learning & Instruction*, *17*(6), 708-721.

Kozlowski, S. W. (2018). Enhancing the effectiveness of work groups and teams: a reflection. *Perspectives on Psychological Science*, *13*(2), 205-212.

Li, Y., Dai, J., Wang, X. & Slotta, J. D. (2020). Active learning designs for Calculus II: a learning community approach for interconnected smart classrooms. *International Journal Smart Technology and Learning*, *2*(1): 66–87.

Madhok, J., Slotta, J. D. & Linn, M. C. (2010). Longitudinal impact of an eighth-grade inquiry curriculum on students' beliefs and achievement in science. *Annual Meeting of the American Educational Research Association*, April 30–May 4. Denver, Colorado.

Magolda, M. B. B. (1992). Students' epistemologies and academic experiences: Implications for pedagogy. *The Review of Higher Education*, *15*(3), 265-287.

Mazur, E. (1997). *Peer Instruction: A User's Manual*. Prentice Hall, Upper Saddle River, NJ.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, *49*(1), 345-375.

Peer, J., & Lourdusamy, A. (2005). Students' epistemological beliefs about science: The impact of school science experience. *Journal of Science and Mathematics Education in Southeast Asia*, *28*(2), 81-95.

Scardamalia, M. & Bereiter, C (1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences*, *3*(3), 265-283.

Slotta, J. D., Quintana, R. M., & Moher, T. (2018). Collective inquiry in communities of learners. *International Handbook of the Learning Sciences*, 308-317.

Slotta, J. D. & Peters, V. (2008). A blended model for knowledge communities: embedding scaffolded inquiry. In Paul, A. K., van Merriënboer, J. J. G. and de Jong, T. (eds): *Proceedings of the 8th International Conference on International Conference for the Learning Sciences*, International Society of the Learning Sciences, The Netherlands, Utrecht, *2*: 343–350.

Slotta, J. D., Tissenbaum, M., & Lui, M. (2013, April). Orchestrating of complex inquiry: three roles for learning analytics in a smart classroom infrastructure. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 270-274).

Vogel, F., Kollar, I., Ufer, S., Reiss, K., & Fischer, F. (2016). Fostering university freshmen's mathematical argumentation skills with collaboration scripts. In C. K. Looi, C. K., Polman, J. L., Cress, U., and Reimann, P. (Eds.), *Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences*, Singapore: International Society of the Learning Sciences (pp. 599-606).

Weinberger, A., Kollar, I., Dimitriadis, Y., Mäkitalo-Siegl, K. and Fischer, F. (2009). Computer-supported collaboration scripts: perspectives from educational psychology and computer science. In Balacheff, N., Ludvigsen, S., de Jong, T., Lazonder, A. and Barnes, S. (Eds.): *Technology-Enhanced Learning: Principles and Products*, Amsterdam University Press, Springer, New York.

Acknowledgments

This study is funded by the Elite Network of Bavaria (K-GS-2012-209). Special thanks to Xinli Wang at the University of Manitoba and Joel Weber at the University of Toronto, they have worked on this study as well.

Exploration of Facilitation Strategies for Intergroup Dialogues in a CSCL Context

Tugce Aldemir, Marcela Borge, Yu Xia
tfa5065@psu.edu, mbs15@psu.edu, yzx64@psu.edu
The Pennsylvania State University

Abstract: Our increasingly diverse society requires learners to develop cultural competencies in a way that they could form relationships across differences. Current educational efforts to support these competencies are either marginalized or demand long-term investments. Thus, as the first step of our long-term efforts to design pedagogical tools for modeling productive intergroup dialogues in CSCL, in this study, we explored the facilitation strategies groups implemented to coordinate the group's content sense-making around sensitive topics such as race, gender, sexual orientation, and oppression in a CSCL context. We collected and analyzed the discussion transcripts of the 13 groups across two sessions and identified four facilitation strategies and cognitive and emotional behaviors associated with these strategies.

Keywords: CSCL, intergroup dialogue, facilitation strategies

Introduction

Our increasingly diverse society requires learners to develop cultural competencies in a way that they could form relationships across differences. This could include critically analyzing and reflecting on one's own and others' ideas and approaches pertaining to social identities and power inequalities. Current educational efforts demand increased attention to diversity education (e.g., identifying similarities and differences between social identities) and social justice education (e.g., developing critical consciousness of social identities and power dynamics in play, and creating allyship across differences) (Bell, 2007; Gurin et al., 2013). However, developing these skills is a long-term endeavor, and thus, learners need to be provided with a psychologically safe context and guidance to engage in discussions to practice these skills repetitively (Watt, 2007). In Social Psychology, there is a growing literature aiming to develop and evaluate programs and courses for intergroup dialogues (IGD) –i.e., long-term, usually face-to-face dialogues facilitated by trained facilitators, where people from diverse social identity groups engage in dialogic and critical discourse around politically charged topics such as race, gender, privilege and social injustices (e.g., Frantell et al., 2019; Gurin et al., 2013; Nagda, 2006). However, sustained face-to-face discussions with trained facilitators, as modeled in the intergroup dialogue literature, may not always be available and accessible for all. Thus, we argue that a CSCL context with pedagogical tools modeling how to engage in productive dialogic and critical discourse with diverse groups could allow students to develop the knowledge and ability to optimize such dialogues and practice the skills needed for multicultural competence.

CSCL literature offers intervention tools and strategies that help students optimize their collaborative activities by modeling high-quality collaboration and supporting regulatory processes (e.g., Borge et al., 2018; Järvelä, & Hadwin, 2013; Järvelä et al., 2016). In previous work, we provided a theoretically informed technological intervention that provides a model of competence to support the development of socio-metacognitive expertise: the knowledge about and ability to regulate collaborative processes at the group level to improve collaborative processes over time (Borge et al., 2018). However, the scope of this previous study did not fully address the impacts of emotions on social interactions nor did it focus on multicultural skills. It is uncertain whether our existing approach and models would help students improve the quality of conversations that are more culturally personal and politically controversial. Thus, in this study, we explored the facilitation strategies diverse groups implemented to coordinate the group's content sense-making around politically charged topics such as race, gender, privileges in our CSCL context. By doing so, we aimed to (1) examine whether our existing system, models, and strategies can afford IGD without the presence of a trained facilitator; and (2) whether groups can exhibit effective facilitation strategies (Gurin et al., 2013) without directly instructed to do so. That will ultimately guide us in our larger efforts to identify how we could revise the current strategies and models to better address the needs of IGD in a CSCL context.

Theoretical framework

Intergroup dialogues (IGD) to build multicultural competence

Multicultural competence can be broadly defined as the abilities needed to communicate effectively across cultures (Mio et al., 2019). Being multiculturally competent would require one to develop (1) awareness of her/his assumptions, views, biases; (2) understanding and appreciation of cultural groups and differences; (3) skills to communicate across differences; and (4) critical consciousness to evaluate the role of social identities and structural dynamics in the daily lives (Watt, 2007). As Watt (2007) further posed, these skills can only be improved with repetitive practices over time through conversations with individuals from different social identities. One intervention is Intergroup Dialogue (IGD) (Gurin et al., 2013). IGD was defined as sustained and usually face-to-face dialogues guided by trained facilitators, where people from diverse social identity groups engage in dialogic and critical discourse around politically charged topics such as race, gender, privilege, and social injustice (Frantell et al., 2019; Gurin et al., 2013; Nagda, 2006).

Multiple benefits of IGD have been highlighted in the literature – e.g., identity and ally development, perspective-taking, attitude changes, critical consciousness, skill development, and action preparedness (see for details: Frantell et al., 2019). Nevertheless, these outcomes of the IGD mainly rely on facilitators' performance and the quality of the courses and the programs implementing IGD (Gurin et al., 2013). Designing and implementing such sustained training programs and courses and training facilitators require lengthy resources and efforts, and so, a productive IGD may not be accessible and available for all students. Thus, we argue that if students could develop the knowledge and the skills to engage in productive IGD, we might minimize the need for external support and guidance to produce these positive learning opportunities offered by IGD.

Socio-metacognitive competence in CSCL

In our previous studies, we created a regulation model that guides how students make sense of and regulate their collaborative discussions: Socio-metacognitive competence is the ability to collectively make sense of and regulate group's collaborative processes to improve the quality of their collaborative discussions (Borge et al., 2018; Borge & White, 2016). We also developed a theoretically informed technological intervention to help students develop their socio-metacognitive competence, along with a model of assessment where we listed concrete communication patterns associated with high to low-quality collaborative sense-making processes, i.e., Verbal Equity, Joint Idea Building, Developing Joint Understanding, Exploration of Alternative Perspectives, Quality of Claims and Norms of Evaluation (Borge et al., 2018; Borge & Shimoda, 2019; Borge & White, 2016). Our models and strategies outlined what high-quality collaborative discussion looks like and prompted groups to compare their processes to this optimal model, actively identify problems in their processes, and collectively identify or select strategies to address these problems (Borge et al., 2018; Winnie & Nesbit, 2009). The intervention succeeded in getting groups to improve their collaborative discussions over time (Borge et al., 2018). However, the focus of these discussions we evaluated the intervention for was limited to information science concepts. Thus, it is uncertain whether our existing approach and models would help students improve the quality of more personal and politically charged discussions.

Facilitation in intergroup dialogues (IGD)

Facilitation in IGD literature refers to external guidance provided during the dialogue to optimize content-learning and structured interactions (Gurin et al., 2013). Nagda (1999, cited in Gurin et al., 2013) developed and validated measurement to evaluate the effectiveness of a facilitator, where he listed facilitation strategies that could guide groups' cognitive, social, emotional, and motivational processes. The facilitation strategies were:

> creating an inclusive climate, modeling good communication skills, actively involving me in learning experiences, intervening when some group or class members dominated discussion, encouraging group or class members to talk to each other, not just to the facilitators/instructors, intervening when some group or class members were quiet, handling conflict situations, helping the clarify misunderstandings, offering their perspectives in a helpful way, bringing in a different perspective when everyone seemed to be agreeing, encouraging us to continue discussing when it became uncomfortable (Nagda, 1999, cited in Gurin et al., 2013, p.389).

However, since it was a self-report scale, it did not offer a consistent protocol showing how these strategies would look like in a dialogue. Thus, the questions of whether our existing system, models, and strategies can afford IGD and support effective facilitation strategies, without the presence of a trained facilitator, would require an explorative approach to identify the facilitation strategies groups used to coordinate each other's cognitive, social, emotional, and motivational processes related to the content. Given that, in this study, we explored the facilitation strategies of diverse groups while they collectively made sense of the sensitive topics such as race, gender, privilege in our CSCL platform. More specifically, our research question was: What facilitation strategies do inter-groups use to coordinate their content sense-making around sensitive topics (e.g., race, gender, privileges, oppression) in a CSCL context?

Course context and participants

The study was conducted in an undergraduate Multicultural Psychology course designed to introduce students to concepts of race culture, ethnicity, bias, cultural competence, oppression, and guided them to explore the meaning and value of these concepts as they pertain to various psychological issues. As part of the course, students were expected to engage in collaborative activities. Developing collaborative discussion skills needed to engage in intergroup dialogues around politically charged topics was one main goal of the course. The participants were 35 undergraduate Psychology students who enrolled in the course (25 females, 71.4%; 9 males, 25.7%, 1 non-binary, 2.9%).

Procedure

The students were assigned to 13 groups of three and two based on their self-identified racial-ethnic and gender identities to create diverse groups. The groups were asked to engage in a set of synchronous collaborative activities four times throughout the semester to collectively make sense of their course concepts. Each set included: (1) a pre-discussion activity: students read the weekly readings and wrote an individual reflection addressing 3-4 questions prompting critical reflection of the readings; (2) a synchronous discussion: students set a meeting time with their group members to synchronously discuss the questions, their individual reflections and readings (60 mins); (3) individual assessment of collaborative discussion: once completed their discussion, students were asked to individually assess the quality of their collaborative discussion and to provide justifications to support their scores, using a collaborative process rubric detailing concrete communication patterns associated with high and low quality collaborative discussion, goals of each criterion, problems that can be associated with each criterion and strategies to address those problems (15 mins); and (4) collective reflection and planning: after individual assessment, group members come together again to reflect on their individual assessments to collectively identify their weakness(es) and strength(s), and then to collectively identify or develop strategies to addresses those weaknesses in their future discussions (15 mins). The discussions and individual and collective reflections were held in a CSCL environment and saved automatically to the system. After each discussion, a trained coder scored the discussions using the same rubric and provided feedback to the groups.

Data collection and analysis

Discussion transcripts of the 13 groups across two sessions were collected and analyzed to identify the facilitation strategies group used as they collectively make sense of the sensitive topics (e.g., race, gender, privileges, etc.) in a CSCL context. The third and fourth discussions were excluded in this study because before the third session, four students from four different groups dropped the class, and we wanted to eliminate any confounding impacts these changes might have caused to the groups' dynamics.

We implemented a bottom-up approach to identify the themes and codes associated with facilitation strategies through multiple iterations of coding and connection to theory (Corbin & Strauss, 2014). Coding focused on groups' facilitation strategies during content sense-making and understanding. We were interested in how groups coordinated each other's cognitive and emotional processes related to the content. Therefore, we coded at both the single turn and episodic levels and excluded socio-metacognitive sense-making and regulation turns (e.g., reflecting on their collaborative processes) and non-task related talk in our analytical focus. Then, by constantly comparing the codes across groups, we sorted the codes into appropriate themes and sought out literature to make sense of emerging themes (Corbin & Strauss, 2014). All ethical guidelines were followed in collecting, analyzing, and reporting the study.

Results

Our thematic analysis of the discussions suggested groups engaged in four facilitation strategies to guide their dialogues: encouraging participation and elaboration, establishing common understanding, navigating through prompt questions, and sharing emotional support to other members.

Groups encouraged their members to actively participate in the dialogue and elaborate ideas and personal experiences, by prompting each other to share and elaborate on their opinions and identity experiences/self-descriptions or/and to consider alternative perspectives to the agreed-upon ideas. For example, in the following excerpt from Group 6 (see Table 1), one group negotiated their perspectives of race when asked to define the construct. The episode started with Cesar prompting other group members to share their opinions of a quote about race, and it got richer as he urged others to elaborate on their ideas and to consider alternative perspectives.

Table 1: Encouraging participation and elaboration

Turn	Speaker	Utterance	Explanation
1	Cesar	So, when experts argue that "*there is only one race and that's the human race*" what do you guys think of that?	Prompting others to share opinions
2	Ashley	I agree that scientifically yes there is only one human race because biologically we are more similar than different. Race was an invented construct. It does not actually matter, people only made it matter.	Sharing an opinion about the prompt
3	Cesar	Are we more similar than different? if there's different classifications of types of people based on ethnicity, could we truly see we are similar? // I think I'm very different than a white person. Not saying that in a bad way, but culturally and ethnically, I would say that although me and a white person are both humans, we are actually different based on numerous factors.	Challenging an idea by offering an alternative perspective & Supporting with personal experience
4	Ashley	yes, but those are learned factors	Sharing an alternative perspective
5	Cesar	hmm interesting, could you elaborate?	Prompting other to elaborate
6	Ashley	speaking strictly biologically all races are more similar than different. When you need an organ transplant or something it can come from any race // culture is learned.	Elaborating on the alternative perspective & Sharing an idea
7	Cesar	Is culture learned? // because one could make the argument that they were born within one type of culture. I was born and raised within Puerto Rican culture.	Challenging an idea by offering an alternative perspective & Supporting with personal experience
8	Emma	a difference in skin color is only due to different amounts of melanin in the body	Supporting Ashley's perspective

In this episode, the prompts were mostly directed to the whole group. In conjunction with such group-oriented encouragement of participation and elaboration, even though less frequent, groups also put efforts to invite a certain member – e.g., silent member, member of minority identity – to share opinion and identity experience/self-description. This pattern can be exemplified with the following episode (see Table 2) where a silent group member was prompted to share her identities.

Table 2: Inviting silent member to conversation

Turn*	Speaker	Utterance	Explanation
1	Amelia	I'm Amelia [anonymized], I'm from a small town in NC. I'm white and have lived in a mostly-white environment growing up. My parents were raised catholic in an even smaller town in Missouri, which was completely white...	Sharing self-identities
4	Elena	Did you feel like it was more of shock or did you ease into it?	Prompting other to elaborate on her identity experience
9	Eric	Im Eric [anonymized]. im Korean, born and raised in America in the suburbs outside Philly. I grew up in a Christian household, but the majority of my hometown was Jewish whites. my parents immigrated from Korea about 40 years ago...	Sharing self-identities

10	Amelia	Did you have a hard time growing up finding a balance between your Korean culture and American culture?	Prompting other to elaborate on his identity experience
11	Elena	Did you ever feel excluded where you grew because it was predominantly Jewish whites?	Prompting other to elaborate on his identity experience
28	Eric	Do you want to talk about your background Elena?	Inviting a silent member to share her identities/ identity experiences

Note: *Some of the turns were not included in the table. In the excluded turns, the group members continued their discussion around the same topic.

As observed, while two members describe their self-identities, Elena demonstrates her interest in their stories by prompting them to elaborate more. Yet, she does not share hers. Noticing that, Eric prompts her to share her background to invite her to the conversation around the identities.

Another facilitation strategy observed in groups' discussions was establishing common understanding. There were five types of cognitive behaviors exhibited to establish common understanding: summarizing and paraphrasing, asking for clarification, examples or confirmation, ensuring all terms discussed have the same meaning for all members, providing examples to support/explain others' opinions, and connecting, synthesizing and reflecting on shared ideas/experiences. The explanations of each behavior along with samples from the discussions are presented in Table 3.

Table 3: Cognitive behaviors associated with establishing common understanding strategy

Behavior	Description	Example
Summarizing and Paraphrasing	Summarizing or paraphrasing shared opinion(s) to make sure all team members have a common understanding about the topic, without adding a new idea.	**[Team 6]** Emma: "*...I think there can be another classification of race kind of based on ethnicity like Canadians as a race, or Australians as a race or Russians as a race// and then within that, it can be almost narrowed down more// like into skin color.*" [Original idea on an ongoing discussion] Cesar: "*so what I'm hearing is that both of you guy believe that ethnicity is strongly influenced in determining someone's race.*" [Summarizing the ideas to establish common understanding]
Asking for clarification, examples	Asking others to clarify or exemplify their opinions.	**[Team 2]** Zhang: "I *would say I believe in individualism which makes me respect the difference between people.*" [Original idea on an ongoing discussion] Riley: "*How did you realize that?*" [Asking for explanation] Zhang: "*Because I grew up in a collectivism culture and I hated that so much and after I came to the US I see so diversities in this country*." [Sharing identity experiences for explanation] Riley: "*Did it take coming to the US to understand that? // Can you give an example?*" [Asking for example for explanation]
Asking for confirmation to a shared idea	Asking for confirmation to a new idea built upon others' ideas. Includes interpretation of previously shared ideas.	**[Team 5]** Olivia: *"...People like to feel safe in their own views and opinions. when someone else comes along and has a different opinion or want things to be change their security is threatened."* [Original idea on an ongoing discussion]

		Jiao: "*so maybe the white people were afraid of the black people because they were used as slaves and did not look like white people?*" [Asking for confirmation to new idea built upon the original idea]
Ensuring all terms used have same meaning for all members	Prompting team members to share how they define the terms discussed or providing a definition of the terms discussed.	**[Team 6]** Cesar: "*so let's quickly define the difference between race, ethnicity and culture.*" [Prompting team members to share their definitions to establish a common understanding for the terms] **[Team 2]** Riley: "*Just a reminder, High Context communication: meaning embedded in the context of the situation or internalized societal rules*." [Sharing a definition for the term discussed/mentioned]
Providing examples to support/explain others' opinions	One member sharing an example to support/explain other(s)' idea (e.g. sharing one's own personal experience)	**[Team 5]** Isla: "*I think it's hard not to have the bias especially when you have grown up with a certain cultural idea and you think that certain idea is common sense.*" [Original idea] Olivia: "*One really basic example of what you are saying is the word football. If you just speak it, for some cultures that is the equivalent of soccer for us*." [referring to a shared culture to exemplify an original idea]
Connecting, synthesizing and reflecting on shared ideas/experiences	Connecting, synthesizing or reflecting on shared ideas/experiences to establish common understanding, with adding a new idea.	**[Team 5]** Olivia: "*I think it is ridiculous for a place to refuse service to people based on their beliefs and whatnot. Like they aren't coming to your business to force their beliefs on you, maybe they just want like a check-up at the doctor, or they just want to buy groceries. they aren't there to bother you.*" [Original idea on an ongoing discussion] Isla: "*Professionalism is important when it comes to these things because you do not want to offend someone especially when you are trying to work in a field like therapy. Others may not see you as genuine*." [Another original idea on the ongoing discussion] Jiao: "*so we could all agree that a professional should remain as unbiased as possible right?*" [Synthesizing the shared ideas – New idea]

Another common facilitation strategy was navigating through prompt questions. At least one member from each group shared the prompts questions to either initiate the discussion around the question or redirect the flow of the discussion to another prompt question. For example, after hearing a group member's opinion about individualism versus collectivism, Ehsan (Group 7) stated: "*This is a good segue into the last part of the reflection, do you have any thoughts on your own values and beliefs that you think might be cultural worldviews?*" to redirect the group's attention to the last reflective question.

The last facilitation strategy observed in groups' discussions was sharing emotional support to the other members. This strategy mainly included appreciation of others' perspectives, inquiries and identity experiences/self-descriptions, and expression of emotional empathy. For example, in an episode from Group 5, Isla– A Latin-American woman who grew up in a diverse city – and Olivia– a White American woman who grew up in a white-dominant town– discussed whether they find the university to be diverse enough, and despite their disagreement, they shared their appreciation of each other's diverse experiences and empathized with their emotions aroused from these experiences. Describing university as a diverse environment, Olivia stated: "*It's definitely way more diverse here which I think is great, but for some others coming from where I'm from, may be a culture shock…*" Isla compared her experiences with Olivia's, by expressing "*I experienced culture shock in a* *different way. It is the first time I realized I was a minority.*" In response, Olivia appreciated Isla's diverse experiences and empathized with her emotions aroused from these experiences: "*Yes I can imagine that being a big shock, especially coming from your school where you are used to being surrounded by people from your own culture.*"

The diverse experiences and opinions were mainly appreciated by the group members. In some episodes where the discussion became politicized, alternative perspectives and experiences created uncomfortable situations for member(s). In these cases, group members put effort to comfort and encourage the uncomfortable member(s) to continue the discussion. For example, while discussing the race construct, two group members [Group 6] started sharing their experiences with racism. Having realized that the white member was silent during this exchange, Ashley asked: "*Emma, how does it make you feel being in the only white person in this conversation. Uncomfortable at all?*" To this question, Emma responded: "*kind of. I don't know what to say// like I just can't relate and I don't want to say anything because I don't have anything to relate to*." Then, both Ashley and Cesar tried to comfort her to make sure she knows her opinions are appreciated, by saying "*it's always good to get different perspectives from everyone though!*" (Cesar) and "*I think you definitely contributed a lot to the conversation even though you felt you could not necessarily relate.*" (Ashley)

Discussion

We identified facilitation strategies diverse groups implemented to coordinate the group's content sense-making around sensitive topics such as race, gender, sexual orientation, privilege in a CSCL context. Our analysis yielded four facilitation strategies groups implement to coordinate and negotiate their content understanding: encouraging participation and elaboration, establishing common understanding, navigating through prompt questions, sharing emotional support to other members.

First and foremost, these findings imply that a CSCL context when offered with effective technological and pedagogical interventions can support the dialogic and critical discourse of IGD without the guidance of a trained facilitator. The fact that groups implemented some of the strategies identified as effective facilitator behaviors (e.g., inviting a silent member to contribute to the discussion) (Gurin et al., 2013) to guide their discussions suggest that a model of regulatory behaviors and tools to support these regulatory behaviors might give students (1) an accessible and affordable space to practice the skills needed for multicultural competence (Watt, 2007) and (2) autonomy to regulate their discussions and content sense-making without the need of an immediate authority figure.

Our findings also suggest that communication and psychological processes are interrelated and inform each other throughout the IGD (Nagda, 2006). As exemplified in the fourth theme, group members bring their repertoires of experiences, values, and assumptions to the discussions, and these repertoires shape how they create their social-self and interpret others' social-selves in this context (Ting-Toomey & Kurogi, 1998). Thus, when the group started to talk about race concept, and how it translates to their daily lives, Emma – a White American woman – did not want to contribute to the discussion as she considered her group members – both non-white–as more knowledgeable in that topic. In this situation, encouraging an off-task group member to re-engage might require more than using reminders or fostering a sense of shared responsibility (Rogat & Linnenbrink-Garcia, 2011) – it might also require socio-emotional regulation with emotional support and appreciation of alternative ideas and experiences.

The findings show that our model of socio-metacognitive competence is in need of revision to fully address the interplay between communication and psychological processes during IGD. Further research is needed to identify how the model can be updated to better conform to the needs of intergroup dialogues.

In this work, we limited our focus on identifying the facilitation strategies and associated behaviors as a means to understand whether our current intervention to support collaborative competence could afford intergroup dialogues (IGD) in a CSCL context without the presence of a trained facilitator and to identify what types of facilitation behaviors we want to support with our models. In our follow-up study, we will examine the dynamics between these strategies and the quality of collaborative processes to determine the range of quality for each facilitation strategy.

References

Borge, M., Ong, Y.S. & Rosé, C.P. (2018). Learning to monitor and regulate collective thinking processes. *International Journal of Computer Supported Collaborative Learning, 13,* 61–92. https://doi.org/10.1007/s11412-018-9270-5

Borge, M., & Shimoda, T. (2019). Designing a computer-supported-collective regulation system: A theoretically informed approach. *Technology, Instruction, Cognition, & Learning, 11*, 193-217.

Borge, M. & White, B. (2016). Toward the development of socio-metacognitive expertise: An approach to developing collaborative competence. *Cognition and Instruction, 34*(4), 323-360. https://doi.org/10.1080/07370008.2016.1215722

Bell, L. A. (2007). Theoretical foundations for social justice education. In M. Adams, L. A. Bell, & P. Griffin. (2nd Eds.), *Teaching for diversity and social justice* (pp. 3-26). Routledge.

Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Sage publications.

Frantell, K. A., Miles, J. R., & Ruwe, A. M. (2019). Intergroup dialogue: A review of recent empirical research and its implications for research and practice. *Small Group Research, 50*(5), 654-695. https://doi.org/10.1177/1046496419835923

Gurin, P., Nagda, B. R. A., & Zuniga, X. (2013). *Dialogue across difference: Practice, theory, and research on intergroup dialogue.* Russell Sage Foundation.

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist, 48*(1), 25-39. https://doi.org/10.1080/00461520.2012.748006

Järvelä, S., Kirschner, P. A., Hadwin, A., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual-and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning, 11*(3), 263-280. https://doi.org/10.1007/s11412-016-9238-2

Mio, J. S., Barker, L., Domenech Rodríguez, & M. M., Gonzalez, J. (2019). *Multicultural psychology: Understanding our diverse communities.* Oxford University Press.

Nagda, B. R. A. (2006). Breaking barriers, crossing borders, building bridges: Communication processes in intergroup dialogues. *Journal of Social Issues, 62*(3), 553-576. https://doi.org/10.1111/j.1540-4560.2006.00473.x

Rogat, T. K., & Linnenbrink-Garcia, L. (2011). Socially shared regulation in collaborative groups: An analysis of the interplay between quality of social regulation and group processes. *Cognition and Instruction, 29*(4), 375-415. https://doi.org/10.1080/07370008.2011.607930

Ting-Toomey, S., & Kurogi, A. (1998). Facework competence in intercultural conflict: An updated face-negotiation theory. *International Journal of Intercultural Relations, 22*(2), 187-225. https://doi.org/10.1016/S0147-1767(98)00004-2

Watt, S. K. (2007). Difficult dialogues, privilege and social justice: Uses of the privileged identity exploration (PIE) model in student affairs practice. *College Student Affairs Journal, 26*(2), 114–126.

International Society of the Learning Sciences

Short Papers

Explicit Sharing of Emotions Improves the Relationship of Groups with Lower Dispositions to Regulate Emotions in Collaborative Problem-solving

Sunny Avry, TECFA-FPSE and SIMS-CVML, University of Geneva & Center for Learning Sciences, Swiss Federal Institute of Technology in Lausanne, sunny.avry@epfl.ch
Gaëlle Molinari, TECFA-FPSE, University of Geneva & UniDistance,
gaelle.molinari@unige.ch

Abstract: The role of emotions in (computer-supported) collaborative learning has gained increasing attention. Self and others' emotions and their consequences are crucial in social interactions but may be overlooked during collaborative problem-solving, as group members may highly focus on the processing of task information. Therefore, some authors have suggested that collaboration can benefit from explicit mutual sharing of emotions. However, other authors also argue that the beneficial effects of emotion sharing depend on the individuals' dispositions to use emotions to change behaviors. This study explores how these two aspects influence how group members perceive their relationship during collaborative problem-solving. Results show that people with lower dispositions to regulate emotions improve their perception of equality, receptivity and productivity when they share emotions explicitly.

Introduction

Emotions are crucial information in social interactions because they shed light on how others interpret the current situation (Van Kleef & Fischer, 2016). In collaborative learning, emotions may provide learners with information about various dimensions, for example, socio-cognitive tension (cognitive dimension; Andriessen et al., 2011), collective efficacy beliefs (motivational dimension; Lent et al., 2006) or trust (relational dimension; Hale et al., 2005). Therefore, the awareness of emotions during interaction may induce mutual adjustments to maximize goal achievement. However, there is a loss of nonverbal cues in remote collaboration that may impair collaborative learning quality in reducing the possibility of using emotional information to build accurate mental models of the partner(s) (Avry, 2021; Molinari et al., 2009).

Explicit sharing of emotions

Individuals could compensate for the limitation of nonverbal cues by communicating more their emotions verbally (Walther et al., 2015). Building on this idea, some authors have studied how the explicit sharing of emotions impacts the afore-mentioned dimensions of collaboration. For example, Eligio et al. (2012) investigated whether teammates can improve their mutual understanding of each other's emotions by sharing emotional labels. Results showed a positive effect of emotion sharing on mutual understanding of emotions, group affect and group performance, especially when participants collaborate remotely. Avry and Molinari (2018) also reported a positive impact of sharing emotions on the number of exchanges dedicated to greetings or expressing courtesy (relational dimension). Overall, the literature suggests that explicit emotional sharing could influence collaborative behaviors and compensate for the lack of emotional cues, especially in emotionally deprived collaborative environments. However, this emotional information needs to be understood and used efficiently, which depends on both the capacity and disposition to do so.

Interpersonal emotion regulation

Emotional competencies refer to four hierarchically organized abilities: perceive and express emotion; use emotions to facilitate thoughts; understand emotions; regulate emotions (Mayer et al., 2000). Interpersonal emotion regulation (IER) refers to the attempt to initiate, maintain, modulate or change emotions in self and others (Zaki & Williams, 2013). Its goal is ultimately to change behavior to promote goal attainment. In collaborative problem-solving, IER may serve cognitive, motivational and relational motives. Some are related to increasing others' performance, and others to relational concerns (Niven, 2016). For example, IER could either encourage affiliation or induce social distancing (Fischer & Manstead, 2010). Depending on the context, regulating emotions properly is an essential aspect of successful collaborative problem-solving. Four types of emotion regulation are identified (Niven et al., 2011): *intrinsic affect-improving* (e.g., thinking about something nice), *intrinsic affect-worsening* (e.g., thinking about my shortcomings), *extrinsic affect-improving* (e.g., doing something nice with someone) and *extrinsic affect-worsening* (e.g., explaining to someone how they had hurt myself).

Main hypotheses

The findings described above lead to two main hypotheses. First, better interpersonal emotion regulation (IER) dispositions should promote better relational quality (*H1*) because group members with higher IER dispositions are more likely to use emotions to regulate self and other's behaviors to achieve task goals. Second, explicit emotion sharing should benefit more group members with lower IER dispositions in increasing their attention to emotional information and their tendency to act accordingly (*H2*).

Method

Participants and experimental design

One hundred twenty-four students (86 women and 38 men; M = 23.2 years, SD = 4.4 years, 2.6 years of postgraduate education on average) voluntarily participated to this study. Each pair received 40 CHF as inconvenience allowance. In the registration phase, participants completed the EROS (Emotion Regulation of Others and Self; Niven et al., 2011) questionnaire, assessing their dispositions to regulate their own- and the others' emotions. The four types of affect regulation strategies (intrinsic affect-improving/worsening, extrinsic affect-improving/worsening) described above were concerned by the questionnaire. The EROS items were translated into French following the forward-backward procedure (*Process of Translation and Adaptation of Instruments*, n.d.). Participants were automatically assigned with an unknown partner of the same gender and a similar EROS score. Pairs were then randomly assigned to one of the two conditions: emotion sharing *versus* no emotion sharing. No difference of age ($t(116.26) = -0.76$, p =.44) or level of education ($t(120.98) = 0.13$, p =.89) were found between both conditions.

Procedure

Participants introduced each other briefly and then were seated in front of a computer screen. They collaborated remotely: they could not see each other and communicate only in a written form. The collaborative problem-solving task lasted sixty minutes, interspersed with five breaks of about two minutes dedicated to the emotions assessment. At the end of the task, participants completed the post-task questionnaires and then received feedback on their answers, the optimal solutions, and the computed group performance. A collaborative problem-solving task inspired by a report stating the food and agriculture challenges in 2050 was designed for the experiment. Participants collaborated in a CSCL environment including a statistic table with numeric information, a chat, a notepad and an area to submit joint answers. Depending on the condition, the software was also displaying either emotions (*emotion sharing*) or life habits (*no emotion sharing*) to be assessed five times during the collaboration. In the variant on emotions, participants were asked to focus on the previous 10 minutes of collaboration and evaluate the intensity of their own- and their partner's emotions (from 0-*not at all* to 6-*very strongly*) through verbal labels (*frustration*, *interest*, *boredom*, *enjoyment*, *confusion*). Immediately after the submission of participants' answers, a graph was displayed to contrast their own estimation of their partner's emotions and the actual partner's emotions. At the end of the task, participants also completed a computerized version of the questionnaire developed by Hale et al. (2005) to assess the relational quality of social interactions. Participants assessed – using 7-point Likert scales from 0 (no agreement) to 6 (total agreement) – how they perceived their partner (1) as treating them as equal (*equality*), (2) as friendly (*affection*), (3) as attentive, accessible, open, and interested (*receptivity*), (4) as trying to create a sense of familiarity between them (*depth*), (5) as influential (*dominance*), (6) as involved (*implication*), and (7) as contributing equitably (*productivity*).

Results

Table 1 describes the global mean intensities of the explicitly shared emotions. The intensity of negative emotions was significantly lower compared to positive emotions.

Table 1: Mean (standard deviation) intensity of emotion (out of 6) for each emotion shared

	Frustration	Interest	Boredom	Enjoyment	Confusion
Mean (self)	1.88 (1.34)	4.53 (0.82)	0.93 (0.81)	3.89 (0.91)	1.90 (1.17)
Mean (other)	1.97 (1.21)	4.34 (0.80)	1.05 (0.82)	3.76 (0.81)	1.98 (1.07)

To test *H1* and *H2*, the different measures of relational quality were regressed against the different types of emotion regulation dispositions, taking into account the effect of the explicit sharing of emotions (Figure 1). There were no significant regression equations for *affection*, *depth*, *dominance* and *implication*. Two significant regression equation were found for *equality*. First, emotion sharing and extrinsic affect-worsening explained a significant proportion of variance of *equality* ($F(3, 120) = 2.94$, $p < .05$, $f^2 = 0.07$). The association between *equality* and *extrinsic affect-worsening* was positive in the *no emotion sharing* condition ($r(58) = 0.25$, $p = .05$) and negative in the *emotion sharing* condition ($r(62) = -0.27$, $p = .03$). Second, emotion sharing and intrinsic affect-worsening also explained a significant proportion of variance of *equality* ($F(3, 120) = 2.81$, $p < .05$, $f^2 = 0.07$). A significant negative relation was found between *equality* and *intrinsic affect-worsening* in the *emotion sharing* condition ($r(62) = -0.27$, $p = .02$). This relation was marginal in the *no emotion sharing* condition ($r(58) = 0.23$, $p = .07$). Concerning *receptivity*, a significant regression equation was found ($F(3, 120) = 2.60$, $p = .05$, $f^2 = 0.07$), indicating that both emotion sharing and extrinsic affect-worsening explained a significant proportion of variance of *receptivity*. There was a significant positive relation between *receptivity* and *extrinsic affect-worsening* in the *no emotion sharing* condition ($r(58) = .38$, $p = .00$), which was not found in the *emotion sharing* condition ($r(62) = -0.13$, $p = .28$). Concerning *productivity*, a significant regression equation was found ($F(3, 120) = 3.46$, $p < .05$, $f^2 = 0.08$), indicating that both emotion sharing and extrinsic affect-improving explained a significant proportion of *productivity*. There was a significant positive relation between *productivity* and *extrinsic affect-improving* in the *no emotion sharing* condition ($r(58) = 0.39$, $p < .01$), which was not found in the *emotion sharing* condition ($r(62) = 0.02$, $p = .83$).

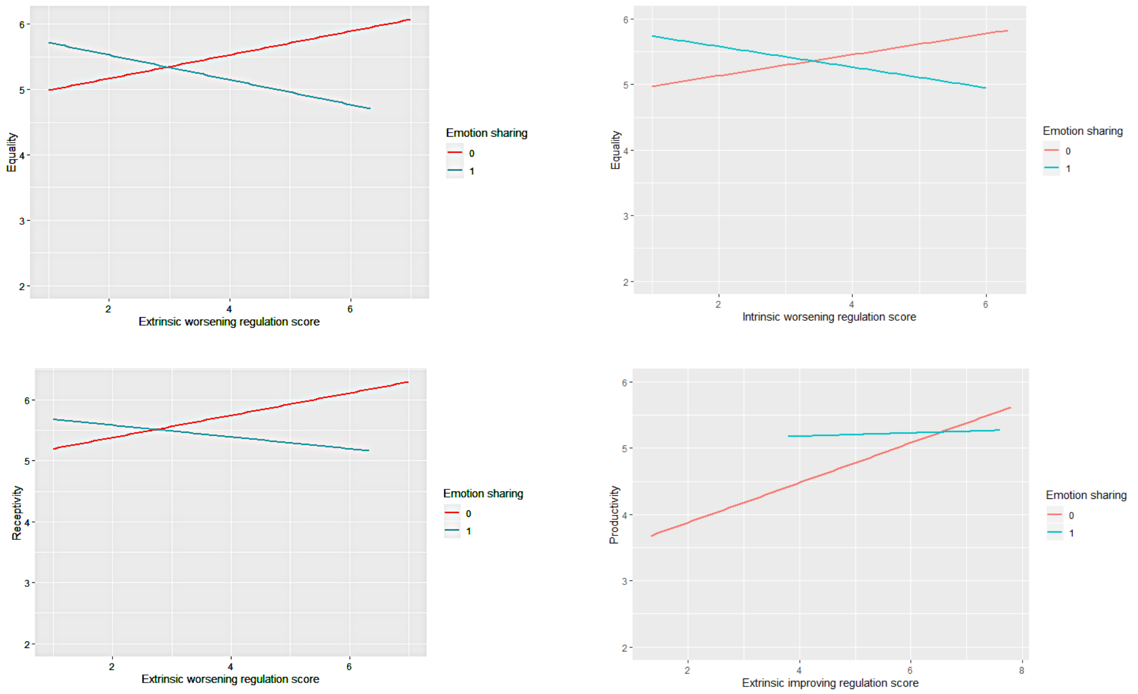


Figure 1. Effect of the interaction between emotion sharing and interpersonal emotion regulation dispositions on the perception of *equality, receptivity and productivity* during collaboration

Discussion

On the one hand, results show an effect of interpersonal emotion regulation dispositions on the relational quality of collaboration (*H1*), and this effect concerns only three relational dimensions, *equality*, *receptivity* and *productivity*. Participants with higher dispositions to worsen both their own – and their partner's emotions report being treated more equitably by their partner. The perception of the partner's *receptivity* also correlates with higher dispositions to worsen the other's emotions. Interestingly, this suggests that worsening emotions is a strategy that may increase a mutual feeling of equality (Yang & Kelly, 2016), and that may lead collaborators to perceive their partner as more attentive, open and interested. In contrast, the perceived *productivity* (equal work contribution) appears to be related to the tendency to improve the partner's emotions. On the other hand, however, the opportunity to explicitly share and compare each other's emotions (*emotion sharing* condition) cancels or reverses these results (*H2*). As a bias toward positivity when people explicitly share emotions in CSCL environments was shown in our previous studies (Avry, 2021; Avry et al., 2020), this may induce a differential effect on the perception of some relational dimensions in people with lower *versus* higher emotional dispositions to regulate emotions. Indeed, participants with lower dispositions may rely more strongly on the displayed emotional information (biased toward positivity) to judge the group's relational quality. However, in participants with higher dispositions, what is mutually shared may conflict with what they actually perceive, reducing their feeling of equality and receptivity. All in all, giving the opportunity to mutually share emotions during collaborative problem-solving seems to impact how people perceive the quality of their relationship. However, it seems to benefit mainly people with lower dispositions to regulate emotions.

Conclusion and further research

The results uncovered suggest that groups with lower dispositions to regulate emotions may benefit from sharing and comparing their emotions explicitly throughout the task to better regulate the collaboration, especially in emotionally deprived CSCL environments. However, although the explicit sharing of emotions could improve relational aspects of collaboration, it could also interfere with the processing of task information and provoke unwanted cognitive load. Therefore, there is still research to be carried out to find the best way to improve emotional awareness in CSCL environments without interfering with the task's resolution. We suggest that emotion awareness tools should optimally mimic how emotions are conveyed in face-to-face collaboration (non-verbal sharing of emotions through technologies) while making possible explicit sharing of emotions when needed (e.g., for regulating others' behaviors).

References

Andriessen, J., Baker, M., & van der Puil, C. (2011). Socio-cognitive tension in collaborative working relations. Learning across Sites: New Tools, Infrastructures and Practices, 222–242.

Avry, S., & Molinari, G. (2018). Sharing emotions impacts computer-supported collaborative processes: effect of an emotion awareness tool. Travaux Neuchâtelois de Linguistique, 68, 85–96.

Avry, S., Molinari, G., Bétrancourt, M., & Chanel, G. (2020). Sharing emotions contributes to regulating collaborative intentions in group problem-solving. Frontiers in Psychology, 11, 1160.

Avry, S. (2021). Beyond the dichotomy between the socio-cognitive and socio-emotional spaces : The pervasive role of emotions in collaborative problem-solving. Doctoral thesis, doi:10.13097/archive-ouverte/unige:149507 Retrieved from https://nbn-resolving.org/urn:nbn:ch:unige-1495076.

Eligio, U. X., Ainsworth, S. E., & Crook, C. K. (2012). Emotion understanding and performance during computer-supported collaboration. Computers in Human Behavior, 28(6), 2046–2054.

Fischer, A. H., & Manstead, A. S. R. (2010). Social functions of emotion. Handbook of Emotions, 3, 456–468.

Hale, J. L., Burgoon, J. K., & Householder, B. (2005). The relational communication scale. The sourcebook of nonverbal measures: Going beyond words, 127-139.

Lent, R. W., Schmidt, J., & Schmidt, L. (2006). Collective efficacy beliefs in student work teams: Relation to self-efficacy, cohesion, and performance. Journal of Vocational Behavior, 68(1), 73–84.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2000). Models of emotional intelligence. RJ Sternberg (Ed.).

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. In European Journal of Psychology of Education (Vol. 24, Issue 2, pp. 129–144).

Niven, K. (2016). Why do people engage in interpersonal emotion regulation at work? Organizational Psychology Review, 6(4), 305–323.

Niven, K., Totterdell, P., Stride, C. B., & Holman, D. (2011). Emotion Regulation of Others and Self (EROS): The development and validation of a new individual difference measure. Current Psychology, 30(1), 53–73.

van Kleef, G. A., & Fischer, A. H. (2016). Emotional collectives: How groups shape emotions and emotions shape groups. Cognition and Emotion, 30(1), 3–19.

Walther, J. B., van der Heide, B., Ramirez, A., Burgoon, J. K., & Peña, J. (2015). Interpersonal and hyperpersonal dimensions of computer-mediated communication. The Handbook of the Psychology of Communication Technology, 1, 22.

Yang, I., & Kelly, A. (2016). The positive outcomes of 'socially sharing negative emotions' in workteams: A conceptual exploration. European Management Journal, 34(2), 172–181.

Zaki, J., & Williams, W. C. (2013). Interpersonal emotion regulation. Emotion (Washington, D.C.), 13(5), 803–810.

 © ISLS

Click Restraint: Teaching Students to Analyze Search Results

Sarah McGrew, Alex C. Glass
mcgrew@umd.edu, aglass12@umd.edu
University of Maryland, College Park

Abstract: Young people often rely on search engines to find news or answer questions, yet research suggests that they need more support to learn to effectively navigate search results. We present findings from an intervention study in which one lesson focused on teaching students to evaluate search results and select websites to open. Findings suggest that, although teachers attempted to teach strategies for evaluating results modeled on fact checkers' approaches, students often fell back on less robust strategies: they used their familiarity with a website and a website's top-level domain to judge its trustworthiness. Our findings suggest avenues for further supporting young people to build effective strategies for navigating search results.

Introduction and framing

When young people need information or are looking for news, they are likely to turn to search engines. Research suggests that young people often click on the first or second result on a search engine results page (SERP) and express belief that those results are the most reliable (Gwizdka & Bilal, 2017; Hargittai et al., 2010). In reality, SERPs are the products of proprietary search algorithms that can be gamed by savvy website owners (Ledford, 2015) and may produce politically segregated search results (Flaxman et al., 2016). While there is still much to learn about the overall effect such algorithmic biases have on one's flow of online information (e.g., Dubois & Blank, 2018), teachers could play a critical role in helping students understand the consequences of navigating algorithmic spaces uncritically (Noble, 2018) and to navigate SERPs effectively to find quality information.

What does effective navigation look like? Rather than clicking on the first result, skilled information seekers spent time on the SERP and engaged in *click restraint*: They carefully read titles, URLs, and snippets and used their background knowledge to make decisions about which sites to click and which to avoid (Wineburg & McGrew, 2019). As they scanned search results, fact checkers asked themselves whether websites were familiar and if they met their information needs, deeming them promising or suspect depending on the task at hand. In one search scenario, fact checkers investigated who provided funding for plaintiffs in a court case by scanning most of the SERP and prioritizing articles from news sources. As one fact checker explained, "I'm coming up with a lot of different information. I'd rather click on some press reports" (Wineburg & McGrew, 2019, p. 28). A clear picture of their information needs, combined with knowledge of potential sources that might best meet those needs, informed fact checkers' choices about where to click.

Existing research suggests that educational interventions can help students learn to spend more time reading the SERP and analyzing the strengths and weaknesses of their searches (Bakke, 2020; Salmerón et al., 2020). However, more research is needed on how to support students to select sites to open from the SERP, as well as on what students learn from this instruction. Researchers have included navigating search results in efforts to teach students to find and evaluate online information but have not offered details about what these sessions focused on—or what, specifically, students learned (e.g, Ibieta et al., 2019; Terrazas-Arellanes et al., 2019). We began an investigation into the question of how we might teach click restraint as a strategy for navigating search results. We asked, In classes dedicated to analyzing search results, how do students and teachers reason about elements of the SERP? On what do they base decisions about where to click?

Method

Data were drawn from a larger study conducted by the Stanford History Education Group to investigate the efficacy of a curricular intervention to teach high school students to evaluate online information (see Wineburg et al., 2019). The study took place in a large school district in the Midwestern United States. As part of this study, six 12th-grade government teachers at three district high schools taught six lessons in *civic online reasoning* over the course of 10 weeks. *Civic online reasoning* is conceptualized as the ability to effectively search for and evaluate online information and centers on asking three questions of digital content: Who is behind this information? What is the evidence for the claim? What do other sources say? The lessons were developed and piloted by the Stanford History Education Group. Evidence from pilot studies and from pre- and posttest data collected in this intervention suggested that the lessons were effective in improving students' civic online reasoning (McGrew, 2020; Wineburg et al., 2019). Before participating teachers taught the lessons, they participated in a day-long professional development workshop led by two members of the research team. Teachers were introduced to civic online reasoning, participated in modeled portions of the lessons, and collaborated with colleagues and the research team to review and modify the lessons.

Teaching click restraint

The final lesson in the six-lesson sequence was devoted to teaching click restraint. This lesson began with teachers asking students where they usually click first within search results. Teachers then delivered a short (5-8 minute) lecture on the factors that influence the order in which search algorithms display results. Although search engines claim to factor both relevance and quality into the ordering of websites in the SERP, teachers emphasized that students bore the responsibility for evaluating how well sites actually met these standards—and which sites to ultimately click on. Next, classes practiced analyzing a sample SERP together. Teachers led students in evaluating the results one at a time, discussing the details they could glean from each result in relation to their information needs. Finally, students practiced analyzing another SERP on their own. They worked from a custom search engine that returned the same results to each student, aided by a graphic organizer that prompted them to examine aspects of the SERP before making a decision about where to click.

Participants and data

This study focused on the classes of six teachers who taught the six civic online reasoning lessons. Participating teachers video recorded the civic online reasoning lessons in one of their class periods. Pursuant to the district's guidelines for videotaping classes, teachers positioned the video camera so that it recorded them teaching but did not capture students' faces. Video and accompanying audio recordings thus captured students' and teachers' comments during full-class instruction portions of each lesson but not during group or individual work time. Teachers also shared student work from the class period they recorded if the student had assented, and parents or guardians had consented, to participate in the study. Across the six teachers' classes, there were 54 completed graphic organizers from the click restraint lesson available for analysis.

Analysis

Portions of the click restraint lesson that included full class instruction and discussion were transcribed. The authors read these transcripts multiple times to identify themes in the ways students and teachers reasoned about elements of the SERP. Because instructional practices that support students to evaluate search results have not previously been theorized in depth, we used inductive coding and constant-comparative analysis (Strauss & Corbin, 2015). We constructed analytic memos to track themes and key episodes that emerged regarding how teachers and students reasoned about the SERP and which sites to open or avoid. We then read and analyzed students' graphic organizers, adding to and amending the memos based on evidence of student thinking available in their responses. Based on this exploratory analytic work, we identified two themes that surfaced across the lesson videos and graphic organizers, which we detail in the findings.

Findings

Two themes emerged from our analyses of dialogue and student work from the click restraint lesson. First, teachers regularly advised students to avoid clicking on sites they did not recognize. All six teachers differentiated between sites they "knew" and sites they didn't as they discussed search results. For example, a teacher had the following exchange with students as they examined a SERP for a search on Internet filtering policies in schools:

Teacher:	So the top result is "Why schools block websites and you should too" from Net Nanny.
	Okay. Does anybody know Net Nanny is a website? Are you familiar with this one?
Student 1:	No.
Teacher:	Never heard of it. Me, I've never heard of it. Ooh, but number two is ALA.org. The American Association of School Libraries. Might they be a quality source?
Student 2:	Sure.
Teacher:	Okay. Which one would perhaps be least reliable for information?

In this exchange, the teacher moved beyond the result displayed first, Net Nanny, after confirming that he had "never heard of it," implying that being unfamiliar with a website justified passing over it. His reaction to the American Association of School Libraries website, which the teacher did seem to recognize, was more positive: After seeking confirmation from a student that it might be a "quality source," the teacher shifted the line of questioning from deciding which sites to open to finding the least reliable site. Over the course of the lesson, teachers expressed recognition of—and comfort with—sites including the *New York Times* and *Business Insider*. Meanwhile, teachers said they were not familiar with and therefore would not click on sites like Final Edition, Empire News, and Net Nanny. In these cases, the strategy of avoiding the unfamiliar functioned efficiently: the *New York Times* and *Business Insider* are likely better sources of information than parody or satirical websites (Final Edition and Empire News) or a company that sells internet filtering software (Net Nanny). The reasoning that teachers modeled in these instances mirrored fact checkers' approaches to search results and was, in fact, what the lesson plan advised.

Evidence from the lesson calls into question whether this reasoning was equally effective for students. Students likely have more limited and less differentiated knowledge of "known" websites than their teachers. Thus, they may not be able to distinguish between sites like Net Nanny and the American Association of School Librarians based on background knowledge alone. Further, some students selected or rejected websites based on their familiarity with the sites. When students worked independently to evaluate sources from a SERP on school uniform policies, aided by the graphic organizer, they offered explanations for why sites may or may not be trustworthy that relied heavily on their familiarity with the site (see Table 1).

Table 1: Examples of students' reasons for trusting (or not trusting) websites based on familiarity.

	Reasons Provided to Distrust a Website	Reasons Provided to Trust a Website
Examples of Student Responses in Graphic Organizer	"I know nothing about LiveAbout though so maybe not the most trustworthy." "I don't think this is a good source because I haven't heard of the source before." (Frenchtoast.com) "From a not known source." (Vogue)	"I think this source is probably going to be one of the best sources. It is one that many people know of." (USA Today) "It is from a well-known source (Vogue)." "Because they are from sources I know." (USA Today)

Students embraced sites they were familiar with and rejected those they did not recognize. "I haven't heard of the source" became a reason to distrust—not just temporarily pass over—a website, while "sources I know" were trusted. Students appeared to skip the stage of using what they knew about a site to decide if it met their information needs. Instead, student responses suggested that they equated familiarity and trustworthiness.

The second theme to emerge was that students showed evidence of judging sites on the SERP based on their top-level domains. When teachers asked what details available on the results page they might use to decide where to click, students in each class mentioned the top-level domain. Teachers often problematized the assumptions that were implied in these exchanges; for example:

Student 1: The dot com.

Teacher: But did we say dot coms or dot orgs really matter anymore on reliability?

A few talk turns later, after the teacher asked what kind of site students hoped to find in the search results, another student offered:

Student 2: Like a dot org.

Teacher: Okay, could be. But again, does a dot org always—is it infallible? Could it actually be untrustworthy? In this day and age? At one time dot coms were "Ooh, they're bad; dot orgs are good." But in this time period of now, present day, does it really matter?

Evidence of reasoning about the top-level domain also appeared frequently in students' graphic organizers, generated while students worked independently or in groups. For example, students wrote "The site is only a .com and the title is 'French Toast' which has no credibility tied to it" and "Yes because the first one is a .org." In these cases, students substituted one weak heuristic—assuming the top search result is the most reliable—for another, drawing conclusions based on a top-level domain.

Discussion and conclusion

Teachers advised students to practice *click restraint* in order to avoid the problematic heuristic of quickly clicking on the top result. However, evidence of students' reasoning in the lesson suggests that students substituted other, possibly equally problematic heuristics: They embraced sites they recognized, avoided sites they didn't, and used the top-level domain to make choices about where to click.

This study suggests that students need more practice slowing down on SERPs to consider whether websites, even those with which they are familiar, meet their information needs. To do this without falling back on fallible heuristics like judging websites by their top-level domains, students likely need to build their knowledge of websites and the information they are likely to offer. Instead of advising click restraint from the beginning, teachers might encourage students to click on more sites, even those they do not recognize, to learn more about the sites and the information they provide. Students might curate a catalogue of websites they investigate through such clicking to collaboratively build knowledge of websites. Further, teachers might encourage students to work together so they can discuss SERPs and benefit from each other's knowledge.

These findings lay the groundwork for continued research on how to support students to develop click restraint. Evidence from this lesson suggests that students need continued practice analyzing search results; teachers may thus need help planning for ways to integrate lessons in which students practice click restraint across the curricula. In particular, future studies should probe how teachers can support students to build the knowledge about sources that helped fact checkers effectively use click restraint (Wineburg & McGrew, 2019). Such studies might examine the efficacy of the strategies suggested above: supporting collaborative inquiry so students share knowledge as they examine search results and encouraging students to click and learn about websites in SERPs. Search engines and SERPs are a frequent presence in students' lives, and this study contributes to a growing body of literature on how schools and teachers might better support students to navigate and evaluate these resources.

References

Bakke, A. (2020). Everyday Googling: Results of an observational study and applications for teaching algorithmic literacy. *Computers and Composition*, *57*. https://doi.org/10.1016/j.compcom.2020.102577

Dubois, E., & Blank, G. (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society*, *21*(5), 729-745. https://doi.org/10.1080/1369118X.2018.1428656

Flaxman, S. R., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, *80*, 298–320. https://doi.org/10.1093/poq/nfw006

Gwizdka, J., & Bilal, D. (2017). Analysis of children's queries and click behavior on ranked results and their thought processes on Google search. In *Proceedings of the Conference on Human Information Interaction and Retrieval* (pp. 377-380). ACM. https://doi.org/10.1145/3020165.3022157

Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. Y. (2010). Trust online: Young adults' evaluation of web content. International Journal of Communication, *4*, 468-494. https://doi.org/1932–8036/20100468

Ibieta, A., Hinostroza, J. E., & Labbé, C. (2019). Improving students' Information Problem-Solving skills on the Web through explicit instruction and the use of customized search software. *Journal of Research on Technology in Education*, *51*(3), 217–238. https://doi.org/10.1080/15391523.2019.1576559

Ledford, J. L. (2015). *Search engine optimization bible (2nd ed.).* Sage.

McGrew, S. (2020). Learning to evaluate: An intervention in civic online reasoning. *Computers & Education*, 145. https://doi.org/10.1016/j.compedu.2019.103711

Noble, S. (2018). *Algorithms of oppression*. New York University Press.

Salmerón, L., Delgado, P., & Mason, L. (2020). Using eye-movement modelling examples to improve critical reading of multiple webpages on a conflicting topic. *Journal of Computer Assisted Learning*, *36*(6), 1038–1051. https://doi.org/10.1111/jcal.12458

Strauss, J., & Corbin, J. (2015). *Basics of qualitative research techniques and procedures for developing grounded theory (4th ed.).* Sage.

Terrazas-Arellanes, F. E., Strycker, L. A., & Walden, E. D. (2019). Web-based professional development model to enhance teaching of strategies for online academic research in middle school. *Journal of Research on Technology in Education*, *51*(2), 118–134. https://doi.org/10.1080/15391523.2018.1564637

Wineburg, S., Breakstone, J., Smith, M., McGrew, S., & Ortega, T. (2019). *Civic Online Reasoning: Curriculum Evaluation* (Working Paper 2019-A2, Stanford History Education Group, Stanford University).

Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teachers College Record*, *121*(11).

Acknowledgements

This work was supported by Google.org (Grant no. 1712-49805, Sam Wineburg, Principal Investigator). Special thanks to members of the Stanford History Education Group and to participating students and teachers.

Investigating Gaze Behavior of Dyads in a Collaborative Explanation Task Using a Concept Map: Influence of Facilitation Prompts on Perspective Taking

Yugo Hayashi, Ritsumeikan University, Japan, y-hayashi@acm.org
Shigen Shimojo, Ritsumeikan University, Japan, cp0013kr@ed.ritsumei.ac.jp

Abstract: In this study, we investigated dyads gaze behavior from an educational dataset on collaborative learning which used a learning support system which provided facilitation prompts. By using two eye-trackers, this study focused on (1) how the prompts influenced a reconsideration of prior knowledge and (2) the degree of perspective-taking. To investigate these points, this study considered two learning dyads working on a jigsaw-like task, wherein leaners explained a psychological theory by generating a concept map. The learners worked by showing each other's concept map prior to the task. Two eye-trackers were used to analyze how dyads looked at each other's concept map. The results show that learners receiving facilitation prompts exhibited the gaze behavior of reconsidering each other's concept map while making explanations. Additionally, learners who were using prompts when giving argumentative and conflicting suggestions exhibited a higher tendency to take other's different perspectives.

Introduction

Research on collaborative learning in cognitive science has shown that self-explanation activities, in which learners explain learning material to themselves to gain an understanding of that material, are effective in triggering metacognition (Chi, Leeuw, Chiu, & Lavancher, 1994). However, researchers have also recognized that learners gain a deeper understanding of the concepts in the learning content and generate abstract concepts by engaging in explanatory activities (Greeno, de Sande, 2007). Collaboration facilitates reconsideration based on differing knowledge and perspectives shared by collaborating individuals; in other words, collaboration leads to thinking in a manner that involves reinterpretation of objects from an abstract perspective (Roschelle, 1992). In a collaborative problem-solving study comprising experimenters (Hayashi, 2018), factors such as cognitive bias negatively affect consensus formation. Previous studies have also shown that these types of cognitive processing can be analyzed through eye-tracking technologies and that they can be useful as an index for analyzing pedagogical instructions and the effects of facilitation prompts (Hayashi, 2020).

In the field of psychology, gaze behaviors are studied as a measure for understanding a learner's mental process (Schober, 1993). Research in learning science (LS) and computer supported collaborative learning (CSCL) indicates that eye-trackers are effective for elucidating the nature of collaborative activities (Schneider, & Pea, 2014). Hayashi, (2016) investigated effective design of conversational agents through modalities such as using gaze gestures together with verbal facilitation prompts. Two eye-trackers were used to detect the learning process of learners who used different types of knowledge and perspectives displayed on a computer screen. These studies show how gaze behaviors can be a useful index to understand coordinative behavior during collaborative learning in learning systems with shared visual spaces, during which each learner presents the other learner's perspective. Based on this aspect, the present study investigates gaze behavior related to perspective-taking in a learning environment wherein learners have an opportunity to observe each other's differing viewpoints through a representation in a shared interface. The findings contribute on developing intelligent tutoring systems(Aleven, & Koedinger, 2002) that can detect the cognitive behaviors of the learners in collaborative learning tasks. Using gaze behavior as a tool to evaluate facilitation prompts is potentially useful for capturing the influence on cognitive behaviors such as perspective taking, and therefore this needs to be further investigated in this area.

Current study: Goal and hypotheses

This study investigates how facilitation prompts presented during learner-learner collaborative learning influence gaze behavior related to perspective-taking. The first goal of this study is to understand how the use of facilitation prompts influences gaze behavior such as a reconsideration of self/other knowledge. In our task, we predict that the presentation of prompts promotes reflective thoughts, thereby facilitating gaze behavior focusing on existing knowledge, that is, the concept maps (H1). The second goal in this study is to investigate how prompts can facilitate perspective-taking behavior as reflected through gaze behavior. This study will investigate the effects of two types of prompts that were designed in our preliminary study as well as how they influence perspective-taking gaze behavior. We predict that our designed prompts, such as prompts for facilitating successful coordination and conflicting argumentation, will influence learner's perspective-taking gaze behavior (H2).

Method

Participants

For this study, we utilized data from a previous study we conducted (Shimojo & Hayashi, 2020). Ninety-four university students majoring psychology participated in a laboratory-based experiment in exchange for course credit. Seven were excluded from the sample due to loss of data, resulting in a total of 87 participants. Hereafter, we refer to these participants as learners. Learners were randomly assigned to three types of conditions: the control condition, the grounding-prompt condition, and the conflict-prompt condition.

Experimental task and procedure

The present study focuses on dyads involved in an explanation-based activity (Weinberger & Fischer, 2006); this experiment was conducted in a remote environment that allowed participants to communicate with each other using a concept-map tool(Engelmann, & Hesse, 2011). The experiment included the following three phases: an individual learning phase, a collaborative learning phase, and a phase wherein they gave a shared representation of each other's different perspectives of the learning material. After entering the laboratory, one member introduced the other and they were given instructions regarding provision of informed consent. This study was reviewed for adherence to ethical standards by the authors' university. The learners' goal in this experiment was to theoretically explain a particular case event using psychological theories. Specifically, they were required to apply the attribution theory to a problem case story of a student who participated in a school counseling program and describe why the student has anxiety about the new academic year. The learners used the concept-map tool to reason through the issues of the student in the problem case story and explain the student's mental process using the attribution theory. Two monitors were connected to the PCs; two video-recording devices (Sony, HDR-CX680) were set up, and the Cmap software (https://cmap.ihmc.us/) was installed on the PCs for developing and synchronizing concept maps.

Experimental task and procedure

This set-up allowed for the simultaneous production and sharing of concept maps, thereby enabling each learner to see the other's concept and enabling the learners to develop a new concept map together. In the individual learning phase, the learners were required to read a text describing the case story of the student and to subsequently read another description of the attribution theory. In the collaborative learning phase, the learners worked together by giving oral explanations; they were instructed to explain each other's thoughts and develop another concept map(15 minutes). As described in Figure 1, the participants were able to see each other's concept maps (depicted in the right-hand side with dotted lines) they had developed in the individual learning phase while working on the common concept map (left-hand side). The learners gave explanations and developed a new concept using the mouse and keyboards. To effectively generate the common concept map, including integrating each other's perspectives, it was necessary to briefly refer to each other's concept maps.

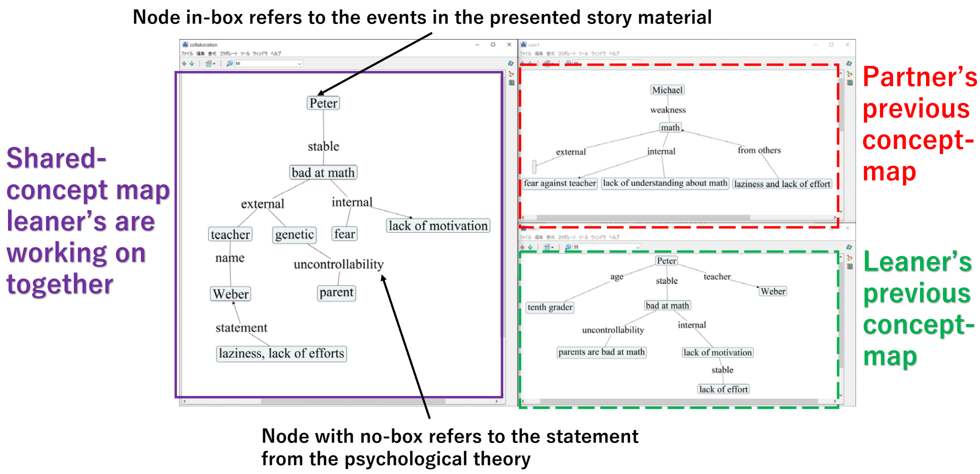


Figure 1. Example screen of the participant.

System

The two types of prompts used in this study were developed from the authors' previous study. The prompts were designed based on previous studies that investigated coordination during collaborative learning activities (Meier, Spada, & Rummel, 2007) and argumentation (Asterhan, & Schwarz, 2009). For the grounding-prompt condition, eight types of prompts were developed to facilitate successful coordination; they were grounded in the definition of the collaboration process provided by Meier, et.al, 2007). This definition includes the following concepts: communication (sustaining mutual understanding and dialogue management), joint information processing (information pooling and reaching consensus), coordination (task division, time management, and technical coordination), interpersonal relationship (reciprocal interaction), and motivation (individual task orientation). For the conflict-prompt condition, 13 types of prompts were designed to facilitate conflicting discussions and developed in line with the study of Asterhan, & Schwarz, (2009). These prompts included important aspect of discussions such as challenges, oppositions, rebuttals, elaborations, and requests for information. These prompts were presented 16 times by the experimenter through the local server connected with the two learners' PCs.

Measures

This study focused on how the prompts influenced participants' observations of their partner's concept map and the impact on the degree of perspective-taking. This study used two eye-trackers (Tobii X2-30) to analyze where the learners were looking during the task. Figure 1 shows that the learners were expected to view their partner's concept map on the right-hand side and their own map on the lower right-hand side. The screen was divided into three parts and the number of fixations per area was counted (area 1: other; area 2: self; area 3: shared area). We investigated the gaze degree of learners on areas 1 and 2 while they generated the common concept map in area 3. Observation of these areas indicated how learners were using self/other information and taking different perspectives while engaged in collaborative interaction. Moreover, we captured the degree to which one partner looked at another partner's screen; this resembles the behavior of considering (taking) the other person's perspective. This was evaluated based on how often the learners paid attention to their own and their partner's concept maps on their screens, leading to the following ratio calculation:

$$(1) \qquad b = \frac{n_1 - n_2}{n_1 + n_2}$$

where n_1 is the number of fixations on area 2 (the other learner's concept), and n_2 is the number of fixations on area 1 (the participant's own concept). If b is higher than 0, this suggests that learners were primarily looking only at their own concept map while working on the common concept map. However, if b is lower than 0, this suggests that learners were primarily looking only at their partner's concept map and indicates the learner's consideration of the other's concept map, which can be interpreted as perspective-taking.

Results

Gaze Fixation

To investigate if the prompts were functioning properly for activities such as referring to each other's concept maps while generating the new concept map, we investigated the ratio of gaze fixation on each of the three areas. A 3 (condition: control vs. grounding-prompt vs. conflict-prompt) X 3 (area: self vs. other vs. shared) analysis of variance (ANOVA) was conducted on the gaze fixation ratio. There was significant interaction between the two factors (F(4, 168) = 0.244, p = .001, η_p^2 = .264). The results of the simple main effect show that there were differences among all conditions in each area- self, other, and shared (F(2, 252) = 0.111, p = 0.01, η_p^2 = .911; F(2, 252) = 0.220, p = 0.01, η_p^2 = .953; and F(2, 252) = 0.644, p = 0.01, η_p^2 = .983). Next, multiple comparisons were conducted for each area using Ryan's method. The results show that grounding- and conflict-prompts led to more fixations than did the control condition in the self-area(p = 0.01 and p = 0.01, respectively) and the other area (p = 0.01 and p = 0.01, respectively). However, there were fewer fixations for grounding- and conflict-prompts than that for the control in the shared area (p= 0.01 and p = 0.01, respectively).

Perspective-taking

A one-way ANOVA was conducted to assess the differences among the three conditions on the gaze plot index b. However, no differences were found (F(2, 84) = 0.00, p = .94, η_p^2 = .001). The standard deviations for each condition- control, grounding-prompt, and conflict-prompt- were quite high (0.3, 0.4, and 0.5, respectively); therefore, we divided the learners into two groups, high b(b>0) and low b(b<0). A 2 (b: high vs. low) X 3 (condition: control vs. grounding-prompt vs. conflict-prompt) ANOVA was conducted on the gaze plot index b. There was a significant interaction between the two factors (F(2, 81) = 4.481, p = .001, η_p^2 = .547). Simple main effects were found only on the low b (F(2, 81) = 3.166, p = .001, η_p^2 = .985). Next, multiple comparisons were conducted for each condition on the low *b* using Ryan's method. The results show that the conflict-prompt had lower *b* than that of the control ($p = 0.01$). These results indicate that the group of learners who were paying more attention to other learners' perspectives ($b<0$)) were more likely to pay attention to their partner's perspective when they received conflicting prompts.

Discussion and Conclusions

This study investigated learners' gaze behavior on (1) the frequency of fixation on concept maps and (2) the degree of fixation on self/others' differing perspectives. The first goal of this study was to understand how the use of facilitation prompts influences gaze behavior such as reconsideration of self/others' knowledge. The results of our analysis on the frequency of fixation on concept maps show that either the use or non-use of prompts leads to an increase in the frequency of fixation. This supported hypothesis H1 and showed that when learners received facilitation prompts, they referred to self/other knowledge visible on the screen. Based on this aspect, we found that the prompts facilitated behaviors related to reconsideration, which are activities related to metacognition. Our second goal in this study was to investigate what kinds of prompts facilitate perspective-taking gaze behavior. The results of our analysis supported hypothesis H2, as the group of learners was paying more attention to others' perspectives than to their own. Further, they were more likely to pay attention to their partner's perspective when they received conflicting prompts. This study provided implications regarding how analysis of gaze behavior can be used for analyzing cognitive behaviors such as perspective-taking. Further analysis combining multimodal communication channels, which is left to future study, may provide more details.

References

Aleven, V., & Koedinger, K. (2002). An effective metacognitive strategy: learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, 26(2), 147-179.

Asterhan, C., & Schwarz, B. (2009). Argumentation and Explanation in Conceptual Change: Indications From Protocol Analyses of Peer-to- Peer Dialog. *Cognitive Science*, 33(3), 374-400.

Chi, M., Leeuw, N., Chiu, M., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18(3), 439-477.

Engelmann, T., & Hesse, F. (2011). Fostering sharing of unshared knowledge by having access to the collaborators' meta-knowledge structures. *Computers in Human Behavior*, 27(6), 2078-2087.

Greeno, G., de Sande, C. (2007). Perspectival Understanding of Conceptions and Conceptual Growth in Interaction. *Educational Psychologist*, 42(1), 9-23.

Hayashi, Y. (2016). Coordinating Knowledge Integration with Pedagogical Agents: Effects of Agent Gaze Gestures and Dyad Synchronization. *In Proceeding of the 13th International Conference on Intelligent Tutoring Systems(ITS2016)*. 254-259.

Hayashi, Y. (2018). The Power of a "Maverick" in Collaborative Problem Solving: An Experimental Investigation of Individual Perspective- Taking Within a Group. *Cognitive Science*, 42, 69-104.

Hayashi, Y. (2020). Gaze awareness and metacognitive suggestions by a pedagogical conversational agent: An experimental investigation on interventions to support collaborative learning process and performance. *International Journal of Computer-Supported Collaborative Learning*, 15, 469-498.

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.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences*, 2(3), 235-276.

Schneider, B., & Pea, R. (2014). Toward Collaboration Sensing. *International Journal of Computer-Supported Collaborative learning*, 4(9), 5-17.

Schober, F. (1993). Spatial perspective-taking in conversation. *Cognition*, 47(1), 1-24.

Shimojo, S., Hayashi, Y. (2020). Prompting Learner-Learner Collaborative Learning for Deeper Interaction: Conversational Analysis Based on the ICAP Framework. *Proceedings of the 28th International Conference on Computers in Education(ICCE2020)*, 177-182.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, 46(1), 71-95.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP20H04299.

International Society of the Learning Sciences

Examining how Three Network Visualizations Influence Student Engagement in Online Discussions

Fan Ouyang, Si Chen, Xu Li, Pengcheng Jiao, Dan Sun
fanouyang@zju.edu.cn; sichen327@zju.edu.cn; lixu2019@zju.edu.cn; pjiao@zju.edu.cn; dansun@zju.edu.cn
Zhejiang University

Abstract: Network visualization tools are designed to demonstrate varied aspects of online discussions in order to foster student engagement. However, empirical studies have indicated controversial results of the effect of tools on student engagement. This research designs network visualizations to demonstrate student engagement from varied perspectives, and further uses the mixed method to empirically investigate the effects of network visualizations on students' cognitive structures. Empirical research results indicate the social network visualization has the best effect on *leader* students' cognitive engagement; topic network visualization lessens the cognitive differences between *leader* and *regular* students; although *peripheral* students consistently have a low level of cognitive engagement, the cognitive network visualization improves their cognitive engagement. Based on the results, this research proposes integrated implications by considering learning theory, pedagogical support and tool development.

Introduction and literature review

In Computer-Supported Collaborative Learning (CSCL), students usually encounter difficulties to sustain social interactions, to advance knowledge, and to sustain collaborative learning momentum (Ouyang & Chang, 2019; Tsiotakis & Jimoyiannis, 2016). To improve the online discussion quality, network analytics, as one of the primary social learning analytics (SLA) methods, have been used to turn student learning data into actionable insights by analyzing and demonstrating student relations, positions, and characteristics (Bodily & Verbert, 2017; Hoppe, 2017; Ouyang, 2021). However, empirical studies indicate controversial results. Chen et al.(2018) found that social learning analytics tool named *CanvasNet* did not have significant effects on increasing students' social and cognitive engagement, whereas the research form Kwon and Park (2017) suggested that social network diagram promoted socially desirable responses while the cognitive diagram produced more cognitive responses. Taking a step forward, this research designs a student-facing network visualization tool to demonstrate social, topic, and cognitive aspects through social network, topic network and cognitive network. Moreover, three major methodological approaches, namely network-oriented analysis (e.g., social network analysis; SNA), content-oriented analysis (e.g., content analysis; CA) and process-oriented analysis (e.g., time series/temporal analysis; TA) were integrated as analytical approaches in CSCL research (Hoppe, 2017). Echoing this trend, this research uses a mixed method to examine the effects of three visualizations on students' engagement, particularly the temporal changes of cognitive structures by students with different social participation patterns.

Research methodology

Research context, participants and intervention

The research context was an online graduate-level, 8-week course titled "*Information Technologies and Education*" offered in the summer semesters in 2020.19 Ed.D. students (10 females, 9 males) from the College of Education enrolled in this course. The instructor designed the course discussion as open-ended, inquiry-based discussions, including synchronous discussions hosted in *DingTalk* and asynchronous discussions hosted in *XueZaiZheDa* forum. The tool intervention includes four phases, i.e., the control phase of non-intervention (Week 1-2), the social network intervention (Week 3-4), the topic network intervention (Week 5-6), and the cognitive network intervention (Week 7-8) (see Figure 1). During the discussion duration, we updated the network visualizations twice a week and embedded the tool as a webpage in the forum. Our research question is: *whether, to what extent and how did three network visualizations influence students' social-cognitive engagement?*

Data collection and analysis

The total number of posts from online asynchronous and synchronous discussion were 972 (Phase 1: N=207, Phase 2: N=234, Phase 3: N=374, Phase 4: N=157). First, SNA was used to identify three social participatory roles, i.e., *leader*, *regular*, and *peripheral*, which were identified in terms of the levels of participation (reflected by outdegree and outcloseness), influence (reflected by indegree and incloseness) and mediation (reflected by betweenness) (Ouyang & Chang, 2019). A *leader* student has high-levels of participation, influence and mediation

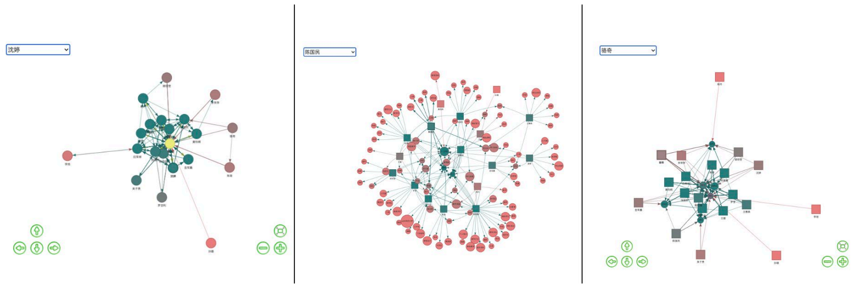


Figure 1. The network visualization intervention

(at least two of them are at the high level). A *regular* student has medium levels of participation, influence and mediation. A *peripheral* student has at least two low level of them. Second, content analysis was used to examine students' post content from a cognitive perspective. Referring to a previous work (Wise et al., 2014), the final coding scheme included content dimension (Information sharing, Sha; Perspective expression, Exp; Perspective elaboration, Ela) and discursiveness dimension (Perspective elicitation, Eli; Peer response, Res). Three trained raters coded the whole dataset separately, had multiple meetings to resolve discrepancies, and consulted with the first author to reach an agreement. Cohen's Kappa of inter-rater reliability was k=0.895. Finally, ENA was performed to examine the differences of cognitive structures between roles. We used an online ENA Tool (epistemicnetwork.org) (Marquart et al., 2018) and a R package named rENA to perform the ENA analyses.

Results

ANOVA results showed that, *leader* students had more contributions than *regular* and *peripheral*, on information sharing (Sha) ($F = 6.63$, $p < .01$), perspective expression (Exp) ($F = 6.73$, $p < .01$), and peer response (Res) ($F = 5.20$, $p < .01$) during Phase 1, on peer response (Res) ($F = 3.51$, $p < .05$) during Phase 2, on information sharing (Sha) ($F = 12.25$, $p < .001$) and peer response (Res) ($F = 7.63$, $p < .001$) during Phase 3, and on perspective elaboration (Ela) ($F = 4.40$, $p < .50$) and peer response (Res) ($F = 11.60$, $p < .001$) during Phase 4. ENA results demonstrated the differences of cognitive structures from the between-role perspective four phases (see Figure 2). We compared two roles that had statistical significances of cognitive structure as followings.

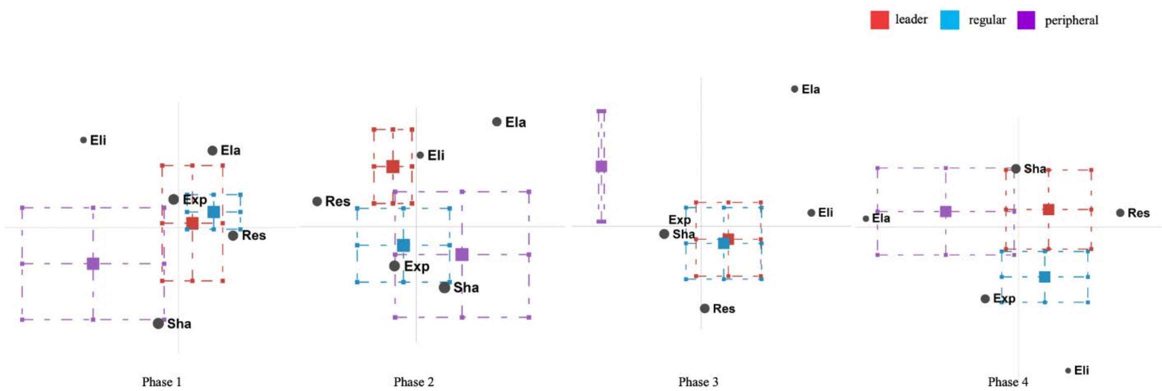


Figure 2. ENA mean plots of three social participatory roles during four phases

In Phase 1, *leader* had more cognitive connections than peripheral (see Figure 3). *Leader* had more Res-Sha, Res-Exp, Exp-Sha, and Res-Ela than *peripheral*, but *peripheral* had more Sha-Eli connection than *leader*. Similarly, *regular* had more cognitive connections than *peripheral*. *Regular* and *peripheral* had differences on Res-Ela, Res-Sha, and Res-Exp. But *peripheral* had more Sha-Eli connection than *regular*. Therefore, *leader* and *regular* tended to share information, express and elaborate perspectives when responding to peers. *Peripheral* tended to share information while eliciting questions rather than expressing or elaborating perspectives.

Under Phase 2 social network intervention, there were statistical significances between *leader* and *regular* and between *leader* and *peripheral*. *Leader* had more cognitive connections than *regular*. The most notable differences between *leader* and *regular* were the Ela-Eli, followed by Res-Eli and Res-Ela. But *regular* had a stronger Exp-Sha connection than *leader*. When comparing *leader* and *peripheral*, we did not see any connection of *peripheral* outweighing the connection of *leader*, suggesting that *leader* had more active cognitive engagement than *peripheral* overall. The most notable differences between *leader* and *peripheral* were Res-Exp and Res-Ela. Therefore, in Phase 2, *leader* tended to elicit peers' responses when sharing information, expressing and elaborating perspectives, while *regular* tended to sharing information and meanwhile expressing perspectives.

Under Phase 3 topic network intervention, there were significant differences between *leader* and *peripheral* and between *regular* and *peripheral*. *Peripheral* did not show any connections when comparing to *leader* and *regular*, indicating there were no cognitive connections of *peripheral* outweighing *leader* and *regular*. There were differences on Res-Sha, Res-Exp, Sha-Eli, Exp-Eli, and Res-Eli. Taken together, in Phase 3, *leader* and *regular* had similar cognitive structures to share information and express perspectives when responding to peers, while *peripheral* did not have cognitive connections that outweighed *leader* and *regular*.

Under Phase 4 cognitive network intervention, there were statistical significances among three roles. *Leader* had more Res-Sha and Res-Exp than *peripheral*, while *peripheral* had more Sha-Exp and Sha-Ela than *leader*. *Regular* had more Res-Sha and Res-Exp than *peripheral*, while *peripheral* had more Sha-Ela, Sha-Exp and Exp-Ela than *regular*. Compared with *regular*, *leader* had more connections of all codes except Res-Sha, although the difference was very small. The results suggested that like previous phases, *leader* students engaged more in cognitive activities than *regular* and *peripheral* students; but *peripheral* students tended to elaborate thoughtful perspectives although their low-level social participation.

Phase 1

Leader & Peripheral *Regular & Peripheral*

Phase 2

Leader & Regular *Leader & Peripheral*

Phase 3

Leader & Peripheral *Regular & Peripheral*

Phase 4

Leader & Peripheral *Regular & Peripheral* *Leader & Regular*

Figure 3. The comparison of ENA between two social roles in four phases

Discussions and implications

This research implements network visualizations to make available social, topic, and cognitive information to students with an expectation to improve student engagement in online discussions. We conclude that this network visualization tool influences cognitive structures in different ways for students with different social participatory roles. Our results show that the social network visualization has better effect on socially active students while cognitive network visualization has better effect on socially inactive students. Therefore, the design of SLA tools needs to consider learner agency, disposition and characteristic in the local context. From a pedagogical perspective, since the core of the student-facing SLA tool is its metacognitive feature (Durall & Gros, 2014), instructors should make deliberate stimulation to align students' metacognition with information demonstrated in SLA tools (Rodríguez et al., 2015; Ouyang et al., 2020). When peripheral students are provided with cognitive information, they develop a better self-awareness of their learning processes and thus their engagement increases overall. Future tool design should further integrate multiple analytic results to reflect learning in varied ways.

References

Bodily, R., & Verbert, K. (2017). Trends and issues in student-facing learning analytics reporting systems research. In *Proceedings of the seventh international learning analytics & knowledge conference* - LAK'17 (pp. 309-318). New York, NY: ACM Press.

Chen, B., Chang, Y. H., Ouyang, F., & Zhou, W. (2018). Fostering student engagement in online discussion through social learning analytics. *The Internet and Higher Education, 37,* 21-30.

Hoppe, H. U. (2017). Chapter 2: Computational methods for the analysis of learning and knowledge building communities. In C. Lang, G. Siemens, A. Wise & D. Gašević (Eds.), *Handbook of learning analytics (First edition)* (pp. 23-33). Creative Commons License 4.0.

Kwon, K., & Park, S. J. (2017). Effects of discussion representation: Comparisons between social and cognitive diagrams. *Instructional Science, 45*(4), 469-491.

Marquart, C. L., Hinojosa, C., Swiecki, Z., Eagan, B., & Shaffer, D. W. (2018). Epistemic network analysis [Software] Version 1.6.0. Website: epistemicnetwork.org.

Ouyang, F. (2021). Using three social network analysis approaches to understand computer-supported collaborative learning. *Journal of Educational Computing Research*. Online

Ouyang, F. & Chang, Y. H. (2019). The relationship between social participatory role and cognitive engagement level in online discussions. *British Journal of Educational Technology, 50*(3), 1396-1414.

Ouyang, F., Li, X., Sun, D., Jiao, P., & Yao, J. (2020). Learners' discussion patterns, perceptions and preferences in a China's massive open online course (MOOC). *The International Review of Research in Open and Distributed Learning, 21*(3). 264-284.

Rodríguez-Triana, M. J., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2015). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *British Journal of Educational Technology, 46*(2), 330-343.

Tsiotakis, P., & Jimoyiannis, A. (2016). Critical factors towards analysing teachers' presence in on-line learning communities. *The Internet and Higher Education*, *28*, 45-58.

Wise, A. F., Hausknecht, S. N., & Zhao, Y. (2014). Attending to others' posts in asynchronous discussions: Learners' online "listening" and its relationship to speaking. *International Journal of Computer-supported Collaborative Learning, 9*(2), 185–209.

Acknowledgments

The first author acknowledges the financial support from the National Natural Science Foundation of China (61907038).

Adapting Interaction Analysis to CSCL: A systematic review

Rolf Steier, Oslo Metropolitan University, rolf.steier@oslomet.no
Jacob Davidsen, Aalborg University, jdavidsen@hum.aau.dk

Abstract: Interaction Analysis (IA) (Jordan & Henderson, 1995) is a fundamental reference in the learning sciences, and a core method within the International Journal of Computer Supported Collaborative Learning. Surprisingly, despite the vast number of citations and labs around the world practicing forms of interaction analysis, there have been few if any substantial efforts to articulate its central premises in the context of CSCL. Following a systematic review method, the purpose of this preliminary study is to provide an overview and foundation for investigating the ways that IA has been interpreted and applied in the field of CSCL. Our findings suggest that IA has been applied in a variety of computer-mediated learning contexts and arrangements which have required extending and adapting the method in novel ways. Our broader goal is to consider ongoing methodological and technological developments for the future directions of interaction analysis within CSCL.

Introduction: The relations between Interaction analysis and CSCL

CSCL is generally concerned with the triadic relationship between human-technology-human (Ludvigsen & Steier, 2019) in a diverse number of learning settings and at different scales. In CSCL a wide range of methodological approaches have been used to study collaborative learning mediated by computers (Jeong, Hmelo-Silver and Yo, 2014) with different analytic foci. Within these diverse methodological approaches, however, it is also clear that one of the popular methods used for studying collaborative learning processes is Interaction Analysis (IA) (Jordan & Henderson, 1995). In 1995 the first CSCL conference was held and Jordan and Henderson's paper "Interaction Analysis: Foundations and Practice" was published in the journal of the Learning Sciences. Since then, researchers in CSCL have applied and adapted IA to studies of different collaborative arrangements mediated by technology. Hall and Stevens (2015) presented IA as a method for studying 'knowledge in use' which perhaps partly explains the popularity of IA in CSCL studies. The emergence of CSCL as a field over the past 25 years similarly reflects a growing interest in interactional meaning making as opposed to the mental representations of individuals (Stahl, Koschmann, & Suthers, 2006). In this short paper, we provide a brief description of the central premises of IA, followed by our methodological procedures for developing a corpus of relevant articles in ijCSCL. By examining and identifying the range the learning contexts in which IA methods have been adapted, we create a foundation for the re-specification of such methods in the field of CSCL.

Framing assumptions and practices of Interaction analysis

On a general level, IA is a set of foundations and practices that describe how video can be used for studying social processes of learning. IA concerns all the steps involved in using video - including camera work, the process of curating the video into data, how to transcribe and present the video. In addition, Jordan and Henderson also summarized seven core analytical foci (structure of events, temporal organization of activity, turn taking, participation structures, trouble and repair, spatial organization of activity, and artifacts and documents) - or ways of looking, that have proven to be fruitful in their own analytical work. The foci for analysis described by Jordan and Henderson are furthermore grounded in theories highlighting social aspects of learning processes, Conversation Analysis and Ethnomethodology. It is worth noticing that these foci each carry some historical legacy from the before mentioned theoretical positions, but the key point for Jordan and Henderson is to provide a more comprehensive way of working with and looking at the video data, which reflects their 'practice' of working with video from being affiliated with different labs in the early 1990's. The foci for analysis are not codes or categories that should be added to the video data - each of the foci offer a way of looking at the interaction at the different scales (the structure of events, the temporal organization of activity and turn-taking) and at specific resources used in the environment by the participants to organize their activity (participation structures, trouble and repair, the spatial organization of the activity, and artifacts and documents).

IA does not exclusively deal with learning in institutional settings like schools, but broadly speaking uses video to understand learning in a diverse number of co-located settings including human-machine interaction, hearing children and their deaf parents and design-group work, etc. In all of these settings IA has proven to be a useful method for studying how people go about learning together in different situations. The core interest in IA is "human activities, such as talk, nonverbal interaction, and the use of artifacts and technologies, identifying routine practices and problems and the resources for their solution." (Jordan and Henderson, 1995, p. 39) Hence, a key element for researchers performing IA is the focus on what people actually do, and not a retrospective account of what happened.

In IA, video is the primary means for capturing learning where learning is taking place - and Jordan and Henderson highlight several affordances of video, e.g. repeated viewing, permeance of the recording and the possibility to engage in collaborative analysis of the recording. In addition, IA researchers are also collecting other types of resources as a part of the research design, including notes, task description from the setting and general ethnographic information. The process and outcome of data collection has of course changed dramatically in the past 25 years because of innovations in video technology and one of our broader aims of reviewing the application of IA in ijCSCL is to document some of the ongoing developments.

Surprisingly, despite the vast number of citations and labs around the world practicing forms of interaction analysis, there have been few if any substantial efforts to articulate its central premises in the context of CSCL. Thus, the purpose of this study is to provide an overview and foundation for investigating the ways that IA has been interpreted and adapted in the field of CSCL specifically. Our broader goal is to consider ongoing methodological and technological developments for the future directions of interaction analysis within CSCL. We pose the following question: In what research contexts, including, environmental, technological, and collaborative arrangements, have IA methods been applied in the International Journal of Computer Supported Collaborative Learning? Extending this line of inquiry into the future will allow us to reflect on and respecify our methodological premises for investigating new collaborative learning interactions.

Methods: Systematic review and building a corpus

The current short paper presents preliminary analyses of a larger project investigating the role of IA methods in CSCL. The project is conducted as a systematic review (Gough, Oliver, & Thomas, 2017), and the data for this paper include a selected corpus of 48 articles published in ijCSCL between 2006 and 2019 (14 years). This corpus was produced based on the following search queries and criteria:

We first identified papers that cite Jordan & Henderson 1995 directly. This produced 32 articles. We then searched for "interaction analysis" which produced 65 results. Removing duplicates (29); editorials (12); superficial references to "interaction analysis" (e.g., reviewing another study, or otherwise peripheral to the methods of the study); and uses of "interaction analysis" from a different research tradition (e.g., Gundawara, Lowe, & Anderson, 1997) produced 3 additional papers for the corpus.

Finally, because our purpose is to identify methodological trends in ijCSCL and not statistically analyze the self-labeled methods of researchers, we wish to ensure a comprehensive and broad corpus. Thus, we searched for related terms including "Video", "Conversation Analysis", and "Ethnomethodology" to identify additional papers rooted in the same methodological tradition as Jordan & Henderson (1995). Such studies needed to be qualitative, based on a sequential, turn-based analysis of interaction, and should generally rely on some form of transcription of video data or digital recordings. Studies with methods primarily based on coding, thematic analysis, or content analysis were not included. This added (17) more papers to the corpus resulting in a total of 52 articles. For the current paper, which is investigating empirical conditions for performing IA, we removed 4 articles without a clear empirical setting to arrive at 48 articles used in this analysis.

We intentionally took an expansive approach to identify these studies which means the role of interaction analysis and the 1995 text is more central in some studies than others. While a limitation of this approach is that making precise categorical claims about the collection of studies becomes problematic, the advantage is that we are able to take a more holistic view of methodological developments in the journal. Another reason for including an expanded search beyond the initial 32 references is grounded in the many voices Jordan and Henderson integrated in the original paper; IA builds upon a diverse set of theoretical and methodological positions and some researchers in CSCL are informed by some of these voices, and do not cite Jordan and Henderson explicitly.

The broader project from which this study is based follows a systematic review of the above corpus. For this initial study, the authors reviewed the selected 48 articles in the corpus with attention to the empirical settings of the research, noting 1) where the learning activities were occurring (e.g. school, museum, etc.); 2) the primary mediating technology (e.g. computer, mobile phone, etc.); and 3) the group size of the participants (e.g. small group, whole class, etc.). These categories were collaboratively refined by the authors. We note that this is a subjective process, and in many cases, such categories are not clear. For example, many articles in the corpus incorporate a variety of collaborative technologies or occur across multiple settings. In such instances, we have tried to emphasize the primary empirical focus of the studies - see Table 1 for definitions of these categories.

Data and analysis

Since the inauguration of ijCSCL roughly 300 papers have been published and 48 of the papers are referring to or performing an analysis of interactions confirming the popularity of IA in ijCSCL. Based on our review it is evident that IA has been used to study collaborative learning in many different settings (Figure. 1), but it is also clear that more than half of the studies are located in primary/secondary education. Three studies are conducted in experimental laboratory settings, whereas the rest take place in what can loosely be determined as interaction occurring in natural settings (in most cases as part of a research project). The fact that researchers have used IA in many different settings shows that there is a broad interest in understanding learning as social interaction and that IA is not exclusively used to study collaborative learning in one particular setting. In our review of the papers, we furthermore see that different studies do not necessarily apply the same elements from IA. While some studies adopt particular foci for analysis, others adopt relevant procedures for working with data, thus adapting the method to fit the particular study and the research question being addressed in the paper.

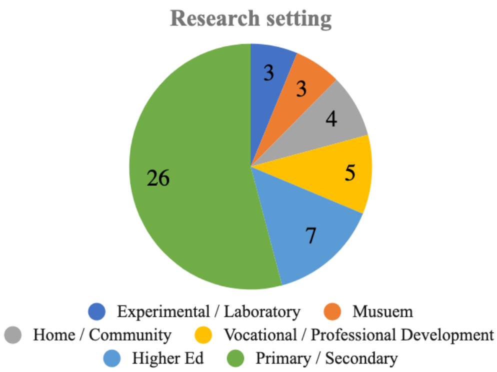


Figure 1. Research setting of interaction analysis and related methods in ijCSCL corpus.

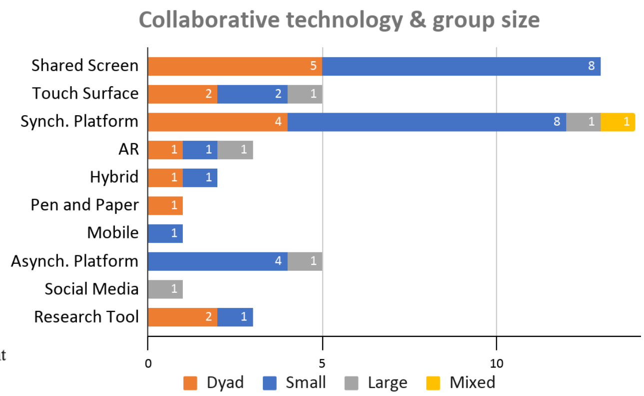


Figure 2. Comparing collaborative technology & group size by paper in ijCSCL corpus. Dyad, Small, and Large refers to group sizes of 2, 2-8, and 8+ respectively. Mixed refers to a combination of group sizes.

Looking into the relations between the collaborative technologies and group size (Figure 2.), we see that the majority of studies work with either dyads and small groups using a shared screen, touch surface or a synchronous platform. However, it is also evident that IA studies in ijCSCL are not exclusively dealing with one group size using a particular technology. New collaborative arrangements have emerged in CSCL as a result of new technologies, e.g. touch surface, augmented reality and social media platforms. Thus, the adaptation and innovations of IA should be viewed as genuine interest in understanding collaborative learning in different constellations of togetherness with different technologies. We only see two studies conducted in hybrid environments and only one study is working pedagogically with a mix of group sizes (small and large group).

Table 1: Category definitions

Shared Screen	Multiple users sitting/standing around the same computer screen or device.
Touch Surface	Touch tables, multi-user touch screens, and interactive white boards. Does not include tablets or mobile phones.
Synch. Platform	Computer /web-based platform in which participants collaborate on their own individual device, and interaction mostly takes place in the platform (not in front of the screen).
AR	Augmented Reality, Mixed Reality, or Simulators with physical and digital components.
Hybrid	Analytic focus is on a range of mediating technologies.
Pen and Paper	Analogue tools like pen and paper.
Mobile	Mobile technology like mobile phones. Does not include laptop computers or handhelds unless mobility is a part of the design/ task.
Asynch. Platform	Platforms and learning management systems mainly using asynchronous communication.
Social Media	Publicly available social media platforms.
Research Tool	Tool for visualizing/ working with CSCL data. The tool is not used by the participants, but the primary goal of the study is to develop a tool to describe participant interactions.

Another interesting finding is that IA has not only been applied to co-located collaborative activities but is also used to study collaborative learning in synchronous platforms (14), asynchronous platforms (5) and social media platforms (1). In these papers the authors have adapted and innovated IA to study social interaction in different media. In some of these papers social interaction is mediated by text in chats or online forums, which is in contrast to the co-located studies presented by Jordan and Henderson (1995). Such studies depict significant methodological development in terms of data collection methods that do not rely on video, instead turning to screen capture, data logging, and related procedures. These studies similarly demonstrate analytic innovation in terms of how interaction is conceptualized across novel temporal and spatial arrangements. Further analysis may better articulate such developments in CSCL.

In 3 studies, IA has been used as part of a broader methodological stance to inform the development of a research tool. The purpose of developing the research tools is primarily a search for finding new ways of visualizing collaborative learning activities. Such tools might be viewed as extensions of Jordan and Henderson's emphasis on transcription and representation methods.

Finally, it is important to note that IA has been used to study collaborative activities mediated by a range of different technologies. For Jordan and Henderson artifacts/technologies was one among many possible analytic foci, whereas the mediating technologies tend to be foregrounded in ijCSCL. This is perhaps not surprising as the core interest in many of the papers is to investigate what role the mediating technology is having on collaborative learning. But it is important to emphasize that such technologies strongly inform the ways that collaboration is conceptualized. In other words, these findings suggest that much more nuance is needed when examining the role of mediating technology in ijCSCL than is provided by the single analytic foci described by Jordan & Henderson. Synchronous and asynchronous collaboration are fundamentally different activities, for example, and there is also a clear distinction between collaboration that occurs *within* a computational artifact (e.g., Synch. Platform) and that which occurs outside the artifact (e.g., Shared Screen). Overall, it seems fair to state that IA has proven useful for uncovering social mechanisms of collaboration in the triadic relationship between human-technology-human.

Discussion and Conclusion

The findings reported in this short paper are part of a larger ongoing systematic review on the use of IA in ijCSCL. They suggest that Interaction Analysis is a popular method in ijCSCL used in different settings to better understand how learners go about learning together using different technology. The majority of papers in our corpus research how dyads or small groups learn together, a few papers work with collaborative activities in larger groups, and only 1 paper is looking across different constellations of collaboration. It is less surprising to see the number of studies with dyads and small groups as CSCL researchers have promoted the idea of small groups as being a primordial site for studying meaning making processes (Stahl, 2006). In addition, this study also shows that researchers in ijCSCL have developed innovations for studying different collaborative arrangements moving beyond the face-to-face arrangements originally described by Jordan and Henderson (1995). As we move into the future, new kinds of technologies will emerge, and it is important that we develop and adapt IA to capture these different forms of reality/interaction (virtual, augmented, mixed, etc.) without forgetting the foundations and practices of IA. We have identified some general trends with this preliminary study; however, we have only scratched the surface and the next steps would be to identify the qualitative nature of the methodological developments in the field - to better understand the history and future practices of IA in CSCL. Moreover, a significant interest is also to articulate the status of collaboration across the different interaction analytical studies.

References

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews*. Sage.

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.

Hall, R., & Stevens, R. (n.d.). Interaction Analysis Approaches to Knowledge in Use. In A. A. diSessa, M. Levin, & N. J. S. Brown (Eds.), *Knowledge and Interaction A Synthetic Agenda for the Learning Sciences* (Vol. 1, p. 37).

Jeong, H., Hmelo-Silver, C. E., & Yu, Y. (2014). An examination of CSCL methodological practices and the influence of theoretical frameworks 2005–2009. *International Journal of Computer-Supported Collaborative Learning*, 9(3), 305–334. https://doi.org/10.1007/s11412-014-9198-3

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundations and Practice. T*he Journal of the Learning Sciences*, 4(1), 39–103.

Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. *Cambridge Handbook of the Learning Sciences*, 2006.

Ludvigsen, S., & Steier, R. (2019). Reflections and looking ahead for CSCL: Digital infrastructures, digital tools, and collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, 14(4), 415–423. https://doi.org/10.1007/s11412-019-09312-3

Stahl, G. (2006). *Group cognition computer support for building collaborative knowledge*. MIT Press.

International Society of the Learning Sciences

Examining the Influence of Instructor Interventions on Group Collaboration

LuEttaMae Lawrence, Carnegie Mellon University, llawrenc@andrew.cmu.edu
Taylor Tucker, University of Illinois at Urbana–Champaign, tdtucke2@illinois.edu
Emma Mercier, University of Illinois at Urbana–Champaign, mercier@illinois.edu

Abstract: Collaborative problem solving is often used in STEM higher education courses to support conceptual knowledge and teamwork. However, course teaching assistants (TAs) often lack the collaborative pedagogical knowledge necessary to orchestrate this form of learning. In this paper, we examine TAs' orchestration strategies and technology used to understand how these factors influence groups' collaboration. Contributions from this paper describe the interplay among technology, strategies, and groups' collaboration toward understanding how to support collaboration in these courses.

Introduction

Collaborative problem-solving has become a common pedagogy in postsecondary STEM courses (e.g., Freeman et al., 2014). However, instructors often lack the skills needed to facilitate effective collaboration (Greiffenhagen, 2012). Efforts to increase the use of collaborative learning in engineering courses have been driven by research indicating that this form of pedagogy allows students to both deepen their conceptual knowledge and develop better team skills (e.g., Barron & Darling-Hamond, 2008). Prior research has shown that graduate teaching assistants (TAs), who frequently teach core engineering courses, often lack the pedagogical knowledge to monitor, assess, and support groups' real-time collaborative interactions (Shehab, 2019; Lawrence, 2020). There is an ongoing need to support TAs in identifying groups' progress and orchestrating collaborative interactions; thus, it is necessary to present TAs with actionable information and recommendations to help them navigate groups who may need collaborative support. In this paper, we describe a study that investigates TAs' orchestration strategies with groups while using a real-time, supportive orchestration tool and examines how these strategies affected groups' interactions.

Perspectives

Orchestration technology has been shown to support teachers' monitoring and awareness of student behaviors and needs in real-time (Holstein, Aleven, & McClaren, 2018; Martinez-Maldanado et al., 2017). Using collaborative orchestration technology requires both technical proficiency and a strong grasp of collaborative pedagogical practices (Dimitridas, 2012). While engineering TAs are often equipped with sufficient knowledge for general technology use in the classroom, they have a wide range of views and experiences with collaborative learning that impacts how they interact in classrooms that embed both (Lawrence, 2020). In light of this range, it is necessary to provide TAs with resources that can help them learn about and facilitate collaboration.

To support small-group collaboration, research suggests that TAs should use monitoring strategies to observe and identify the nature of groups' interactions and, when necessary, follow up with interventions that prompt groups to talk (Hoffmann & Mercer, 2016; Kaendler et al., 2016; Shehab, 2019). However, research indicates that, without training, implementation of these strategies does not occur naturally (Kaendler et al., 2016). This instruction gap creates the opportunity for orchestration technology that embeds features to alert and advise TAs of instances that may require a collaborative intervention (van Leeuwen, Rummel, & van Gog, 2017). We hypothesize that, with resources to help identify and support collaborative interactions, instructors' interventions can support change in students' collaboration. In this paper, we study TA-student interactions to understand the interplay among teacher's orchestration strategies, technology, and groups' collaboration. We will explore these interactions to answer the following research questions:

1. How did TAs interact with the orchestration tool and groups of students during discussion sections?
2. How did the TAs' interventions and tool use affect groups' collaborative interactions?

Methods

Design

This study is part of a multi-year design-based implementation research project (Penuel, Fishman, Cheng, & Sabelli, 2011) that focuses on supporting collaborative problem-solving in undergraduate engineering discussion

 © ISLS

sections. In these courses, students worked in groups of two to four on tasks presented through synchronized drawing software on 11" tablets, which allowed group members to view and modify each other's work. The tasks, which were designed using a literature-based framework (Shehab et al., 2017), delivered ill-structured, real-world problems. The orchestration tool, which was co-design with TAs (Lawrence & Mercier, 2019), used machine learning models (Paquette et al., 2018) to 1) provide alerts of groups' probable need for support that could be confirmed or denied, and 2) provide strategies for intervention (Figure 1). Additionally, TAs could join groups to view students' work in real-time. Video and log file data of students and TAs were collected; two weeks of data are analyzed in this paper.

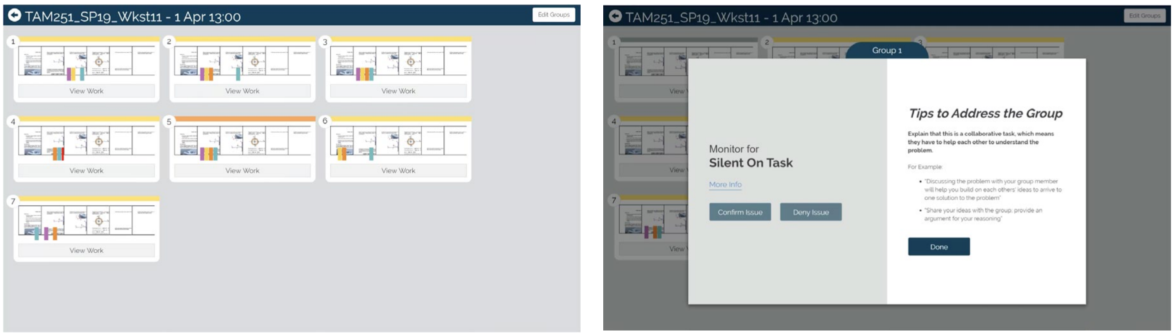


Figure 1. The image on the left shows the orchestration tool interface visualizing seven groups. The image on the right shows an alert that was selected and confirmed by an instructor and strategies to address the group.

Participants

Participants were 90 undergraduate engineering students (20 females, 70 males) who were registered for a required introductory engineering course. Students were separated into 26 groups across five discussion sections. Groups remained consistent throughout the entire semester. Each discussion section had three TAs consisting of one graduate student and two undergraduate students. In total, eight TAs (two graduate and six undergraduate students) taught the five classes.

Analysis

To understand how TAs' interventions affected groups' interactions, interventions were identified in the video data and transcribed in playscript form (Sullivan & Forrester, 2018). Each transcribed intervention was framed by 20 seconds of student dialogue before and 30 seconds after, building on prior monitoring and intervention analysis (Shehab, 2019; Lawrence, 2020). Log file data were reviewed to identify tool use during interventions. Each intervention was coded for the presence of orchestration strategies that were derived from literature and past research (Table 2; percent agreement ranged from 89% to 99%; Cohen's Kappa ranged from 0.25 to 0.96). Students' behaviors before and after the intervention were coded for talking or silent, engaged (e.g. nodding, making eye contact with the speaker) or not, and on-task or off-task (IRR = 0.91). Using these codes, each group was categorized as collaborative or non-collaborative. A collaborative group was identified when the majority of students were engaged and talking on task; a non-collaborative group was identified when the majority of students were silent, talking off task, or split across codes, indicating a lack of cohesion.

Results

Across the two weeks, the TAs engaged in 223 interventions with groups (Table 1). The machine learning models presented 678 alerts; 374 were opened by an instructor. Of those 374, 80 were confirmed–meaning the instructor confirmed that they perceived the detected behavior as correct. An instructor viewed a group's work 79 times. Thirty-nine instances of tool use led to an intervention with students, including 38 alerts and one "view work."

To understand how TAs interacted with groups, each intervention was coded for the presence of orchestration strategies. Table 1 illustrates the frequency of each strategy across all types of interventions. During the majority of interventions, the instructors initiated by probing for the groups' understanding, held the whole group's attention, and chose to explain content after students asked questions or expressed confusion (as compared to explaining without prompting). Interventions, where the TA used the tool, were less likely to explain content without students asking questions or expressing confusion. These interventions were also more likely to be initiated by the instructor. Several orchestration strategies were infrequently enacted by TAs across all interventions, including monitoring a group before intervening, prompting group members to talk to each other, and ending the intervention by checking for students' understanding.

Table 1: Instructors' orchestration strategies across all interventions, those that had tool use and those that did not.

Orchestration Strategies	Frequency (%)		
	All Interventions	Interventions with Tool Use	Interventions without Tool Use
Total interventions	223 (100%)	39 (100%)	184 (100%)
Instructor monitored the group	58 (26%)	14 (36%)	44 (24%)
Instructor initiated intervention	111 (50%)	27 (69%)	84 (46%)
Instructor initiated intervention by probing for the groups' understanding	132 (59%)	22 (56%)	110 (60%)
Instructor explicitly prompted the group to talk	22 (10%)	3 (8%)	19 (10%)
Instructor had the whole group's attention	154 (69%)	21 (54%)	133 (72%)
Explanations were preceded by question or confusion from a student	136 (61%)	17 (44%)	119 (65%)
Instructor ended by checking for the group's understanding	43 (19%)	3 (8%)	40 (22%)

We coded before and after each intervention to understand if the instructors' intervention and tool use affected students' collaborative interactions. Across all groups, 56% (N = 134) were in a collaborative state (e.g., talking on task and engaged) before an intervention started. After an intervention occurred, 60% (N = 89) of groups were in a non-collaborative state. Of the groups that were in a collaborative state before the TA engaged in an intervention, half remained in a collaborative state and half transitioned into a non-collaborative state (Table 2). The majority (73%) of groups that were in a non-collaborative state pre-intervention transitioned into a collaborative state post-intervention.

We also compared groups' collaborative states pre- and post-intervention specifically for those in which the TA used the tool. Most groups (60%) that were in a collaborative state before such an intervention transitioned into a non-collaborative state once the TA left. Of groups that were in a non-collaborative state before a tool-based intervention, 57% transitioned into a collaborative state post-intervention while 43% remained in a non-collaborative state.

Table 2: Groups' collaborative states pre- and post-intervention for all interventions and for those where the instructor used the orchestration tool.

Status Before an Intervention	Status After an Intervention	Frequency (%) of all Interventions	Frequency (%) with the Tool
Collaborative	Collaborative	62 (50%)	10 (40%)
	Non-collaborative	62 (50%)	15 (60%)
Non-Collaborative	Collaborative	72 (73%)	8 (57%)
	Non-collaborative	27 (27%)	6 (43%)

Conclusions and implications

The goal of this study was to understand how instructors' orchestration strategies and tool use during interventions affected groups' collaboration. Several orchestration strategies, including monitoring and checking for understanding, have been shown to be beneficial for collaborative activities (Hoffmann & Mercer, 2016; Kaendler et al., 2016; Shehab, 2019). Prior research shows that these strategies do not occur spontaneously and require support for successful enactment (Kaendler et al., 2016). Our findings indicate that the orchestration tool prompted the TAs to enact some orchestration strategies more frequently as compared to interventions without the tool.

Findings from the groups' interactions show that groups who were in a non-collaborative state before an intervention benefited from interventions with and without the orchestration tool, meaning that they transitioned into a collaborative state after the TA engaged in an intervention. While some tool-based interventions helped move students toward collaboration, our initial findings have not revealed a concrete difference in benefits between the two types of interventions. Future work will examine the relationship between individual orchestration strategies and groups' changes in collaboration to further understand what strategies were most useful in supporting groups' collaborative states. Further analysis will be used to inform how and what strategies are presented through the orchestration tool to support instructors. This work has important implications regarding the interplay among technology, orchestration strategies, and groups collaboration, and sheds light on the need to understand how to support novice instructors (Greiffenhagen, 2012) as they facilitate collaboration in higher-education engineering courses.

References

Barron, B., & Darling-Hammond, L. (2008). *Teaching for Meaningful Learning: A Review of Research on Inquiry-Based and Cooperative Learning*. Book Excerpt. George Lucas Educational Foundation.

Dimitriadis, Y. A. (2012). Supporting teachers in orchestrating CSCL classrooms. In Research on E-Learning and ICT in Education (pp. 71-82). Springer, New York, NY.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410-8415.

Greiffenhagen, C. (2012). Making rounds: The routine work of the teacher during collaborative learning with computers. *International Journal of Computer-Supported Collaborative Learning*, *7*(1), 11-42.

Hofmann, R., & Mercer, N. (2016). Teacher interventions in small group work in secondary mathematics and science lessons. *Language and Education*, *30*(5), 400-416.

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Informing the design of teacher awareness tools through causal alignment analysis. *International Society of the Learning Sciences, Inc.* [ISLS].

Kaendler, C., Wiedmann, M., Rummel, N., & Spada, H. (2015). Teacher competencies for the implementation of collaborative learning in the classroom: a framework and research review. *Educational Psychology Review*, *27*(3), 505-536.

Lawrence, L. & Mercier, E. (2019). Co-design of an orchestration tool: Supporting engineering teaching assistants as they facilitate collaborative learning. *Interaction Design and Architecture(s) Journal,* (42), 111- 130.

Lawrence, L. (2020). The design process of a collaborative orchestration tool and its implications for instructor uptake. (Doctoral dissertation, University of Illinois at Urbana–Champaign). ProQuest Dissertations Publishing.

Martinez-Maldonado, R., Yacef, K., Dos Santos, A. D. P., Shum, S. B., Echeverria, V., Santos, O. C., & Pechenizkiy, M. (2017, July). Towards proximity tracking and sensemaking for supporting teamwork and learning. In *IEEE 17th International Conference on Advanced Learning Technologies* (ICALT) (pp. 89-91). IEEE.

Paquette, L., Bosch, N., Mercier, E., Jung, J., Shehab, S., & Tong, Y. (2018). Matching data-driven models of group interactions to video analysis of collaborative problem solving on tablet computers. *International Society of the Learning Sciences, Inc.* [ISLS].

Penuel, W. R., Fishman, B. J., Cheng, B., & Sabelli, N. (2011). Developing the area of design-based implementation research. *Menlo Park, CA: SRI International*.

Shehab, S., Mercier, E., Kersh, M., Juarez, G., & Zhao, H. (2017). Designing Engineering Tasks for Collaborative Problem Solving. In *Making a Difference—Prioritizing Equity and Access in CSCL: The 12th International Conference on Computer Supported Collaborative Learning*. Philadelphia: The International Society of the Learning Sciences.

Shehab, S. S. (2019). Collaborative problem solving in higher education classrooms: Exploring student interactions, group progress, and the role of the teacher (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Sullivan, C., & Forrester, M. A. (Eds.). (2018). *Doing qualitative research in psychology: A practical guide*. Sage.

van Leeuwen, A., Rummel, N., & van Gog (2017). Teacher regulation of collaborative learning: research directions for learning analytics dashboards. In *Making a Difference: Prioritizing Equity and Access in CSCL*, *2*(805-806), 1939-1382.

Acknowledgments

The project was funded by the National Science Foundation's Cyberlearning and Future Learning Technologies Program (Award Number: 1628976).

Examining Contrasting Collaborative Programming Behaviors among Three Pairs

Dan Sun, Fan Ouyang, Yan Li
dansun@zju.edu.cn, fanouyang@zju.edu.cn, yanli@zju.edu.cn
Zhejiang University

Abstract: Pair programming is a collaborative learning mode to foster novice learners' computer programming. Previous empirical research results in contrasting conclusions about the effect of pair programming on student learning. To further understand students' pair programming, this study uses interaction analysis approaches to examine three contrasting pairs' collaborative behaviors in Minecraft programming platform. The results show that the high-ranked student pair is characterized as the interactive and goal-oriented pair; the middle-ranked student pair is characterized as the highly-interactive and process-oriented pair; and the low-ranked student pair is characterized as the lowly-interactive and programming-distracted pair. Based on the results, this research proposes pedagogical and theoretical implications for future instructional design and empirical research of collaborative programming.

Introduction

Computer programming is one of the main learning modes to improve students' computational thinking (CT) in K-12 schools (Brennan & Resnick, 2012; Lye & Koh, 2014; Wing, 2006). Compared to solo programming, pair programming, as a collaborative learning mode, is a practical strategy for students to solve challenging problems and generate creative ideas. However, previous empirical research shows discrepancies about the effect of pair programming on student learning. Primary factors that influence the effectiveness and quality of pair programming are individual's programming ability (Salleh et al., 2011), social atmospheres (Werner et al., 2005), and the group configurations (Demir & Seferoglu, 2020). Given the complex factors that may influence the pair programming, it is necessary to conduct a fine-grained, multi-dimensional analysis of pair programming in order to provide research, analysis and practice implications.

Literature review

Grounded upon the social, cultural, situated perspectives of learning, collaboration is defined as a group of people participating in coordinated activities to maintain shared understandings, to solve shared problems of a project, and to create new knowledge or relevant products (Goodyear et al., 2014). Pair programming, as a computer-supported collaborative learning mode, supports two programmers working together at one workstation to solve the same programming problem (Braught, Wahls, & Eby, 2011). Empirical studies have indicated that pair programming, under favorable conditions, is beneficial for students to develop the solution to ill-structured programming tasks (Liebenberg et al., 2012), to advance their programming knowledge (Umapathy & Ritzhaupt, 2017) and to foster their computational and creative thinking skills (Zhong, Wang, & Chen, 2016).

Although pair programming can improve student learning under some conditions, previous empirical research shows mixed results for students' collaborative behaviors. For example, through analyzing and identifying code patterns, Hwang et al. (2012) discovered that students had different coding patterns during pair programming, including increasing, decreasing and no transition of six programming behaviors during the problem-solving processes. Satratzemi et al. (2018) found that students had varied levels of programming performances, significantly related to a student's previous programming experiences and confidence in programming. Using discourse analysis and epistemic network models, Wu, Hu, Ruis and Wang (2019) revealed that a high-performing team exhibited programming with a systematic approach, whereas a low-performing team used tinkering or guess-and-check approaches. Overall, although collaborative strategies are widely used to improve students' programming, relevant research indicates different results in terms of students' behaviors.

This study uses a pair programming strategy to improve students' programming within the Minecraft programming environment in China's secondary education contexts. Using an interaction analysis approach (i.e., click stream analysis and classroom video analysis), we investigate three contrasting pairs' collaborative behaviors. Based on the results, this study proposes pedagogical and theoretical implications.

Method

Research purpose and question

The overarching research purpose is to empirically investigate pair programming behaviors during collaborative programming. Among ten pairs of students, we identify three contrasting pairs in terms of their individual procedural performances. Our research question is: *What* are *the differences of collaborative behavior of three contrasting pairs during the pair programming process?*

Research context and participants

The research context is an optional course titled "*The Interactive Programming in Minecraft*" offered at a junior high school during Spring 2019 in the Eastern area of China. Twenty 7th graders (2 females, 18 males) enrolled in this 12-week course; they were all novice programmers with no text-based Python language programming experience. Students were designated into ten pairs at the beginning under the instructor's arrangement. Minecraft was adapted as the programming learning environment to facilitate students' learning of the Python language. Students built basic structure using mouse click and created interactive function with structures by Python in Minecraft, and Python was connected with Minecraft through StartServer.bat in Adventures In Minecraft. Among ten pairs, we identified three contrasting pairs based on students' individual procedural performance scores (see Figure 1). Two students of pair 1 (the high-ranked pair) achieved the scores of 86 and 70 (M = 78.00, SD = 9.87); two students of pair 2 (the middle-ranked pair) had the scores of 70 and 65 (M = 67.50, SD = 7.57); two students of pair 3 (the low-ranked pair) had the scores of 30 and 50 (M = 40.00, SD = 19.57). The average score for all students in the course was 63.

(a) Pair 1 (b) Pair 2 (c) Pair 3

Figure 1. Three pairs' collaboration during a class session.

Data collection and analysis approaches

We recorded students' online and offline behaviors through the computer screen-running videos and a whole class video (without sounds). We used click stream analysis (Filvà, Forment, García-Peñalvo, Escudero, & Casañ, 2019) and classroom video analysis (Kersting, 2008) to analyze those two types of data in order to identify pairs' programming behaviors. Two coders followed open coding analysis, first individually watched the computer screen-running videos and the classroom videos, identified initial codes of online and offline behaviors, then had discussions to achieve an agreement of the final coding framework. Shown in Table 1, eight behaviors were identified. Then two coders independently coded the data again in a chronological order based on the coding framework, and reached an inter-rater reliability with a Cohen's Kappa of .888.

Table 1. The coding framework for online and offline behaviors.

Code	Description
Project Understanding (PU)	A student transferred to the task window to understand the programming projects.
Python Coding (PC)	A student wrote codes with the Python language in the system.
Minecraft Debugging (MD)	A student debugged the code in Minecraft.
Minecraft Gaming (MG)	A student played games in Minecraft.
Programming Assistance (PA)	A student looked at the computer and assisted his/her partner to program.
Partner Discussion (PD)	A student discussed with his/her partner during the programming.
Instructor Communication (IC)	A student(s) talked with the instructor.
Classmate Communication (CC)	A student(s) talked with other classmates in the class.

 © ISLS

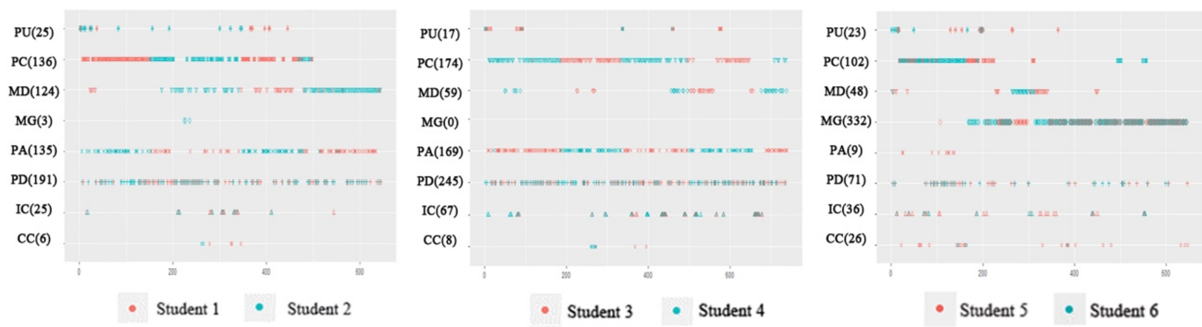
Note. The first four codes were identified through the clickstream data as online behaviors; the last four codes were identified through the video data as offline behaviors.

Results

Pair 1, a high-ranked pair among three pairs, was identified as *the interactive* and *goal-oriented pair*. Pair 1 not only had the highest scores of highest frequency of codes, but also solved all the tasks for the final programming project with the highest scores of 96 among three groups. The most frequent behaviors were Partner Discussion (PD; frequency = 191), Python Coding (PC; frequency = 136), and Programming Assistance (PA; frequency = 135), which all ranked at the middle level among three groups (see Fig. 2). Moreover, the temporal graph showed that the programming behaviors of PD and PA distributed evenly across the time period, which indicated that students in Pair 1 consistently communicated with and assisted each other during the pair programming. Among three pairs, Pair 1 had the highest frequency of Minecraft Debugging (MD; frequency = 124), which were interwoven with the Python Coding (PC) behavior.

Pair 2, the middle-ranked pair among three pairs, was identified as *the highly-interactive* and *process-oriented pair*. The most frequent behaviors were Partner Discussion (PD; frequency = 245), Python Coding (PC; frequency = 174), and Programming Assistance (PA; frequency = 169), which also ranked highest among three pairs (see Fig.2). According to the temporal graph, Pair 2 first attempted to understand the programming project (PU), discussed with the partner (PD), and then turned to the Python coding (PC) process. Pair 2 also debugged in Minecraft (MD) several times to modify the codes, and further continued the Python coding (PC) process.

Pair 3, the low-ranked pair among three pairs, was identified as *the lowly-interactive and programming-distracted pair*. The most frequent behaviors were Minecraft Gaming (MG; frequency = 332), Python Coding (PC; frequency = 102), and Partner Discussion (PD; frequency = 71). Among three pairs, Pair 3 had the highest frequency of Minecraft Gaming (MG; frequency = 332) and the lowest frequency of Partner Discussion (PD; frequency = 71), Minecraft Debugging (MD; frequency = 48), and Programming Assistance (PA; frequency = 9). According to the temporal results (see Figure 2), Pair 3 first attempted to understand the programming project (PU), and then focused on Python coding (PC) through partner discussions (PD). However, from the middle to the end of the programming period, both of the students were constantly attracted by Minecraft Gaming (MG), which was a sign for giving up the programming problem-solving.



(a) Pair 1 (b) Pair 2 (c) Pair 3

Figure 2. The temporal graph of collaborative behavior changes.
Note. The x-axis represents the time period; y-axis represents online behaviors. The total frequency of eight behavior codes for pair 1, pair 2 and pair 3 were 645, 739 and 647 (M = 677.00, SD = 53.70).

Discussion

Since previous empirical studies resulted in contrasting conclusions about the effect of pair programming on students' learning quality, this study conduct a fine-grained analysis of three pairs' programming behaviors. Echoing previous studies (Yang, Chen, & Hwang, 2015), this research reveals discrepancies among different pairs and complex correlations between programming behaviors, which may have significant influences on the collaborative programming quality, performance and experience. Based on the results, this study proposes pedagogical and theoretical implications for instructional design and empirical research of collaborative programming. On the pedagogical level, the instructors should use the process-oriented interventions to foster pair programming. The instructors should identify the specific learning situations that are appropriate to provide instructional interventions during the collaborative programming process. Our results showed that the low-performing groups could easily get distracted by irrelevant activities (e.g., gaming), such that the instructor needed to provide on-time assistance to guide students on the programming track (Wang & Hong, 2018). More incentives and assistance need to be given for low-achievers at the early stage of a problem-solving period to intervene around their programming behaviors and motivations (Hwang et al., 2012). On the theoretical perspective, student agency is demonstrated by the students' intentionality toward and their action of taking learning initiations (Bandura, 2001), which should be promoted in novice learners' programming process. The results showed that high-performing student pairs took actions to initiate questions, share and negotiate knowledge and create the programming solutions. They together achieved cognitive accomplishments that exceed the knowledge of any individual members for the higher-order programming work (Stahl, 2005). In other words, group members' synergistic coordination of the peer interactions and programming behaviors can lead to the high quality of the problem-solving process.

Conclusion

Collaborative programming is a promising yet challenging for novice programmers. This research selects three contrasting pairs in China's secondary education and conducts a fine-grained analysis of the pair programming processes. The results reveal differences among three pairs in terms of collaborative behaviors during the pair programming process. Although this research merely focuses on a small size of student sample, it makes contributions on the pedagogical, analytical, and theoretical perspectives. Since the intrinsic value of programming centers on its process, relevant research and practice should take a process-oriented perspective to investigate, advance, and assess students' programming in order to foster a sustainable learning.

References

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1–26.

Braught, G., Wahls, T., & Eby, L. M. (2011). The case for pair programming in the computer science classroom. *ACM Transactions on Computing Education, 11*(1), 1-21.

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. *Proceedings of the 2012 annual meeting of the American educational research association* (pp. 1-25). Vancouver, Canada.

Filvà, D. A., Forment, M. A., García-Peñalvo, F. J., Escudero, D. F., & Casañ, M. J. (2019). Clickstream for learning analytics to assess students' behavior with scratch. *Future Generation Computer Systems, 93*, 673-686.

Hwang, W. Y., Shadiev, R., Wang, C. Y., & Huang, Z. H. (2012). A pilot study of cooperative programming learning behavior and its relationship with students' learning performance. *Computers & Education, 58*(4), 1267-1281.

Kersting, N. (2008). Using video clips as item prompts to measure teachers' knowledge of teaching mathematics. *Educational and Psychological Measurement, 68*(5), 845-861.

Lye, S. Y., & Koh, J. H. L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? *Computers in Human Behavior, 41*, 51-61.

Satratzemi, M., Xinogalos, S., Tsompanoudi, D., & Karamitopoulos, L. (2018). *Examining student performance and attitudes on distributed pair programming.* Hindawi.

Stahl, G. (2005). Group cognition in computer-assisted collaborative learning. *Journal of Computer Assisted Learning, 21*, 79-90.

Umapathy, K., & Ritzhaupt, A. D. (2017). A meta-analysis of pair-programming in computer programming courses: implications for educational practice. *ACM Transactions on Computing Education, 17*(4), 1-13.

Wang, S. L., & Hong, H. T. (2018). The roles of collective task value and collaborative behaviors in collaborative performance through collaborative creation in CSCL. *Educational Technology Research and Development, 66*(4), 937-953.

Wing, J. M. (2006). Computational Thinking. *Communications of the ACM, 49*(3), 33-35.

Wu, B., Hu, Y., Ruis, A. R., & Wang, M. (2019). Analysing computational thinking in collaborative programming: A quantitative ethnography approach. *Journal of Computer Assisted Learning, 35*(3), 421-434.

Yang, T. C., Chen, S. Y., & Hwang, G. J. (2015). The influences of a two-tier test strategy on student learning: a lag sequential analysis approach. *Computers & Education, 82*, 366-377.

Zhong, B., Wang, Q., & Chen, J. (2016). The impact of social factors on pair programming in a primary school. *Computers in Human Behavior, 64*, 423-431.

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China (19ZDA364; 61907038), and the Fundamental Research Funds for the Central Universities, China (2020QNA241).

Idea Improvement Processes Leading to High Learning Outcomes and the Development of Regulation in Collaboration

Anthony J Taiki Kawakubo, Jun Oshima, Ritsuko Oshima
kawakubo.taiki.15@shizuoka.ac.jp, joshima@inf.shizuoka.ac.jp, roshima@inf.shizuoka.ac.jp
Shizuoka University

Abstract: This study examined the relationship of idea improvement in knowledge-building practice with learning outcomes and the development of collaboration regulation. Questionnaires were administered to 70 university students engaged in project-based learning (PjBL) with a computer-supported collaborative learning (CSCL) environment for evaluating students' metacognitive knowledge and skills of regulation before and after their PjBL. Their final poster presentations were assessed as learning outcomes. To identify idea improvement, socio-semantic network analysis was applied to the CSCL discourse. Their discourse topics were identified through clustering analysis of the temporal changes in the degree centralities of word networks. Finally, multi-regression analyses were performed to develop models to predict learning outcomes and regulation development from discourse topics. Results revealed that the learning outcome was higher when learners engaged in the appropriate idea improvement and that learners' collaboration regulation develops through their engagement in collaborative monitoring.

Theoretical background and research purpose

Analysis of the idea improvement process: The development of a socio-semantic network analysis of discourse

The knowledge-building (kb) community has been a representative model of learning as knowledge-creation for decades in the learning sciences (Scardamalia & Bereiter, 2014). In kb practice, learners engage in sharing their ideas, making judgments of their idea promisingness, and further searching for new information to improve the ideas. For evaluating idea improvement, we need to clarify: (1) How a group's collective state of knowledge can be represented, and (2) how its dynamic change of state is represented over time (Oshima et al., 2012).

One direction of research in such development is Knowledge-Building Discourse Explorer (KBDeX) (Oshima et al., 2012). KBDeX can visualize the collective state of knowledge as a network of vocabulary in discourse, allowing us to view its dynamic changes over time through conversation turns as a unit of analysis, thereby aiding visualization by displaying how ideas are connected in learners' discourse. Since the first release of KBDeX in 2012, there have been several modifications of the algorithm to represent the collective state of knowledge in the discourse more accurately (e.g., Kawakubo et al., 2020). An issue is the computation to detect clusters of words in the vocabulary network. Kawakubo et al. (2020) developed an algorithm to detect discourse topics. They used temporal changes in the degree centralities (DCs) of words in discourse over a PjBL course in their clustering analysis. They assumed that students engaged in a variety of discourse topics related to their own ideas or others', and that the clustering analysis of DCs of all the (noun) words rather than keywords could more appropriately represent students' kb practices. They found that groups engaged in similar discourse topics, but high-learning-outcome groups were particularly involved in discourse topics regarding multiple ideas and warrants for their accepted ideas. In this study, we applied a modified algorithm that takes the temporality of the network into consideration to characterize idea improvements.

Development of collaboration regulation as a goal of instruction

In educational psychology, the mechanism of collaborative learning has been examined from the perspective of regulation. In a collaborative task, learners engage in three different layers of regulation: themselves, others, and the group as a whole (Hadwin et al., 2018). In the layer of self-regulated learning (SRL), learners regulate their own learning, not for individual performance but to contribute to group performance. In another layer, namely, co-regulated learning (CoRL), learners engage in regulation of others or by others. Learners consider ways in which their actions and interactions influence one another and the task by monitoring the task perception, goals, and standards of other group members. In the third layer, they collectively regulate their group cognition: This is socially shared regulation of learning (SSRL).

Järvenojä et al. (2013) developed the "Adaptive Instrument for Regulation of Emotions" (AIRE) to identify how learners are involved in the three layers of regulation during collaboration. Their questionnaire covers 14 socio-emotional challenges that learners might face in their collaborative learning, including items asking learners which layer of emotion regulation they engaged in, and found that individual, shared, and other forms of regulation were used to maintain group work when students encountered a socio-emotional challenge in a teacher training program at a university. Although many studies have examined the development of learners' regulation of collaboration from the perspective of the acquisition metaphor of learning, few studies have discussed the relation between the development of regulation and kb practices from the knowledge-creation perspective. In this study, therefore, we attempted to examine how learners develop their collaboration regulation through the idea improvement.

Research purpose
In this study, we examined the relationship of the idea improvement in kb practice with learners' development of collaboration regulation, as well as learning outcomes. To this end, we proposed a new algorithm for the socio-semantic network analysis to evaluate learners' idea improvement. We further administered a questionnaire to assess learners' metacognitive knowledge and skills of regulation in the pre- and post-test session so that we could examine its development. Finally, we developed models to predict learning outcomes and the development of regulation from their idea improvement using multi-regression analyses.

Study design

Study context
Seventy university students (18 groups consisting of 3 to 4) engaged in group work to develop their original happiness indices in their PjBL course over 15 weeks. Students tried to propose new perspectives to create their original happiness indices over a six-week period (weeks 9–14). In week 15, the final week, students had a poster session in which they discussed their ideas with the other groups in the form of poster presentations.

During weeks 9–14, students were instructed to report their idea improvement by writing down their progress in group activities and building their individual comments on their group notes each week in Knowledge Forum (Scardamalia & Bereiter, 2014), a computer-supported collaborative learning environment. In their group progress reports, they described the ideas they had considered, the criteria they had selected to evaluate their ideas, and how they had created and improved upon their ideas. In their individual notes, students reported their individual comments on the group activities by describing their thoughts on how to further improve their group ideas and what other ideas they might consider in the next week.

Data and analysis

Poster presentation as learning outcome
Groups' poster presentations were used as their learning outcomes. Four independent raters, including the authors, evaluated their poster presentations based on the following criteria, with five-point Likert scales: (1) appropriateness of the names of their indicators; (2) how well they described the unique natures of their proposed indicators; (3) how much evidence (data) was used to calculate their proposed indicators; (4) how well structured the data were to represent the nature of their proposed indicators; (5) how well their ranking of prefectures in Japan based on their indicators were explained; (6) how well their results were presented; and (7) how well they discussed their results of prefecture ranking. All the correlations across raters on their scores of the seven criteria were significant (rs = .53~.91, ps < .05). Scores by each rater were standardized, and average scores across raters were used as the groups' learning outcomes.

Questionnaire to assess students' regulation in collaboration
We asked students how they think of conflicts in groups in four typical contexts happening in collaboration and what they say to solve the conflicts in the questionnaire conducted in the first week (the pre-test) and a week after the final week (the post-test). Their thoughts on the conflicts were examined as their metacognitive knowledge about the conflict contexts and their discourse actions were examined as their skills to solve the conflict contexts. Two independent raters coded all the responses from the socio-cognitive and socio-emotional perspectives by the coding scheme from "no articulate understanding of the conflict in the scene" (score 1) to "recognition of the conflict based on appropriate understanding of both sides" (score 5) for metacognitive knowledge, and "no productive action was taken" (score 1) to "taking discourse actions to solve the conflict in the scene with appropriate understanding" (score 5) (Cohen's Kappa = .77). The disagreements were resolved by discussion. The development scores of each student's regulations in collaboration were calculated as the differences in pre-test and post-test scores. Group scores for development were calculated as the mean scores of group members.

Socio-semantic network analysis (SSNA) and clustering analysis to detect temporal discourse topics

To examine the idea improvement process in each group, we used all nouns to detect the discourse topics discussed during a jigsaw group activity. By using KBDeX, a socio-semantic network of vocabulary was created based on the co-occurrence of nouns. We calculated the *degree centrality coefficient*, a measure of the cohesiveness of a network structure, of each noun every time a new KF note was added, and examined the changes in its coefficient over time (Figure 1). Based on the datasets of temporal changes in coefficients, we further conducted a clustering analysis of nouns in each group with the Ward method.

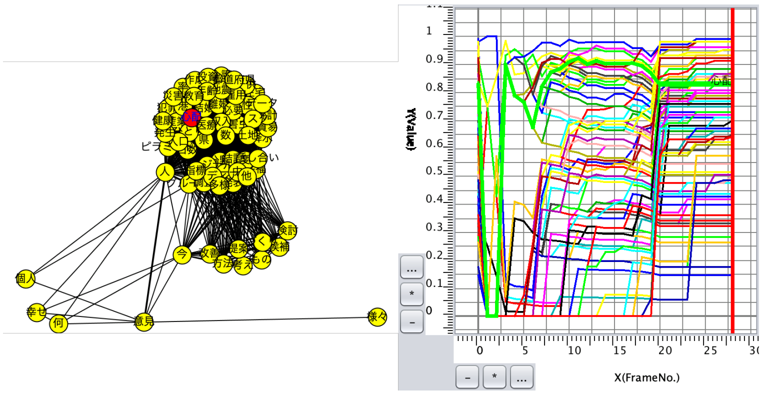


Figure 1. A socio-semantic network of vocabulary (left) and
temporal changes in the degree centrality coefficients of vocabulary (right).

Multi-regression analysis to develop models for predicting learning outcomes and regulation skills from temporal discourse topics

We further conducted stepwise multi-regression analyses to develop models for predicting students' learning outcome scores or regulation development scores from the scores of discourse topics. Standardized DCs of discourse topics were used as the discourse topic scores.

Results

Learning outcomes and development of regulation in collaboration

The standardized scores of poster presentations by 18 groups ranged from 34.79 to 66.37. The scores for the development of regulation in collaboration were .06 (SD = .07) for socio-cognitive metaknowledge of the conflicts, .18 (SD = .11) for socio-cognitive skill, .13 (SD = .19) for socio-emotional metaknowledge, and .16 (SD = .12) for socio-emotional skill.

Detection of discourse topics in the idea improvement

Clustering analyses of the nouns used in discourse based on temporal changes in DCs were performed to categorize all the nouns into discourse topics. In each group, nouns were categorized into seven topics: (1) *presented idea*, (2) *topic related to the presented idea*, (3) *rejected idea*, (4) *warrant*, (5) *reflection*, (6) *report*, and (7) *preparation of presentation*.

Predictive models of learning outcomes and development of regulation in collaboration

Stepwise multiple linear regression analyses were performed to develop models for predicting the learning outcome scores and four aspects of regulation development scores from the seven discourse topic metrics. For the learning outcome, discourse topics such as *rejected idea*, $t(13) = 2.09$, $p < .10$, *warrant*, $t(13) = 2.67$, $p < .05$, *preparation*, $t(13) = -2.34$, $p < .05$, and *topic related to the presented idea*, $t(13) = -2.14$, $p < .10$, were included as significant explaining variables to predict the learning outcome in the final model, $F(4, 13) = 4.48$, $p < .05$, adjusted $R^2 = .45$. For the development of socio-cognitive knowledge, *topic related to the presented idea*, $t(14) = 3.16$, $p < .01$, *preparation*, $t(14) = 1.75$, $p = .10$, and *reflection*, $t(14) = -1.693$, $p = .11$, were included in the final model, $F(3, 14) = 7.60$, $p < .01$, adjusted $R^2 = .54$. For the development of socio-cognitive skills, only *report*, $t(16) = -2.89$, $p = .11$, was included in the final model, $F(1, 16) = 8.35$, $p = .11$, adjusted $R^2 = .30$. For the development of socio-emotional knowledge, no topics were included and a predictive model was not created. For the development of socio-emotional skills, only *report*, $t(16) = -1.894$, $p < .10$, was included in the final model, $F(1, 16) = 3.59$, $p < .10$, adjusted $R^2 = .13$.

Discussion

How the idea improvement process was related to the learning outcome

Results revealed that three discourse topics were significantly related to the learning outcome. Having *rejected ideas* means that students discussed multiple ideas and made a judgment of idea promisingness. As discussed in a previous study (Kawakubo et al., 2020), having multiple ideas is the critical condition of idea improvement. It may be easier for learners to judge the promisingness of their ideas if they develop and compare multiple ones. Learners should not only have multiple ideas but also engage in the appropriate process of improving their selected ideas. The significant relation of the *warrant* to the learning outcome suggests that students engaged in discourse on how to make their ideas more robust and convincing by examining rationales and collecting evidence for their claims. Finally, the discourse topic of preparation was negatively correlated with high learning outcomes. Our further examination of the discourse revealed that high learning-outcome groups engaged in reporting a variety of issue for preparing their presentations and that the degree centralities of words in the network were consequently lower than those in low learning-outcome groups.

How the idea improvement process was related to the development of regulation in collaboration

First, in the models of both socio-cognitive and socio-emotional regulatory skills, only the *report* was found to be negatively correlated. The result may suggest that learners could develop their regulation skills when they had more fluent opportunities to discuss their group work.

Second, in the model of metacognitive knowledge of socio-cognitive conflicts, *topic related to the presented idea*, *preparation*, and *reflection*, were found to be critical to predicting development. *Topic related to the presented idea* was a cluster of words representing a variety of vocabularies discussed around the idea. *Preparation* was a cluster of words used in reporting their preparations for their final presentations. *Reflection* was a cluster of words used in individual reflection notes on their group progress, and the discourse topic was negatively correlated to the regulation development. The results suggest that learners could develop their metacognitive knowledge of socio-cognitive conflicts if they engaged in intensive and productive group work as an object for them to reflect on. In other words, engagement in good group work and intentionally rich reflection on the work are necessary conditions for learners to successfully develop their metacognitive knowledge of socio-cognitive conflicts.

References

Hadwin, A., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Educational psychology handbook series. Handbook of self-regulation of learning and performance* (pp. 83–106). Routledge/Taylor & Francis Group.

Järvenoja, H., Volet, S., & Järvelä, S. (2013). Regulation of emotions in socially challenging learning situations: An instrument to measure the adaptive and social nature of the regulation process. *Educational Psychology*, *33*(1), 31–58.

Kawakubo, A. J T., Oshima, J., & Oshima, R. (2020). Differences in idea improvement processes between high and low learning-outcome groups in project-based learning. In Gresalfi, M. and Horn, I. S. (Eds.), *The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences (ICLS) 2020, Volume 1* (pp. 505–508). Nashville, Tennessee: International Society of the Learning Sciences.

Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse. *Educational technology research and development*, *60*(5), 903–921.

Scardamalia, M. & Bereiter, C. (2014) Knowledge building and knowledge creation: Theory, pedagogy, and technology. In Sawyer, K. (Ed.) *The Cambridge handbook of the learning sciences: Second edition* (pp. 397–417). Cambridge University Press.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 16H0187, 18K18639 and 20KK0046.

Operationally Defining Turn-taking in Collaborative Online Documents

Mik Fanguy, School of Humanities and Social Sciences, Korea Advanced Institute of Science and Technology, mik@kaist.edu
Jaemyong Chang, Georgia Institute of Technology, jchang376@gatech.edu

Abstract: Turn-taking is an important aspect of collaboration, but turns are difficult to operationally define within online collaborative documents. For example, turns can be taken by multiple group members simultaneously, and in some cases, their turns may seem to meld together into a single written contribution. Thus far, no clear definition for a turn has been provided in this context. This paper proposes that a turn be defined as an instance of writing that begins with the first character contributed by a member and ends with the last character of an uninterrupted string of text contributed by the same member or by the end of the document. A computer system was developed in order to automatically calculate the number of turns within a text. This turn-taking data will be of use to practitioners who strive to encourage high levels of written interaction among members of collaborative writing groups.

Introduction

Interactivity is an important facet of both communication collaborative work. When engaging in a discussion or group work, participants should interact with one another, contributing with an awareness of and responsiveness to the contributions of their group members. In a number of contexts, turn-taking can provide a measure of interactivity in communication and collaboration among group members (McKinlay et al., 1993). Most of the literature on turn-taking focuses on spoken conversation, where Sacks's (1992) three maxims for conversations apply: 1) speakers contribute one at a time, 2) speakers contribute in non-overlapping turns, and 3) speakers contribute in a turn-wise manner. Such conventions make turns easily identifiable in spoken conversation. However, in online collaborative writing contexts, these first two maxims are often irrelevant, as contribution tends to occur in a free-for-all manner (Gibson, 2009). Multiple contributions may be added simultaneously, and some turns may seem to meld together into a single contribution. Because of this lack of clear boundaries between turns, the best of our knowledge, turn-taking has never been operationally defined in the context of writing online collaborative documents. Therefore, this paper provides a simple operational definition for a turn within an online collaborative document and describes a custom computer system that can automatically count turns within such documents according to the proposed definition. This operational definition and counting technique offer researchers a clear way to assess the amount of interaction taking place among group members when collaboratively writing documents online.

Literature review

Compared to writing by oneself, collaborative writing offers a number of advantages to students, including exposure to various viewpoints, development of collective knowledge, and improved writing quality (Beck, 1993; Ede et al., 1990). In addition, writing collaboratively provides a number of social and cognitive benefits (Rice & Huguley, 1994; Sullivan, 1994), as the social processes learners engage in can help them to improve their abilities to think critically, reflect, and exchange ideas (Picciano, 2002; Scardamalia & Bereiter, 2003; Stahl, 2006). Various technological tools can be used by learners to write collaboratively online, and these include wikis (Aydin & Yildiz, 2014), online word processors (Kessler et al., 2012), and blogs (Sun & Chang, 2012). Using such internet-based tools enables users to contribute writing from any location at any time. Furthermore, group members do not need to be logged in simultaneously to a shared online document in order to collaborate and may work at their convenience. As a result, learners tend to contribute asynchronously to an online collaborative document (Weng & Gennari, 2004).

When writing collaboratively, learners may choose not to write their entire contribution in one sitting but may instead write in an interactive or turn-taking manner, responding to the written contributions, edits, and comments of other members. Online word processors can increase the interactivity of collaboration among members, as such document interfaces allow collaborators to view the contributions of other members, thereby enabling them to write, edit, and comment with greater frequency than was previously possible (Yim et al., 2014). Learners' cognizance and planning of turn-taking has been shown to correlate positively with the quality of the writing they produce (Erkens et al., 2005).

However, identifying turns within a collaborative document can be challenging because written contributions can occur simultaneously and may seem to blend together into a single contribution, lacking discrete boundaries. Prior work has operationalized various aspects of collaboration in online documents, such as the evenness of contribution among group members, the amount of edits each member makes to their own writing and the writing of other members, and the number of times each member logs in to an online document in order to contribute writing (Wang et al., 2015). However, due to the difficulty of defining turn-taking within collaborative documents, no clear definition has been proposed so far.

Proposed definition of a turn in online collaborative documents

The present paper proposes a definition for a turn within online collaborative documents as follows. A turn is defined as an unbroken string of writing contributed by a group member to a shared document that other group members have permission to read, edit, and respond to, regardless of whether any fellow group member reads, edits, or responds to the contribution. However, the proposed definition does not count embedded comments, oral feedback, oral discussion, or any other forms of backchannel communication as turns. Instead, the proposed definition only considers the contributions by each group member to the final written product. Operationally defined, a turn begins with the first character contributed by a member and ends with the last character of an uninterrupted string of text contributed by the same member or when the end of the document is reached. This definition enables the identification of turns that are taken simultaneously and/or meld together. To do so, one would simply count the changes in authorship within a given online collaborative document. A higher frequency of authorship changes indicates a greater degree of interaction among group members in composing the document, and a lower frequency of authorship changes indicates a lesser degree of interaction among group members in composing the document. To facilitate the identification of authorship changes, practitioners can require students to select and use a unique font color, so that each member's contribution is clearly distinguishable from that of his/her fellow contributors. In this way, each change of font color within a document would represent the beginning of a new turn.

A group of three students collaborated in writing a set of notes using Google Documents, a widely-used online word processor, and a screenshot of their work is shown in Fig. 1. Each member of the group distinguished his/her work by writing in a unique color (brown, green, and blue). According to the proposed definition, a total of ten turns was taken by the group members who composed this section of text shown in the figure. The end of each of the ten turns is marked by a change in authorship, signified here by a change in font color, or by the end of the document.

Notes on the use of quotation marks
- Single and double quotes are used differently in Korean and English
- American English use double quote, British english use single quote
(American English use almost always "double quotes", hardly ever single quotes.)
-exception : "~~~'`~~~~'~~~~",   '~~~~"~~~~~"~~~~~' > In double quotes, in the case of we have to use quotes again, we can use single quotes.
Example: smith explained, " we found ~ 'slips' during~ ." -> It is not Smith who said about slip. ("",'' Vice versa) In British English, it is different
Rules when quoting someone else's words directly
-Capitalize the first word when quoting complete sentence; however, do not capitalize the first word of a part of a sentence
-If a quote is broken up, do not capitalize the second part of the quotation
-Do not use quotation marks for indirect quotes(including paraphrasing)
-In a quote, use brackets to add words and ellipsis to remove them

Figure 1. Sample notes taken collaboratively by a group of students using Google Docs.

Development of an automatic turn-counting computer system

In order to automatically count turns taken within Google Documents, we developed a computer system, Collab_Notetaking (https://github.com/porkchop-jim/Collab_Notetaking), that downloads multiple Google Docs from folders and sub-folders in Google Drive for the purpose of counting the number of words and the number of turns each contributor takes. Collab_Notetaking stores the data in a local database file. The system counts the number of words that each of the contributors wrote. Each contributor is assigned a font color at the beginning of the document. The black default font is designated for instructors to give instructions or feedback to the students.

Architecture of the collab_notetaking turn-counting System

The system is written in the Python (https://www.python.org) programming language and uses various Python libraries that are publicly available. The system consists of the following files: collabo_db.py, g_drive_list_folders.py, student_revisions.py, and font_counter.py.

The user must run the collabo_db.py file to create a database file named collabo.db. The user must give permission for the system to interact with the Google Drive API through Google's security setup procedure and obtain Google's Client ID and the credentials.json json (JSON, JavaScript Object Notation) file. The credentials.json file should be stored in the same folder as the font_counter.py file. The user must also obtain the parent folder's id by running g_drive_list_folders.py file. The folder ID can be obtained from the output. The user must put the folder id into the font_counter.py file where stated.

When the font_counter.py runs for the first time, the web browser will be opened for further authentication. Two authentication files will be created to eliminate the need for future authentications. Once the authentication is finished, the system searches through folders and sub-folders to download the Google Doc files as a .docx (.docx, MS Word document file) file to keep the formatting intact. After each download completion, the system reads the file and counts the number of changes in font color within the document, which represent changes in authorship among group members. This tallied number serves as the turn-taking variable.

Discussion and conclusion

The present paper has proposed an operational definition for turn-taking in the context of online collaborative documents and has introduced a computer system that automatically counts the number of turns taken by each group member. Measuring the frequency of turn-taking within a group in composing a collaborative document can provide a useful metric to practitioners who strive to encourage high levels of written interaction among members of collaborative writing groups. Existing visualization tools for online collaborative documents operationalize a number of facets of collaboration in writing (Wang et al., 2015) but have thus far ignored turn-taking, resulting in a potential blind spot with regard to the interactivity of the collaborative writing processes students engage in. Therefore, as a future work, the user interface of the proposed turn-counting computer system should be improved in order to provide a visualization tool, perhaps as an add-on to the Google Docs platform, in order to supplement existing collaborative writing visualization tools. Practitioners and researchers would then be able to consider turn-taking data in context with other relevant collaboration data in order to gain deeper insights into the processes learners engage in when composing documents together online. Furthermore, a visualization of turn-taking could also provide benefit to learners participating in collaborative writing groups, as prior research has shown that learners' awareness and planning of turn-taking is positively correlated with the quality of the document that they produce (Erkens et al., 2005).

Although the turn operationalization method proposed here provides a valuable measure of written interaction at the group level, it provides a potentially problematic measure of turn-taking at the level of the individual. The problem stems from the fact that the proposed definition allows for the number of turns taken by an individual to be increased by other group members without his/her knowledge or effort. This may occur when a large chunk of text contributed by an individual early on in the writing process gets split up into smaller chunks by fellow group members who intersperse their own writing (turns) within the original section of text. In such a case, the initial author is credited with having taken multiple turns, though he/she may have written the entire contribution in a single session. While such an example does not resonate with conventional notions of turn-taking from other forms of communication such as spoken conversation, one could view such unintended turns as legitimate since the original contributor was attempting to interact and achieved interaction by inducing a written contribution to the document by another member. We recognize that the proposed definition of a turn is controversial and imperfect, and for this reason, we recommend practitioners and researchers use it to measure levels of written interaction at the group level rather than literal turns taken at the individual level. When used in this way, such a metric will be useful to researchers in better understanding the processes that co-authors engage in when collaborating.

References

Aydin, Z., & Yildiz, S. (2014). Using wikis to promote collaborative EFL writing. Language Learning & Technology, 18(1), 160–180.

Beck, E. (Ed.). (1993). A Survey of Experiences of Collaborative Writing. Berlin, Germany: Springer Verlag. doi:10.1007/978-1-4471-2007-0_6

Ede, L. S., Ede, L., & Lunsford, A. A. (1990). Singular Texts/Plural Authors: Perspectives on Collaborative Writing. SIU Press.

Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. Computers in Human Behavior, 21(3), 463-486. doi:10.1016/j.chb.2004.10.038

Gibson, W. (2009a). Intercultural communication online: Conversation analysis and the investigation of asynchronous written discourse. Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 10(1): article 49. doi:10.17169/fqs-10.1.1253

Kessler, G., Bikowski, D., & Boggs, J. (2012). Collaborative writing among second language learners in academic web-based projects. Language Learning & Technology, 16(1), 91–109.

McKinlay, A., Procter, R., Masting, O., Woodburn, R., & Arnott, J. (1993). A study of turn-taking in a computer-supported group task. Proceedings of HCI 93: People and Computers VIII, 383-383.

Picciano, A.G. (2002). Beyond student perceptions: issues of interaction, presence and performance in an online course. Journal of Asynchronous Learning Networks, 6, 21–40.

Rice, R. P., & Huguley, J. T. (1994). Describing collaborative forms: A profile of the team-writing process. IEEE Transactions on Professional Communication, 37(3), 163-170. doi:10.1109/47.317482

Sacks, Harvey (1992). Lectures on Conversation, Vols. 1 & 2. London: Blackwell.

Scardamalia, M., & Bereiter, C. (2003) Knowledge building. In Encyclopedia of Education, (ed. J.W. Guthrie), pp. 697–706. Macmillan Reference, New York.

Stahl, G. (2006) Group Cognition: Computer Support for Building Collaborative Knowledge. MIT Press, Cambridge, MA.

Sullivan, P. (1994). Computer technology and collaborative learning. New Directions for Teaching and Learning, 59, 59-67.

Sun, Y. C., & Chang, Y. J. (2012). Blogging to learn: Becoming EFL academic writers through collaborative dialogues. Language Learning & Technology, 16(1), 43-61.

Wang, D., Olson, J. S., Zhang, J., Nguyen, T., & Olson, G. M. (2015). DocuViz: visualizing collaborative writing. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 1865-1874). ACM. doi:10.1145/2702123.2702517

Weng, C., & Gennari, J. H. (2004, November). Asynchronous collaborative writing through annotations. In Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (pp. 578-581). ACM. doi:10.1145/1031607.1031705

Yim, S., Warschauer, M., Zheng, B., & Lawrence, J. F. (2014). Cloud-based collaborative writing and the common core standards. Journal of Adolescent and Adult Literacy, 58, 3, 243–254. doi:10.1002/jaal.345

International Society of the Learning Sciences

Embodied Transmission of Ideas: Collaborative Construction of Geometry Content and Mathematical Thinking

Veena Kirankumar, Hanall Sung, Michael Swart, Doy Kim, Fangli Xia, Oh Hoon Kwon, Mitchell Nathan
kirankumar@wisc.edu, hanall.sung@, mswart@, doy.kim@, fxia24@, kwon@math.wisc.edu, mnathan@
University of Wisconsin – Madison
Candace Walkington, Southern Methodist University, Cwalkington@mail.smu.edu

Abstract: This study looks at how students embody their ideas about geometry conjectures and how those ideas travel within and between student groups. In one classroom of a Title 1 high school, students participated in a three-part program in which they: (1) played *The Hidden Village*, a motion-capture video game where they assess the veracity of geometric conjectures (i.e., if it is always true or ever false) while their intuitions, insights, and rationales (including their gestures) are video recorded, (2) designed their own directed actions (i.e., a sequence of movements that represents a body-based interpretation of the structure and transformation of a spatial configuration), and (3) re-played the game with a mixture of previous conjectures combined with the conjectures designed by their peers. Multiple cases revealed ways that simulated enactment and collaborative construction can convey mathematical ideas.

Keywords: Embodiment, Geometry, Collaborative Construction, Transfer

Introduction

Students played a motion-capture video game, *The Hidden Village,* and then were provided opportunities to make new content for the game. Students were invited to think, act, and talk through the ways that their bodies could *represent* geometric objects in the conjectures--statements that are provable false or true. We hypothesize that these embodied sequences, called *directed actions*, can foster mathematical insights crucial for students' understanding. Students designed their own directed actions using their bodies to express their conceptualizations of geometric conjectures to themselves and their fellow group members in the context of the game. Analysis of students' gesture production, simulating the actions of geometric transformation, demonstrated how students explore and explain their thinking.

Theoretical background

Studies have shown that mathematics can be learned through action-based interventions (Abrahamson & Sánchez-García, 2016). *The Hidden Village* (THV; Swart et al., 2020) is an educational video game. It draws on the theory of *Gesture as Simulated Action* (GSA; Hostetter & Alibali, 2019), which asserts that gestures activate perceptual-motor processes in the brain when co-articulated with speech or thought. These sensorimotor experiences can induce cognitive states through the process of *Action-Cognition Transduction* (ACT; Nathan, 2017). From this, Nathan and Walkington (2017) developed the *Grounded and Embodied Cognition* (GEC) framework, which proposed that directing players' bodily movements (via directed actions) will complement learners' verbal expressions of mathematical reasoning.

An embodied theory of transfer (Alibali & Nathan, *in press*) posits that concepts are ultimately represented by the actions, gestures, and other body-based resources embedded in various physical and social settings, like collaborative game play. We call this form of embodied transfer "travel." By prompting players to explain their answers, THV primes players' production of *dynamic depictive gestures* that mentally and physically simulate transformations of mathematical objects through multiple states (Garcia & Infante, 2012) that can "travel" to other players.

In the current study, we explored two research questions: (RQ1) How does a student group designing new game content develop their mathematical ideas and create their own directed actions intended for others to play? (RQ2) How does the intention of the original group's mathematical ideas "travel" to other student groups through subsequent game play, and show up as the embodied transfer of those ideas in other groups' gesture and speech? Thus, we investigate how student groups created directed actions for geometry conjectures and formed their ideas about geometric transformations, hypothesizing that students' embodied mathematical ideas "travel" through player-generated content in the form of direct-action movements.

 © ISLS

Methods

Materials

The hidden village (THV) game
THV is a 3D motion-capture collaborative video game that offers an immersive embodied geometry curriculum in which each player emulates in-game avatar's movements and then reasons about geometry conjectures to prove whether it is either *false* (F) or *always true* (T).

The hidden village (THV) conjecture editor
The THV Conjecture Editor enables students to create new movement-based game content. Students add new conjectures and then design mathematically relevant directed actions by manipulating the sequences of poses of the avatar (Figure 1). Using the Pose Editor, student groups collaboratively generate 2-3 poses (starting, intermediate, and target pose; see middle panel of Figure 1) to create directed actions for each conjecture. Once poses have been designed, players can preview the movements as an animation. Once completed, user-generated actions are stored in the online database of THV and accessible to any other users to access and play.

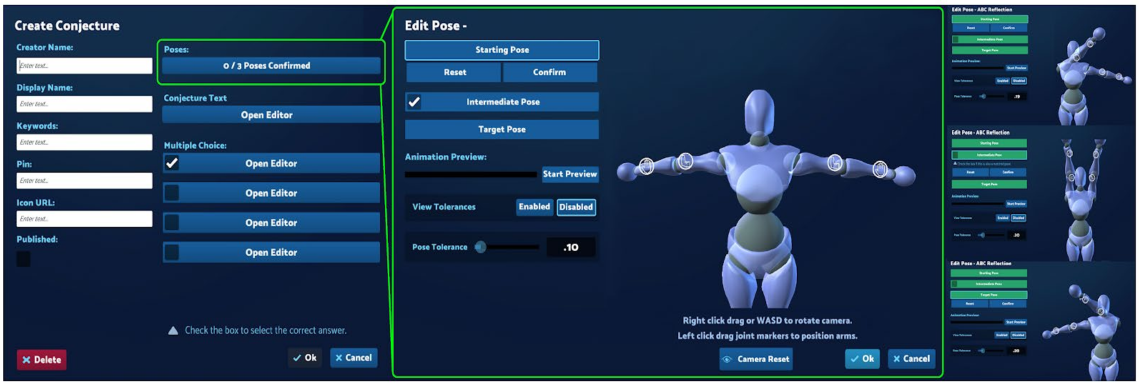


Figure 1. The THV Conjecture Editor and THV Pose Editor (for creating directed actions) with an example of a directed action sequence (far right).

Participants

In this study, 12 students in a Title I high school in the midwestern United States participated in a three-day embodied mathematics curriculum focused on geometric thinking. Students were randomly assigned to groups of three or four. This paper focuses on two of the student groups.

Procedure

The three-day curriculum extended over three class periods over three successive weeks: (Day 1) group members take turns playing six conjectures in THV; (Day 2) student groups collaboratively construct their own directed actions for a new conjecture, and (Day 3) student members take turns playing a new THV curriculum (eight conjectures total; three repeated from Day 1, three designed by student groups on Day 2, and two new (transfer) conjectures). The *in situ* curriculum was administered during normal class time and students' group gameplay data (including student's (1) intuitions, (2) insights, and (3) explanations of conjectures) and co-design activities were video recorded, transcribed and coded by researchers.

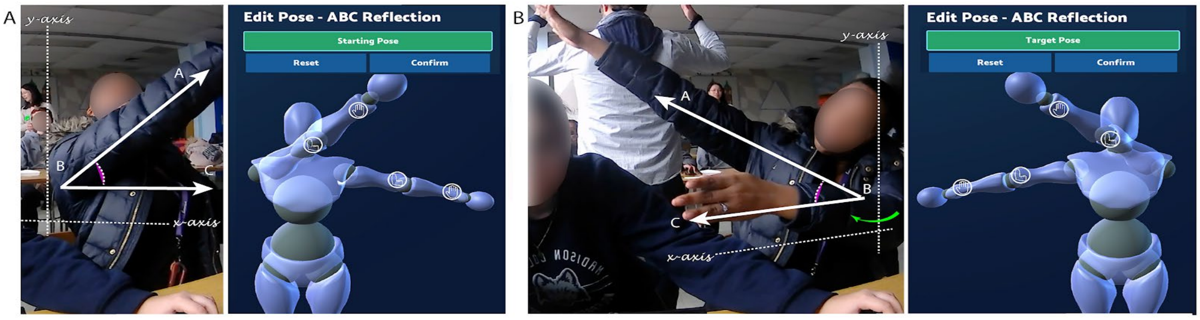
Results

Within-group analysis

To understand learning processes in the co-design activity, researchers analyzed students' (Group 1) collaborative multimodal interactions during group discussions, including both gestural and verbal communication (RQ1). Students in Group 1 co-constructed directed actions for their chosen conjecture. Figure 2 is a photo-illustrated transcript of Group 1's discussion of their mathematical ideas (RQ1).

In the course of designing their directed actions, students used multiple dynamic depictive gestures (i.e., action-speech pairings, Nathan, 2017) while deliberating which directed actions would best assist players to grasp the geometric relations relevant to proving their conjecture, the *ABC Reflection* (which is false):

Given three points A, B, and C, and their reflected images about a line, A', B', and C', then ∠ABC and ∠A'B'C' are not equal.



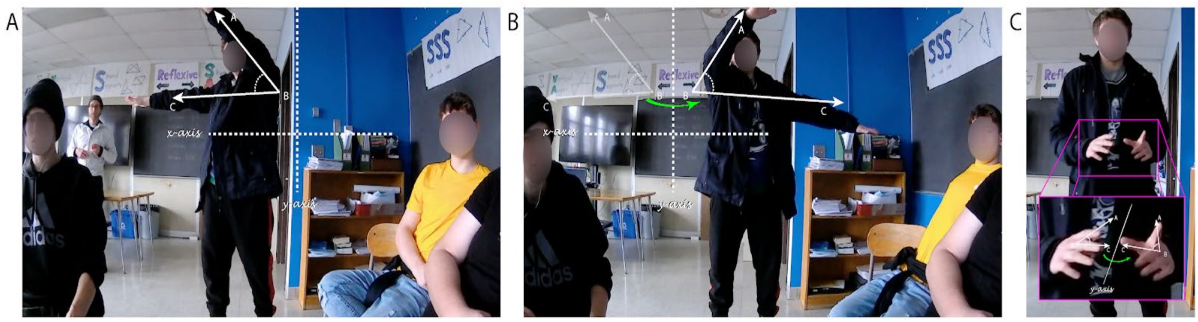
Transcript #1: (*N.B. S1 indicates Student #1;* brackets [...] indicate gestures.)
[1] S1: Oh, wait. This is not the starting pose. Is that the starting pose? [Uses arms to make ∠ABC on the left
[2] side of the body] We are going like, this is the angle [shifted arms directly to the right side of her
[3] body by performing a reflection across the body vertical axis]... Boom💥! That's the angle!

Figure 2. For *ABC Reflection* conjecture, Student 1 in Group 1 embodies the starting pose (also shown as designed in THV Pose Editor, panel A) and S1 performs the entire directed action, finishing on the target pose.

Figure 2 indicates the starting and target poses (see panel B) the student group used for the *ABC Reflection* conjecture. Narrating their actions as they reflect the angle from the right side of the body to the left, S1 embodied the idea of "using your body as the midline" through this directed action. In finalizing these directed actions, the group members solidified their understanding of the conceptual difference between reflection and rotation in the process of designing their pose sequences in the THV pose editor.

Between-group analysis

On Day 3, students played THV with a mixture of conjectures from Day 1, conjectures designed by their peers on Day 2, and two previously unseen conjectures. Students in Group 4 played the *ABC Reflection* conjecture as designed by Group 1. One player per group performed the directed actions prompted within THV, while the other group members observed. To track how Group 1's embodied mathematical ideas traveled to other groups (RQ2), researchers analyzed students' gestures and discourse.



Transcript #2: (*N.B. S2 indicates Student #2;* brackets [...] indicate gestures.)
[1] S2: False. Because it can be proportionally the same, have the same angles [using hands to make an angle]
[2] while being in different locations. [S2 then, selects the correct answer from the multiple-choice options]

Figure 3. Student 2 in Group 4 performs the directed actions (panels A & B) for *ABC Reflection*. In panel C, S2 provides their intuition (i.e., T or F) and rationale, using their hands to represent the reflection of ∠ABC.

After performing the directed actions during game play (see Figure 3), Student 2 (Transcript #2) states their intuition ("False"). S2 provides their rationale (Lines [1-2]) with spontaneous gestures (panel C). In the process of proofing the conjecture, S2's spontaneous gesture demonstrates an embodied conceptualization of the ∠ABC that results from the transformation. In effect, this truncated gesture complements S2's verbal rationale and extends Group 1's original idea for embodying the reflection of an angle over an axis.

Discussion

This study demonstrated instances of how mathematical ideas "travel" through embodied actions. THV served as a vehicle to reify geometric relations as movements of an avatar. Students created content that coupled geometric conjectures with movements intended to help players to embody these mathematical ideas. We found that students used the posable avatar as a way to explore embodied ways of reasoning and then share those ideas through subsequent game play. Performing these directed actions facilitated new players' mathematical intuitions and helped them articulate their justifications for transformational proofs.

Within Group 1, students communicated their ideas to each other about the *ABC Reflection* conjecture through their discussion and design of their directed actions. By embodying the geometric transformation in the conjecture, student's garnered insights about angles and axes in the course of reflecting the angle across the y-axis. Developing these actions enabled them to work through any misconceptions about reflection and contributed to Student 1's reconsideration of how to enact a directed action that more accurately depicted the reflection of ∠ABC. Through collaborative co-construction, this group exemplified how embodying mathematical thinking travels within a group as a design team and helped finalize their directed actions for the conjecture.

Between Group 4 and Group 1, the directed actions in *ABC Reflection* conjecture demonstrated how embodied mathematical ideas traveled successfully. The geometric transformation depicted by the in-game directed actions helped Student 2 interpret and explain the concept of reflecting an angle over an axis. Moreover, after performing these gross-motor movements with their arms, Student 2 generated a truncated spontaneous gesture using only their hands, a type of marking (see Kirsh, 2010) to represent the outcome of the geometric transformation. Nathan et al.'s (2017) *Grounded Embodied Cognition* framework contends that the directed actions primed the sensorimotor stimulation (i.e., feedforward and feedback) that preceded Student 2's gestural reaffirmation that the reflected angle across the y-axis was indeed congruent.

These case studies identify some of the promises of an embodied mathematics curriculum. Directed actions are a malleable factor that can scaffold cognition and produce historical traces that can give rise to spontaneous gestures and task-relevant speech in support of successful mathematical reasoning and proof.

References

Abrahamson, D., & Sánchez-García, R. (2016). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences, 25*(2), 203–239.

Garcia, N., & Infante, N. E. (2012). Gestures as facilitators to proficient mental modelers. In L.R. Van Zoest, J.-J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 289–295). Kalamazoo, MI:WMU.

Hostetter, A. B., & Alibali, M. W. (2019). Gesture as simulated action: Revisiting the framework. *Psychonomic Bulletin & Review, 26*(3), 721-752.

Kirsh, D. (2010). Thinking with the body. In *Proceedings of the Annual Meeting of the Cognitive Science Society, 32.*

Nathan, M. J. (2017). One function of gesture is to make new ideas. In R. B. Church, M. W. Alibali, & S. D. Kelly (Eds.), *Why Gesture?* (pp. 175–196). John Benjamins Publishing Company.

Nathan, M. J. & Alibali, M. W. (in press). An embodied theory of transfer of mathematical learning. In Charles Hohensee and Joanne Lobato (Eds.) *Transfer of Learning: Progressive Perspectives for Mathematics Education and Related Fields.* Springer.

Nathan, M. J., & Walkington, C. (2017). Grounded and embodied mathematical cognition: Promoting mathematical insight and proof using action and language. *Cognitive Research: Principles and Implications, 2*(1), 9.

Swart, M.I., Schenck, K., Xia, F., Kwon, O. H., Nathan, M. J., Vinsonhaler, R., & Walkington, C. (2020). Grounded and embodied mathematical cognition for intuition and proof playing a motion-capture video game. In Ilana Horn & Melissa Gressalfi (Eds.), (pp. 175-182) *Proceedings of the 2020 International Conference of the Learning Sciences*. Nashville, TN.

Acknowledgments

Research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A160020 to the University of Wisconsin-Madison. The opinions expressed are those of the authors and do not represent views of the IES or the U.S. Dept. of Ed.

Examining Collaborative Support for Block-Based Programming with Upper Elementary Students

Jennifer Tsan, University of Chicago, jennifertsan@uchicago.edu
Jessica Vandenberg, Collin Lynch, Eric Wiebe
jvanden2@ncsu.edu, cflynch@ncsu.edu, wiebe@ncsu.edu
North Carolina State University
Kristy Elizabeth Boyer, University of Florida, keboyer@ufl.edu

Abstract: Elementary students' pair programming relationships can be imbalanced in terms of talking and driving time. We investigated the balance of talk distribution between 4th and 5th grade students while they pair program. Students in our study either used a system that supported their switching and talking or the original block-based programming environment. Qualitative examination suggests that the students who received reminders had more balanced talk distribution than students who did not. We investigated this further by examining case studies of four pairs of students. We found that the two more balanced pairs often agreed with each other in regards to the positioning of their roles. The two pairs who were less balanced each had a student who disengaged from the task in their sessions. These findings suggest ways in which support for collaboration may foster young learners' interactions.

Introduction

Collaborative learning is a complex process that involves co-constructing knowledge via discussion between partners (Roschelle & Teasley, 1995). The amount of time students spend actively communicating is correlated with the success of their collaborative learning process (Barron, 2003). Important elements of the discussion involve the students' dialogue-act and social position (van Langenhove & Harré, 1999). Students do not always have balanced relationships when they are learning computer science (CS) concepts collaboratively (Deitrick et al., 2016; Shah et al., 2014). An unbalanced relationship suggests that there is a difference in power and status between students (Shah et al., 2014). This could be a result of how they perceive each other's ability in a subject, such as CS (Yamakawa et al., 2009). In short, unbalanced relationships can be detrimental to the collaborative learning process.

We study pair programming (Williams et al., 2000) where students take on roles as a *driver* (who controls the computer) or *navigator* (who directs the driver, plans ahead, and looks for mistakes). The students switch partway through to ensure both students experience both roles. We have previously shown that imbalances can arise in pair programming with elementary students, even when the instructor prompts students to switch roles (Tsan et al., 2018). We developed an intervention to remind students when to switch and to talk to their partners.

In this current study, we explore how the balance between upper elementary students varied when they received talking and switching reminders through a system called SuCCESs (Support for Collaborative Coding with Elementary Students) compared to students who did not receive the reminders. Our research questions are:

RQ1) *How do students respond to adaptive support features in a collaborative programming environment?*

RQ2) *To what extent do the adaptive support features support the students' talk balance?*

We compared students' interactions and talking balance between students in two pair programming studies. Pairs that received the SuCCESs reminders had more balanced talk times than pairs that did not. Additionally, the pairs that were more balanced seemed to agree on their roles more than students that were unbalanced. The pairs that were less balanced were also pairs with at least one distracted student. These results suggest that SuCCESs may be beneficial to students but more support is needed for the more unbalanced pairs.

Theoretical framework

Our work focuses on the balance between students collaborating in pairs while learning programming. We view our work through the lens of positioning theory (van Langenhove & Harré, 1999). Positioning explores how, over the course of a discussion, students espouse certain roles and related discursive practices, but, roles can shift. Students position themselves and others in accordance with contextual demands the conversation requires. We aim to explore the discursive tension between the students' assigned programming role, how they enacted that role, and how they transitioned into their next role. Positioning theory has been used to analyze the balance of students learning CS concepts in elementary (Shah et al. 2014) and high school (Deitrick et al., 2016).

Collaborative coding interface

In this study, students interacted with a novel set of features within NetsBlox, a block-based programming environment that allows users to create programs by dragging and dropping blocks into a workspace (Broll & Ledeczi, 2017). We implemented two features that make up the SuCCESs system: a switching-reminder that tells the students when to switch roles and a talking reminder that tells the students to talk to their partner. We piloted the features with 4th and 5th grade students and conducted a round of iterative refinement.

Method

Participants. For all participants reported in the paper, we obtained parental consent and student assent within an IRB-approved protocol. Participants in the first study (study 1) were 16 5th grade academically or intellectually gifted students (9 to 11 years old) from a suburban school in the Southeastern United States; 15 participated in the interventions. The intervention activities were taught by one of the authors. The demographics of the school were 54% white, 22% Hispanic, 18% African American, 4% Multi racial, 1% Asian, 1% other.

The participants in the second study (study 2) were 61 4th and 5th grade (8 to 11 years old) students in an urban elementary school in the Southeastern United States. The students were taught by one teacher during the programming lessons. The demographics of the school were 75% white, 10% African American, 6.5% Hispanic, 5% Multiracial, 2% Asian, 1% other.

Procedure. The students in both studies participated in the same set of CS learning tasks that were piloted and refined in previous studies (Zakaria et al., 2019). The hour-long lessons covered algorithms, input, conditionals, loops, and broadcasting. All students used pair programming for all of the activities. The students used versions of NetsBlox without (study 1) or with (study 2) SuCCESs. Each pair in study 2 received a switching reminder and about half the pairs received a talking reminder.

Data collection. During both five week studies, we collected audio, video, and screen recording data of their pair programming process. In order to fully explore the two research questions, student talk was transcribed, with the video and screen recording providing contextual clarity to the transcription process. Due to the intensive nature of annotating talk and video data, we sampled 24 pairs from our full dataset, 12 from each study. We sampled about 40 minutes of audio and video from each pair, eliminating recordings that were not clear enough to transcribe.

Analysis. For RQ1, we annotated the student behaviors when the reminders appeared. We developed a coding scheme to formally analyze how the students reacted to the reminders. First, we coded for verbal acknowledgement of the talking reminder (*acknowledged* and *did not acknowledge*). Then for the switching reminder, we labeled whether they switched (*switched*, *did not switch*, *willingly switched later*, or *teacher had them switch*). For RQ2, we used turns of talk as the unit of analysis to analyze the balance of talk between students.

Results

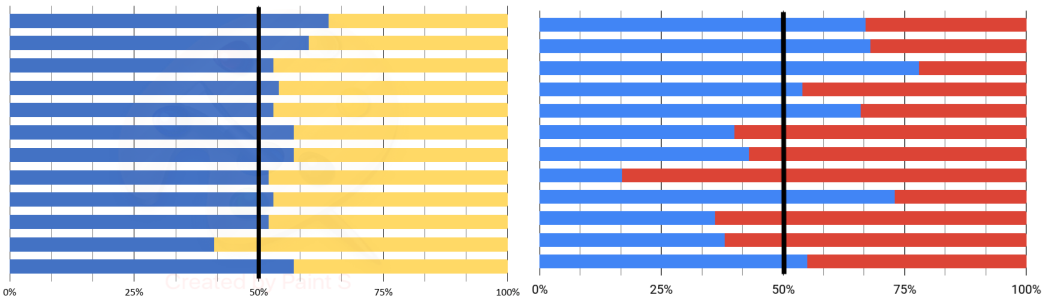


Figure 1. Distribution of talk of students that received SuCCESs reminders (left) and students that received switching reminders from their teacher (right).

The majority of switching reminders from the twelve pairs resulted in a *switched* (26) before the pair continued to work on the task. Pairs 1, 10, and 12 each had one instance of *willingly switched later* (3). Pair 2 (2) and 12 (1) had instances of *did not switch*. Pair 2 had one instance of *teacher had them switch*. In the focus groups, most students stated that they switched immediately; however, other students admitted that they would wait to switch or switched on their own based on how they perceived their and their partner's skills. Instances where pairs *acknowledged* (14) the talking reminder included 12 in which students expressed frustration through utterances such as "We don't care." Instances of *did not acknowledge* (9) showed that pairs either clicked "ok" immediately or left it on the screen.

We then compared the most talkative partners' proportion of talking time (e.g. .67 for Pair 1 and .68 for Pair 2) of Figure 1 between the datasets and found that the difference was significant (Wilcoxon Rank Sums Test; Z = 2.87, p = 0.004).

Case studies from SuCCESs condition (study 2)

Most pairs in the SuCCESs group adhered to the reminder to switch and they demonstrated more balanced talk. Below, we detail four pairs of students with varied reminder adherence and talk distribution.

Pair 2 (Surya and Evan): Did not always follow switching reminders, less balanced

Surya spoke 60% and Evan spoke 40% of the time. They were the second least-balanced of the 12 pairs. Excerpt 1:

Surya (Driver): I just had the best idea. Let's make him super huge so then we'll be able to see him.
Evan (Navigator): Get back on him.
Surya: No, no, no. Make him super huge. **[switching reminder appears; Surya continues working]**
Surya: I just had the best idea. So, no.
Evan: Gimmie it.
Surya: This is actually important [inaudible]. Let me try something. Scripts. **[Surya clicks ok]**

Pair 12 (Amy and Sarah): Did not always follow switching reminders, more balanced

Sarah spoke 57% of the time while Amy spoke 43% of the time. They had three switching reminder instances. Later, when the switching reminder appeared (After Excerpt 2), the students stayed in the same role. Excerpt 2:

Sarah: (Navigator): 90 degrees.
Amy: (Driver): Why 90?
Sarah: Wait, wait, wait. And.. Wait, hold on. I.. And it's this way.
Amy: You're doing it. How can people do this? [time passes] Since I'm the instruction now.
Sarah: Okay, so now it's next. We're at the very bottom.
Amy: Turn 90 degrees.
Sarah: Which way? This way or this way?

Pair 1 (Penny and Kendra) followed the reminders, less balanced

Penny (who spoke 67% of the time) and Kendra (who spoke 33% of the time) strictly followed the switching reminders, although they represent a less balanced pair in terms of speaking distribution. Excerpt 3:

Penny (Driver): Okay, how do I-I want a different costume. You know what, can we just like go back? Delete it? Help me. Help me Kendra, help me. I can't do it. I don't know how to delete anything.
Kendra (Navigator): Okay, okay. Can I have it now, Penny? Okay, so- What did you just do?
Penny: What do you mean what did I just do? Uh, why'd I make this smaller? I can make it bigger again. It's my secret though. **[physically turns laptop away from Kendra]**
...
Kendra: Yeah, I'm not. Not right now. Show me how to do it.
Penny: I will if you just turn around-
Kendra: No. Penny just do it…. Penny can you- I'm not gonna help you at all. **[Kendra sits back and to the side, refusing to look at the laptop]**
Penny: Done. You wanna know what it is? So it's this, Kendra. This button right here.
Kendra: I'm not gonna help you.

[For ~2 mins., Penny works independently and Kendra silently reads next to her.]

Pair 8 (Ariel and Charlotte) - Followed the reminders, more balanced

Ariel talked 52% of the time and Charlotte talked 48% of the time. Even though Charlotte protested as her partner took the laptop away, she quickly fell back into the navigator role by giving directives and suggestions. Excerpt 4:

Charlotte (Driver): Okay, let's do this quick so we can..."say". Backspace, backspace.
Ariel (Navigator): If, "say". No start up here.

Charlotte: So...sorry.
Ariel: It's gonna switch now. **[switching reminder appears]** Oh, let's go.
[grabs the computer and pushes her partner's hand away] Stop I can do it.
Charlotte: Hey. You should.
Ariel: Oh damn it. It's fine.
Charlotte: You should put back on "rotation".
Ariel: Where is "rotation"?

Discussion

This analysis has examined the ways in which students respond to talking and switching reminders and their talk balance during pair programming. Our evidence suggests that the talking reminder was not effective in triggering conversation: 12 out of 23 talking reminders were met with negative student comments. The switching reminder scaffold appeared to be helpful for pairs who agreed that they would adhere to the suggested structure. But, as the case studies illustrate, pairs 12 and 8 varied in how much they followed the switching reminder. Despite this, they were both relatively balanced pairs, with the stricter students (pair 8) being more balanced. The students in both pairs seemed to willingly adhere to their roles. Another important factor appeared to be the students' level of focus and engagement. When students in pairs 1 and 2 disengaged, they often physically positioned themselves away from the laptop (either towards a classmate (Evan) or toward other work (Kendra)); this suggests that they were not in control of the programming work (Deitrick et al., 2016). The behaviors of Penny and Surya may have been one reason their distribution of talk was more imbalanced than the pairs that followed the switching reminders similarly.

Other researchers that have taken a closer look at younger students' collaborative process have found that the equity of a collaborative relationship depends in part on the way the students position each other socially (Shah et al., 2014). We found that the students' way of positioning themselves as driver and navigator is also important to the talk balance in their relationship, as evidenced by the significantly more balanced discourse of the SuCCESs group. A study of girls' collaboration found that one student established herself in a more knowledgeable and authoritative position by speaking more, giving commands, and by maintaining control of the equipment which in this case was a keyboard and mobile phone (Deitrick et al., 2016). In our case studies, we found one pair of students that had similar problems with one student attempting to maintain control of the computer more than his partner. This also led the partner to disengage, which affected the talk balance of the relationship.

References

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, *12*(3), 307-359.
Broll, B., & Ledeczi, A. (2017, March). Distributed Programming with NetsBlox is a Snap!. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education* (pp. 640-640).
Deitrick, E., Shapiro, R. B., & Gravel, B. (2016). How Do We Assess Equity in Programming Pairs?. *Proceedings of the 2016 Annual Conference of the Learning Sciences* (pp. 370-377). Singapore.
Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning* (pp. 69-97). Springer, Berlin, Heidelberg.
Shah, N., Lewis, C. M., & Caires, R. (2014) Analyzing Equity in Collaborative Learning Situations: A Comparative Case Study in Elementary Computer Science. *Proceedings of the 2014 Annual Conference of the Learning Sciences*, (pp. 495-502). Boulder, CO.
Tsan, J., Lynch, C. F., & Boyer, K. E. (2018). "Alright, what do we need?": A study of young coders' collaborative dialogue. *International Journal of Child-Computer Interaction*, *17*, 61-71.
van Langenhove, L. & Harré, R. (1999). Introducing positioning theory. In R. Harré, & L. van Lagenhove (Eds.), Positioning theory: Moral contexts of intentional action (pp. 14-31). Malden, MA: Blackwell.
Williams, L., Kessler, R. R., Cunningham, W., & Jeffries, R. (2000). Strengthening the case for pair programming. *IEEE software*, *17*(4), 19-25.
Yamakawa, Y., Forman, E., & Ansell, E. (2009). Role of positioning: The role of positioning in constructing an identity in a third grade mathematics classroom. In *Investigating classroom interaction* (pp. 179-202). Brill Sense.
Zakaria, Z., Boulden, D., Vandenberg, J., Tsan, J., Lynch, C., Wiebe, E., & Boyer, K. E. (2019). Collaborative talk across two pair-programming configurations.

Acknowledgments

This work is supported by the National Science Foundation through the grant DRL-1721160. Everything expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

Automated Tracking of Student Activities in a Makerspace Using Motion Sensor Data

Gahyun Sung, Tyler Yoo, Edwin Chng, Stephanie Yang, Bertrand Schneider
gsung@g.harvard.edu, tyler_yoo@g.harvard.edu,
chng_weimingedwin@g.harvard.edu, szhang@g.harvard.edu, bertrand_schneider@g.harvard.edu
Harvard University

Abstract: Learning inside makerspaces can be difficult to track and support. Using Kinect data collected from students enrolled in a course for making, we explore ways to track student learning trajectories in an automated way. Namely, by transforming our data into a set of action sequences that span a semester, we are able to find that discouraged students display a statistically distinct type of activity pattern already in the first two weeks. Generating metrics on makerspace use, we also find that time spent alone and the number of transitions between stations are significant indicators for discouragement and motivation levels. We argue that high-frequency location data could provide an accessible, meaningful overview of student learning in a makerspace to all stakeholders, and conclude with limitations and future directions.

Makerspaces are collaborative learning environments that have the potential to promote key 21st century skills such as creativity, curiosity, and problem solving for participants (OECD, 2018). One challenge however is that it can be difficult to track and support students working on various open-ended projects in and outside supervised hours. Further, students who need the most help are often the ones most hesitant to seek it, causing them to fall behind and eventually give up on future making experiences. A pressing need in makerspace education therefore is to help instructors and facilitators track and support dispersed student experiences inside the makerspace.

Using sensor data to track learner actions inside the makerspace could provide a way to automate part of this task, as well as a way for researchers to quantitatively investigate the connection between different student trajectories and targeted outcomes. Cooke and Charnas (2019) suggest gate counters and sign-in systems as ways to generate useful information. We believe that this approach could be taken a step further by collecting fine-grained location data. Indeed, as a precursor to this study, Chng et al. (2020) has shown how location data from a multi-sensor Kinect system can provide insights into the social interactions inside a makerspace.

Generally, however, very few papers to date have used location data to explore makerspaces. One cause of the slow uptake may be that location data from physical learning environments tend to be large and complex, requiring extra cleaning and feature engineering steps. Thus, a pipeline to extract meaningful metrics and action sequences is an outstanding challenge for the widespread use of location data in research and practice. Our paper presents the preprocessing and analysis steps carried out on location data from 24 students enrolled in a maker course. Generating metrics and action sequences from location data, our paper 1) creates proof-of-concept visualizations of student trajectories in makerspaces; 2) tests if different student groups show distinct action sequences; and 3) tests if there are location-based indicators for student outcomes of interest.

Methods

Our setting is a maker course for education graduate students in a private northeastern U.S. university. Students worked freely in the makerspace on weekly tool-specific assignments as well as on a final project that asked students to create a digital fabrication product for educational purposes. The bulk of student making was done independently, with office hours and individual consultations providing instructional support.

The survey data comes from weekly surveys administered after every class. Surveys were crafted via a literature review of surveys that measure student states, namely self-efficacy and motivation (Pintrich & DeGroot, 1990; Williams & Deci, 1996), maker mindset (Clapp et al., 2016), and affective attitude (Watson et al., 1988). After adapting questions for our context, we validated the survey with input from students from a previous iteration of the course. Later, to address correlation between question responses and to improve interpretability of results, we conducted factor analysis on the survey data prior to final analysis. Following the recommendations of Costello and Osborne (2005), we used a scree plot to select the number of factors and affirmed that no items had less than a 0.4 loading. We were thus able to summarize 14 survey items into four factors: motivation, self-efficacy, help perceptions, and discouragement, where factors have by design a mean of 0 and variance near 1.

Student location data was continuously collected with six Microsoft Kinect v2 sensors placed around the walls. We began preprocessing the data by generating student ID labels with OpenFace. Then, using Cv2, data from all sensors was mapped onto a 2D coordinate system of the room. Next, we dropped all but one instance of an individual student observed by multiple sensors. Lastly, we gave each coordinate a location label (e.g., table, 3D printer, etc.), while filtering out transitions between stations and sparse data from the weekends. This facial recognition → homography → deduplication → labeling process gave us, per student, a series of location labels corresponding to their movements in the makerspace for the entire semester.

In making meaning of this data, we focused on hours and co-work as we were most interested in indicators that could help detect students in need of support. We hypothesized that spending relatively large portions of time alone and/or spending long hours working in the makerspace could be linked with high discouragement and low motivation. We were also interested in finding a proxy for iterative work, as it is well known that iterative processes are beneficial to problem-solving (Atman et al., 2007). In our data, we believed the number of times a student crosses location boundaries to work with different tools, or to return to tools after working on a non-tool task at the table (e.g., talk with people, look up information with personal laptop) could be one. Ultimately, we created and explored four metrics per student: *together time* versus *alone time* (i.e., time spent without OR with someone within 1 meter), *total time* in makerspace, and *number of transitions* made between stations. In figure 1, the dotted line in the left image shows a fictional student making 3 transitions (2-3, 3-4, 4-5). As no one else was in the space, 'alone time' equals 'total time' for this student.

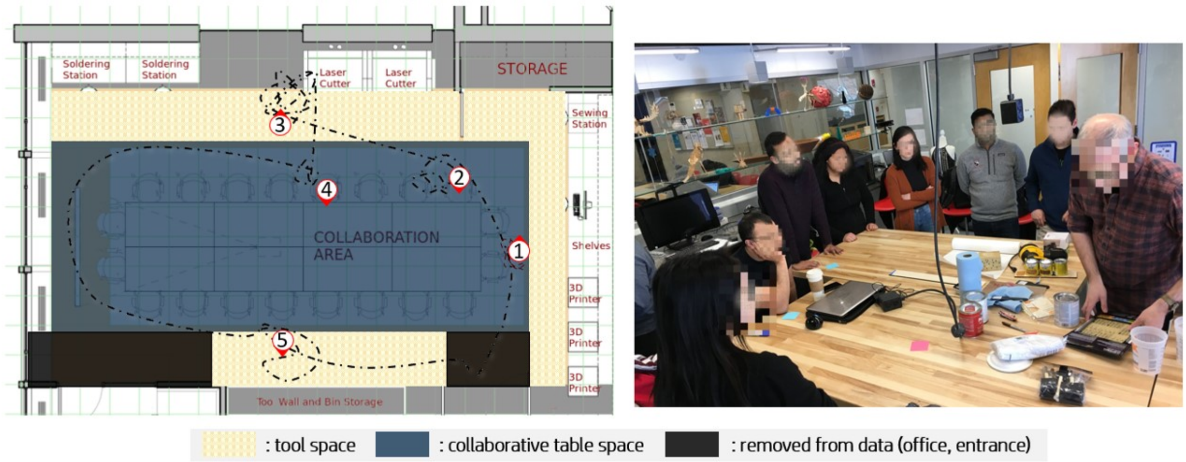


Figure 1. Makerspace layout and photo of actual learners in the space.

In addition to weekly summative metrics, we also generated daily action sequences for each student from the data and investigated whether there are patterns connected to student outcomes. Looking again at total time and co-work proportions, we gave each day of a student one out of five labels: a long day spent alone at the makerspace, a short day spent alone, a long day spent together, a short day spent together, and absence. Long/short and alone/together thresholds were median values for the entire semester. As a result, each student had an action sequence of length 57, the number of days data was collected during the semester excluding the weekends.

For sequence visualization and analysis, we utilized R's TraMiner package (Gabadinho et al., 2011). For analysis on summative metrics, we used R to regress weekly survey outcomes on the weekly summative metrics while controlling for idiosyncratic week and student effects, as both were seen to affect survey outcomes. Data for statistical analysis consisted of 11 weeks' worth of survey and Kinect-derived metrics from 24 students. As we dropped 7 rows from the data due to missing student survey responses, the final dataset had 257 rows (11 x 24 – 7), where each row corresponds to a week of a student.

Results

We first provide an overview of student activities in the makerspace, derived from the Kinect data. Students spent an average of around 5 hours a week in the makerspace (316 minutes), and time spent in the space grew about 8 minutes on average per week ($b = 7.89$, $se = 3.81$, $p = 0.04$). Students tended to underreport the time spent in the makerspace by 3 hours a week (179 minutes, $sd = 391$) on average, likely due to the fact that we asked students to only estimate the time spent in the makerspace working on assignments. A point increase in reported frustration (7-point scale) was linked to an average student's reporting error (reported time – Kinect time for a week) being 70 minutes larger, controlling for student effects ($b = 70.49$, $se = 17.14$, $p < 0.001$). That is, students who reported higher frustration tended to feel they spent more time in the makerspace. Students only spent about 9% of the total time at tool stations. Among the tools, laser cutters were used most frequently and sewing machines the least, which aligns with instructor observations on how students used the space.

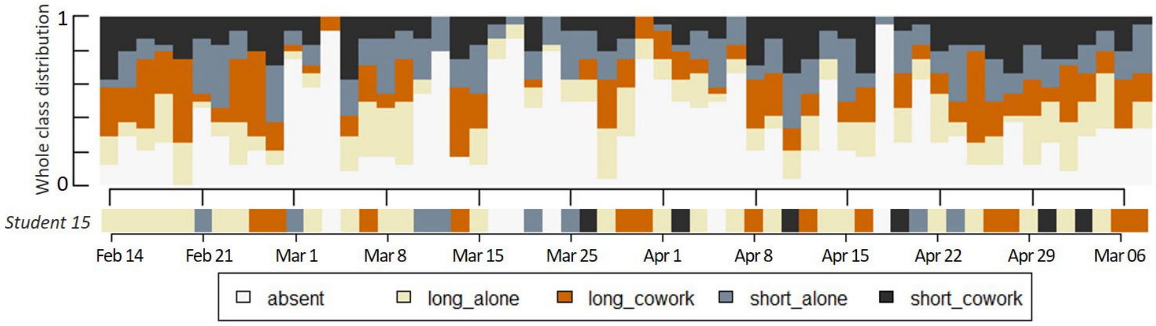


Figure 2. Whole class and individual student trajectories.

Figure 2 demonstrates one way to visualize student learning trajectories over the course of a semester. With the stacked graph above, instructors and facilitators can see whole-class trends from February to May. We see spring break appearing as a spike in absences mid-March, and students spending longer hours for final project work in later weeks (less absent, short_alone, short_cowork days). The single line for student 15 demonstrates how we could zoom in on an individual student for check-in purposes and see e.g., a student spending long hours alone in the makerspace at the start of the semester. This could warrant a check-in based on our sequence analysis, where we found that groups with high and low discouragement levels show statistically significant differences in their action sequences in the first two weeks (discrepancy test on dissimilarity matrix; $F = 1.51$, $p = 0.048$). Visual inspection showed that those who report a level of discouragement lower than the mean have more co-work days in the early weeks. Regression analysis on the first two weeks also show that the discouragement factor is positively associated with alone time and negatively with together time, although these associations are not statistically significant.

The observation regarding the association of discouragement and co-work in the early weeks is contrasted by what we find in the regression analysis for the entire semester, where we regressed each of the four weekly outcome factors on each of the six metrics. Namely, we find that the discouragement factor was negatively associated with alone time, controlling for week and student effect ($\beta = -0.0008$, $se = 0.0003$, $p = 0.003$; note that small scale of effect is due to large X, minutes spent for entire semester, while Y has mean 0 and $sd \approx 1$). Also for the entire semester, we found that groups with low and high help perceptions differ in their sequences ($F = 1.592$, $p = 0.031$). Not surprisingly, students who felt they gave and received less help were characterized by shorter days at the makerspace and more absences. An additional finding is that the number of transitions between stations was positively associated with motivation ($\beta = 0.014$, $se = 0.005$, $p = 0.002$).

Discussion

Our results demonstrate different ways in which high-frequency location data from makerspaces could be used to track students. With sequences and metrics derived from Kinect data, we find that the activity patterns of students with lower discouragement tend to show higher levels of co-work in the early weeks compared to those feeling more discouraged, but that more solo work was indicative of low discouragement for the semester overall. This may imply that students benefit more from peer support in the earlier weeks when frustration is generally higher, but that investing more time working alone on a task becomes more important as the class advances.

We also find that transitions between stations are positively associated with the level of motivation throughout the semester. This aligns with prior research that finds that an iterative design process is important for creating effective solutions (Atman et al., 2007). Conversely, it could also mean that motivated students are more willing to undergo this effortful iterative process. This is a promising metric for a model to automatically detect students who are frustrated and lose motivation without asking for help in makerspaces.

Lastly, we contribute intuitive visualizations that could help instructors and facilitators monitor student activities inside the makerspace, including those carried out outside of supervised hours. While the proof-of-concept visualizations in the current paper only display total time and time spent alone, more sophisticated, interactive visualizations could integrate multiple layers of information inferable from location data. Immediate uses of such representations include learning what stations are underutilized or need additional resources, or getting an objective sense of the workload students are experiencing from a course week by week. This can offer benefits similar to teacher orchestration graphs (Prieto et al., 2018), which is posited to benefit teacher professional development and serve as a novel quantitative method to understand the process of learning in physical learning environments.

Conclusion: Future directions and Limitations

There are several limitations to our paper, the largest one being the coarseness of the labels we create for student states. Table time takes up around 92% of the total time. Co-work, which takes up around 60%, is also too broad as co-work can take many forms in a makerspace. While we generated more elaborate labels such as active building or working side-by-side based on skeletal joint data, we ultimately decided that we do not have a reliable way to validate these labels without ground truth data. Collecting short videos clips along with Kinect data for validation purposes seem essential to fully utilize the rich information latent in Kinect data.

We also note that the outcomes utilized in this study are limited to self-reported states. In future studies, we hope to generate researcher codes with protocols such as BROMP (Ocumpaugh, 2015), or seek input from instructors to triangulate student outcomes. With these improvements to the data, we hope to replicate findings, explore new metrics, and use new analysis methods such as frequent pattern mining or recurrent neural networks to advance our understanding of what action patterns in makerspaces can tell us of student states.

Our results show that location data from makerspaces can provide insights into student learning processes and help track student states. Data from makerspaces are especially well-suited for this approach as most student activities, in and outside of regular class hours, occur inside the makerspace. Additionally, distinct makerspace stations can be labelled in the data and provide contextual information. We believe it can be both feasible and impactful to equip makerspaces with location data-based learning dashboards that visualize student trajectories.

Moving beyond makerspaces, utilizing location data have the potential to yield generalizable information on student collaborations, or help track student states in any hands-on, physically active learning activities. Particularly with the maturation of algorithms that track people's movement in a space with low-cost camera systems (Mulloni et al., 2009), establishing an analysis pipeline for extracting pedagogically meaningful metrics and sequences from location data could offer a new means to understanding what goes on inside a wide variety of physical learning environments.

References

Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. Journal of engineering education, 96(4), 359-379.

Chng E., Seyam M.R., Yao W., & Schneider B. (2020). Using Motion Sensors to Understand Collaborative Interactions in Digital Fabrication Labs. In: Bittencourt I., Cukurova M., Muldner K., Luckin R., Millán E. (eds) Artificial Intelligence in Education. AIED 2020. Lecture Notes in Computer Science, vol 12163. Springer, Cham.

Clapp, E. P., Ross, J., Ryan, J. O., & Tishman, S. (2016). Maker-centered learning: Empowering young people to shape their worlds. John Wiley & Sons.

Cooke, M. N., & Charnas, I. C. (2019). The Value of Data, Metrics, and Impact for Higher Education Makerspaces. IJAMM, 1(1).

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical assessment, research, and evaluation, 10(1), 7.

Gabadinho, A., Ritschard, G., Müller, N.S. & Studer, M. (2011), Analyzing and visualizing state sequences in R with TraMineR, Journal of Statistical Software. 40(4), pp. 1-37.

Mulloni, A., Wagner, D., Barakonyi, I., & Schmalstieg, D. (2009). Indoor positioning and navigation with camera phones. *IEEE Pervasive Computing*, *8*(2), 22-31.

Ocumpaugh, J. (2015). Baker Rodrigo Ocumpaugh monitoring protocol (BROMP) 2.0 technical and training manual. Technical Report. New York, NY: Teachers College, Columbia University. Manila, Philippines: Ateneo Laboratory for the Learning Sciences.

Organisation for Economic, Co-operation and Development. (2018). The Future of Education and Skills Education 2030. Paris: OECD.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. Journal of educational psychology, 82(1), 33.

Prieto, L. P., Sharma, K., Dillenbourg, P., & Jesús, M. (2016, April). Teaching analytics: towards automatic extraction of orchestration graphs using wearable sensors. In Proceedings of the sixth international conference on learning analytics & knowledge (pp. 148-157).

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of personality and social psychology, 54(6), 1063.

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: a test of self-determination theory. Journal of personality and social psychology, 70(4), 767.

Individually preparing learners to perform interactive activities in CSCL: do generative tasks work?

Stephan Mende, Antje Proske, Susanne Narciss
stephan.mende@tu-dresden.de, antje.proske@tu-dresden.de, susanne.narciss@tu-dresden.de
Dresden Technical University

Abstract: CSCL can foster in-depth knowledge acquisition if learners build on each other's knowledge to draw new conclusions. Individually preparing learners with generative tasks has been proposed to foster such interactive activities in subsequent collaboration. However, the scarce evidence provides no clear picture. Thus, we conducted an experimental study where participants went through the following phases prescribed by a CSCL-script: after reading a text learners prepared individually by answering a task in written form, read the co-learner´s task answer, and subsequently discussed the text via chat. According to experimental condition, the preparation phases consisted of either non-generative, compare-and-contrast, or explanation tasks. Results revealed no overall effect of preparation task type on interactive discussion activities. However, generative preparation tasks affected interactive activities indirectly through the learners' and their co-learners' generative preparation activities. This suggests generative tasks can foster interactive discussion as far as the dyad partners actually enact generative preparation activities.

Introduction

CSCL can foster in-depth knowledge acquisition, *if* learners perform interactive activities, that is, when they build on the contributions of their co-learners (*referencing*) to generate new knowledge (*inferencing*; Chi & Wylie, 2014; Fischer, Kollar, Stegmann, & Wecker, 2013). In this way learners can use their co-learners as learning resources in addition to the given instructional material in the service of knowledge construction, thus developing an understanding beyond what they might have achieved when learning alone. Unfortunately, learners do often not spontaneously enact interactive activities when they collaborate without support (e.g., Fischer et al., 2013).

A support strategy often argued as effective in addressing this issue is individual preparation for collaborative learning: having learners process the instructional material by themselves (e.g., taking notes from a text) before collaborating with others (e.g., discussing the text; Lam & Kapur, 2017). Preparing individually may allow the learner to activate relevant knowledge and to develop some initial ideas and conclusions with regard to the instructional material or task before having to invest efforts in communicating and coordinating with others. This in turn may a) facilitate interactive engagement with the additional knowledge and information contributed by co-learners in subsequent collaboration and b) increase the amount of additional knowledge and ideas each co-learner can contribute, so that altogether a wider range of information is available through which interactive activity could be stimulated (e.g., Lam & Kapur, 2017; Mende, Proske, & Narciss, 2020).

However, the related empirical evidence reveals a mixed picture: individual preparation for collaboration sometimes has positive, no, or even negative effects (Mende et al., 2020). This raises the question of *how* learners should prepare or, in other words, how instructors should design an individual preparation phase to raise its potential benefits in view of the performance of interactive activities during subsequent collaboration.

In this regard, prior research has especially addressed the role of a) generative preparation tasks, that is, requiring the learner to infer new knowledge beyond what is given in the instructional material (e.g., Fiorella & Mayer, 2016) and b) having learners inspect each other's individual preparation products before the collaboration (e.g., viewing co-learners' notes) which could be understood as a specific form of awareness induction support (e.g., Janssen & Bodemer, 2013).

While there is already some evidence that awareness induction support can improve the effects of individual preparation on interactive activities during subsequent collaboration, the role of generative preparation tasks is far less clear (Mende et al., 2020). Though, generative tasks have been shown to foster deeper processing of and learning from subsequent lectures in individual learning (e.g., Schwartz & Martin, 2004), their effectiveness in preparing learners for a subsequent collaboration phase has rarely been examined. Moreover, the few existing studies showed a rather mixed picture: generative preparation had either no (Lam & Muldner, 2017) or even negative effects (Lam, 2019) on subsequent collaborative learning.

Two reasons may account for this issue. First, distinguishing between generative and non-generative tasks is a rather rough classification. For example, two commonly used generative tasks are a) comparing and contrasting to-be-learned concepts and b) generating causal explanations for to-be-learned concepts or phenomena (e.g., Chi & Wylie, 2014). Although both tasks are considered to be generative, some prior research suggests explanation tasks to engage learners in generative activities to a higher degree than compare-contrast tasks (e.g., Chin, Chi, & Schwartz, 2016). In addition and regardless of the task type, the provision of generative tasks is no guarantee that learners will actually perform generative activities (Chi & Wylie, 2014). Moreover, though generative activities have been shown to foster deep knowledge acquisition (e.g., Fiorella & Mayer, 2016), it is not clear whether preparing generatively by one's own necessarily fosters subsequent high quality collaboration in terms of interactive activities. Hence, it is necessary to examine a) whether and to what degree generative tasks foster generative activities during individual preparation and b) whether and to what degree these generative activities in turn foster interactive activities during subsequent collaboration.

Second, the extent to which a learner engages in interactive activities during collaboration may also depend on the degree to which their co-learners have prepared generatively. This is even more relevant when an individual preparation is accompanied by awareness induction support. In other words, the extent to which learner A performs interactive activities may also depend on the degree to which the inspected preparation product of co-learner B contains additional knowledge, conclusions, and ideas not already stated in the instructional material since this new information can stimulate learner A to further develop or challenge learner B's contributions (e.g., M. Erkens, Bodemer, & Hoppe, 2016; Lam & Kapur, 2017).

To our knowledge, these issues have not been systematically considered so far. To address these research gaps we conducted an experimental study in which participants were guided by a CSCL script through the following phases: After reading a text, learners individually answered a task in written form (preparation phase). Learners were subsequently requested to read their co-learners' task response to get aware of each other's knowledge, ideas and perspectives (awareness induction phase). Finally, learners were requested to chat on the text collaboratively (discussion phase). Three versions of this script were developed which differ in the type of task provided in the individual preparation phase (i.e. two different generative and a non-generative task) in order to answer the following research questions:

1) What is the effect of the individual preparation task type on the number of interactive activities during the collaborative discussion phases?

2) Does the individual preparation task type indirectly affect the number of interactive discussion activities through the number of the learner`s and/or their co-learner's generative preparation activities?

Method

In this experimental e-learning study 122 undergraduates (72.9 % female, mean age: 22.81 years, *SD* = 3.95) of psychology (49.2%) and educational sciences (50.8%) learned from an expository text about the human circulatory system translated and adapted from Chi et al. (2001). Participants were randomly assigned to stable dyads which in turn were randomly assigned to one of three experimental conditions. Participants followed the above described script for each third of the text. Conditions differed regarding the task type administered in the individual preparation phases.

Participants in the first condition received the task to take notes from the text as for an exam and, thus, were not specifically requested to perform generative activities (non-generative task condition). The remaining conditions were provided with two commonly used types of generative tasks: Subjects in the second condition were required to compare-and-contrast central concepts addressed by the text (comparison task condition). For example, subjects were asked to compare the different types of blood vessels of the circulatory system concerning their components and the processes in which they are involved. Participants in the third condition were requested to provide explanations concerning the same central text concepts as in the comparison task condition (explanation task condition). For instance, the learners were asked to find reasons why we have different kinds of blood vessels instead of only one type in our circulatory system.

The learning activities performed in the individual preparation and the collaborative discussion phases were analyzed through coding schemes. The protocols of each learner's individual preparations (i.e. the written task responses) were assessed with respect to indications of generative activities. To this end, each protocol was coded in terms of the number of sentences containing inferences, that is, topic-relevant information not already contained in the given learning text (Chi & Wylie, 2014). For example, a comparison of the thickness of capillaries and arteries or a causal explanation such as "diffusion is not possible in arteries because of their thick walls" was scored as generative activity because these comparisons or explanations were not explicitly presented in the text. In contrast, mere repetitions of text information were not considered generative activity. A second rater coded 25% of the individual preparation protocols (Krippendorfs $\alpha = .91$). The resulting score represents the sum of generative activities a learner has performed during the individual preparation phases.

The chat discussion protocols were assessed with respect to indications of interactive activities. To this end we segmented the participants chat-messages according to punctuation and "connectives" (G. Erkens &

Janssen, 2008). Each segment was then assessed for whether it contained a) an inference (see above) and b) indications of a reference to a prior contribution of the co-learner in terms of incorporating or taking into account a co-learner´s previous preparation task response or chat message. A second rater coded 25% of the chat discussion protocols (Krippendorfs $\alpha = .82 – .88$). Segments which were coded as containing indications of both inferencing *and* referencing were counted as interactive activities. The resulting score represents the sum of the interactive activities a learner has performed during the collaborative discussion phases.

Since subjects were nested in dyads, we conducted linear mixed regressions for dyadic data (Kenny, Kashy, & Cook, 2006) to address our research questions. As the variables representing the number of generative and interactive activities revealed some deviations from a normal distribution we performed bootstrap analyses with 5000 resamples to estimate the standard errors and confidence intervals for all regression coefficients (e.g., Afifi, Kotlerman, Ettner, & Cowan, 2007). An effect is considered significant at a 5% significance level if the 95% bootstrap interval does not include zero.

Results

Experimental conditions did not differ in prior knowledge. To address research question 1 (effects of task type on interactive discussion activities) a linear mixed model with experimental condition as predictor and interactive activities as dependent variable was conducted. Experimental condition was Helmert-coded so that the non-generative task condition is compared to the generative task conditions (i.e. the comparison and explanation task conditions together) and the comparison task condition is compared to the explanation task condition. None of these comparisons revealed significant effects indicating that the preparation task type did not affect the number of interactive activities during the discussion phases.

To address research question 2 (indirect effects of task type on interactive discussion activities through generative preparation activities) we conducted two linear mixed regressions followed by the computation of the indirect effects according to the recommendations of Hayes (2013). First, the effects of the independent variables (i.e. experimental conditions) on the mediators (i.e. number of individual generative preparation activities) were estimated (a-paths). Results revealed the generative task conditions to positively affect the number of generative preparation activities in reference to the non-generative task condition (B = 8.61, BCa $CI95\%$ = 7.59–9.61). Further, the explanation task condition had a positive effect on the number of generative preparation activities in reference to the comparison task condition (B = 5.12, BCa $CI95\%$ = 3.31–6.92).

Second, we examined the effects of the learners' (actor-effects) and their dyad partners' (partner-effects) individual generative preparation activities on the number of the learners' interactive activities while controlling for the effect of experimental condition (b-paths). Thereby we followed the procedure for estimating an Actor-Partner-Interdependence model for indistinguishable dyads (Kenny et al., 2006). Results showed positive actor (B = .25, BCa $CI95\%$ = .15–.36) and partner (B = .27, BCa $CI95\%$ = .12–.42) effects of generative preparation activities on interactive discussion activities. In other words, learner A´s interactive activities were positively affected by both learner A´s and their co-learner B´s generative preparation activities.

Third, the a-path and the b-path coefficients as well as their bootstrapped standard errors were used to calculate the indirect effects (ab-paths) along with 95% Monte Carlo confidence intervals based on 100000 replications, using the SPSS macro MCMED (Hayes, 2013). Results reveal that – in reference to the non-generative task condition – the generative task conditions had positive indirect effects on interactive discussion activities mediated through the actors (B = 2.15, MC $CI95\%$ = 1.19-3.16) as well as the partners (B = 2.29, MC $CI95\%$ = 1.18-3.45) generative preparation activities. In other words, the results indicate generative preparation tasks to benefit interactive activities to the extent to which these tasks induce the actor and the partner to perform generative preparation activities. Further, compared to the comparison task condition, the explanation task condition had positive indirect effects on interactive discussion activities mediated via the actor's (B = 1.28, MC $CI95\%$ = 0.63-2.07) and the partner's (B = 1.36, MC $CI95\%$ = 0.63-2.25) generative preparation activities.

Discussion

The present study examined the role of generative tasks in individually preparing learners for performing interactive activities during subsequent collaboration. We found no overall effect of preparation task type in this regard. However, mediation analyses revealed generative as compared to non-generative preparation tasks to foster generative preparation activities which in turn benefitted the performance of interactive activities in collaborative discussion. Thereby the learners own as well as their partners' generative preparation activities revealed to be significant mediators. These results contribute in several ways to our understanding of how individual preparation may affect subsequent collaborative learning. First, they indicate that generative activities may not only be suited to prepare learners for subsequent individual knowledge acquisition, but also for productive interactions in collaborative learning (cf. Lam & Kapur, 2017). Second, the crucial factor influencing subsequent collaboration quality seems not to be the preparation task per se, but rather what the learners actually do with the task. Third, for a learner to perform interactive activities, both their own and their co-learner's generative preparation activities seem to be important. Finally, in line with and extending previous research (e.g., Chin et al., 2016), explanation tasks invoked the described mechanisms to a higher degree than compare-contrast tasks.

While the results of the mediation analyses suggest individual preparation through generative tasks to be preferable to preparation by non-generative tasks, we observed no total effect of task type on interactive discussion activities. Among others, the following explanation may account for this. Although the preparation task conditions significantly differed in terms of the number of generative preparation activities, the standard deviations were remarkably high. Thus, the degree to which the particular generative preparation task fostered generative activities strongly varied among the learners. This may be due to task characteristics, learner characteristics, or both. For example, the extent to which learners are able to perform generative activities in response to a generative task may depend on their prior knowledge. Analyses of the results are ongoing and this issue will be examined soon. In addition, we will also investigate the effects of the preparation task type as well as of actors' and partners' generative and interactive activities on learners transfer knowledge as captured in a posttest after a delay of one week.

References

Afifi, A. A., Kotlerman, J. B., Ettner, S. L., & Cowan, M. (2007). Methods for improving regression analysis for skewed continuous or counted responses. *Annual Review of Public Health*, *28*, 95–111. https://doi.org/10.1146/annurev.publhealth.28.082206.094100

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471–533. https://doi.org/10.1207/s15516709cog2504_1

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Chin, D. B., Chi, M., & Schwartz, D. L. (2016). A comparison of two methods of active learning in physics: inventing a general solution versus compare and contrast. *Instructional Science*, *44*(2), 177–195. https://doi.org/10.1007/s11251-016-9374-0

Erkens, G., & Janssen, J. (2008). Automatic coding of dialogue acts in collaboration protocols. *International Journal of Computer-Supported Collaborative Learning*, *3*(4), 447–470. https://doi.org/10.1007/s11412-008-9052-6

Erkens, M., Bodemer, D., & Hoppe, H. U. (2016). Improving collaborative learning in the classroom: Text mining based grouping and representing. *International Journal of Computer-Supported Collaborative Learning*, *11*(4), 387–415. https://doi.org/10.1007/s11412-016-9243-5

Fiorella, L., & Mayer, R. E. (2016). Eight Ways to Promote Generative Learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

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. https://doi.org/10.1080/00461520.2012.748005

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regressionbased Approach*. New York: Guilford Press.

Janssen, J., & Bodemer, D. (2013). Coordinated Computer-Supported Collaborative Learning: Awareness and Awareness Tools. *Educational Psychologist*, *48*(1), 40–55. https://doi.org/10.1080/00461520.2012.749153

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York: Guilford Press.

Lam, R. (2019). What students do when encountering failure in collaborative tasks. *Npj Science of Learning*, *4*(1), 1–11. https://doi.org/10.1038/s41539-019-0045-1

Lam, R., & Kapur, M. (2017). Preparation for Future Collaboration: Cognitively Preparing for Learning From Collaboration. *Journal of Experimental Education*, 1–14. https://doi.org/10.1080/00220973.2017.1386156

Lam, R., & Muldner, K. (2017). Manipulating cognitive engagement in preparation-to-collaborate tasks and the effects on learning. *Learning and Instruction*, *52*, 90–101. https://doi.org/10.1016/j.learninstruc.2017.05.002

Mende, S., Proske, A., & Narciss, S. (2020). Individual preparation for collaborative learning: Systematic review and synthesis. *Educational Psychologist*, *56(*1), 1–25. https://doi.org/10.1080/00461520.2020.1828086

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. https://doi.org/10.1207/s1532690xci2202_1

Crafting Human-material Collaborative Learning Processes and Technology Advances

Anna Keune, University of California, Irvine, Ruhr-University Bochum, akeune@uci.edu
Nickolina Yankova, Kylie Peppler
nyankova@uci.edu, kpeppler@uci.edu
University of California, Irvine

Abstract: Computer supported collaborative learning has a history of investigating learning processes between people and the technological materials that support them. Emergent posthumanist perspectives view technology as actively produced through routinized actions by humans and non-human materials. Applied to CSCL, this view would suggest that materials are active participants. The present inquiry explores materials that are historically connected to technology advancements–fiber crafts–to investigate material collaborative learning processes and the kind of learning they produce. Findings present material-collaborative learning processes that show crafters collaborate with materials to produce physical evidence of learning, i.e., craft-technology advances. This work has implications for theorizing and designing for collaborative learning, expanding who and what can be considered a participant.

Introduction

The field of computer-supported collaborative learning (CSCL) looks as collaboration as a process between people (i.e., pair, small group, whole classroom) often aided by computational materials (Stahl, 2013). Despite a rich array of theoretical approaches in CSCL, several underlying assumptions operate within the field, including that humans are the active participants in collaboration and that learning happens to humans. As a result, the CSCL technological materials have typically acted as mediators of collaborative processes and human learning (Yrjönsuuri et al., 2019). These theoretical assumptions treat CSCL materials as hierarchically inferior to people, potentially obscuring collaborative learning processes that the materials bring to learning.

Posthumanist perspectives (Barad, 2003) emerging in the learning sciences and CSCL are challenging assumptions of what is the active player in collaborative learning. Applied to CSCL, posthumanist perspectives decenter the human and suggest that collaborative learning processes may emerge from a human-material collaboration (Keune & Peppler, 2019). This is an important area to explore because posthumanist perspectives can address deficit notions and broaden what counts as learning (Mehto et al., 2020; Keune, 2020), potentially furthering the design of equitable learning environments. We consider fiber crafts a suitable context for this inquiry of posthumanist perspectives for CSCL because fiber crafts have a historical connection to computation (Essinger, 2004), make computational processes transparently performable (Keune, 2020), and are often performed collaboratively with others, for instance, through crafting circles. What remains unclear is what processes characterize a human-material collaboration and what may be identified as learning. Thus, we asked: *What collaborative learning processes do crafters reference when talking about engaging with materials while crafting? What learning is being produced during this collaboration?*

To answer the questions, we take a posthumanist stance that decenters the human to also positions materials as active contributors in collaboration to analyze 65 semi-structured interviews with adult crafters (e.g., knitters, quilters, sewers) and 397 photographs of crafters' projects through iterative thematic analysis. Findings present two themes that characterize the collaboration of crafter and craft materials: 1) *sensory collaboration,* where collaboration relies on the exchange of sensory information between crafters and materials through touch and sight and 2) *genealogical collaboration,* where expansion of the craft is made possible through collaboration that builds on past and present materials, and across different craft forms. Through the interview analyses, we observed that the craft technologies and techniques expanded and changed over time, presenting a material form of evidence of collaborative learning that neither belongs to the crafter nor the craft alone, but is a product of the shared engagement. This work has implications for theorizing collaborative learning processes and outcomes, expanding what can be considered a participant in collaborative learning.

Background

Physical and digital materials are often considered mediators for human learning (e.g., Yrjönsuuri et al., 2019; Stahl, 2013). Materials have been theorized as instruments that serve specific human learning objectives where materials turn into instruments based on context, practices, and the kind of knowledge that is intended to be gained (e.g., Ritella & Hakkarainen, 2012). Shared assumption center humans as main actor in collaborative learning and materials as in-between humans. By contrast, more recent posthumanist perspectives call to decenter humans (Mehto et al., 2020). With histories in indigenous theories, physics, animal studies and beyond, posthumanist perspectives are concerned with inviting a range of voices into the inquiry of how materials work to produce human worlds (Kuby, 2017). These perspectives argue that non-humans can act on what manifests in the world (Mehto et al, 2020) and point to capture physical changes that are based on routine movements of people and materials (e.g., fiber crafts) and related to disciplinary domains as evidence of learning (Keune, 2020). Applied to CSCL, this questions what materials do in collaborative learning (Keune & Peppler, 2019).

An interesting question then is to explore what physical change happens through collaborative processes of people and materials. Fiber crafts are an interesting context for exploring this because they are historically linked to the fast-changing technological landscape (e.g., Essinger, 2004) and can be used to perform computational concepts (e.g., loops; Keune, 2020). While the discovery of new ideas and techniques can happen through direct experience with materials through trial and error (Gore, 2004), it remains unclear how crafters and craft may collaborate to inform new techniques and projects. Looking at crafting as collaborative technological practice from a posthumanist perspective presents a lens on collaborative learning that plays with established notions within collaborative learning, including processes, players, learning outcomes, and computation, and promises to see whether there is something that an additional set of theoretical tools could bring to the field.

Methods

This qualitative study included 65 adult crafters (20-73-years old) who were experienced in an average of three crafts (e.g., quilting, knitting, weaving). Crafters were recruited through online and offline crafting communities followed by snowball sampling. The majority of participants were female (n=60, 92%) and clustered around the Midwestern United States (n=38, 58%). Most crafters were white (n=57, 88%), four were Black (6%), three Asian (5%), and one did not share their racial identity (2%). Interviews were conducted at crafters' homes, studios, in public places, or through video conference.

Data sources

To investigate collaborative learning processes with materials, we analyzed 65 semi-structured interviews with experienced adult crafters that were conducted as part of an investigation into fiber craft for mathematics learning. Interviews were on average 59 minutes long and included questions related to 1) demographics, 2) connections between math and crafts, and 3) learning crafts. The interviews contributed to our understanding of the collaborative learning processes that crafters referenced as they reflected on their collaborative doing with the materials. The interviews also provided insights into the outcomes of such collaboration. To better understand what learning was produced as crafters and materials collaborated, data sources also included 397 photographs of crafters' projects. Of all, 21 crafters (32%) were comfortable sharing pictures of their crafts. The photographs contributed to visualizing the projects crafters mentioned and to reconstruct collaborative processes and outcomes.

Analytical techniques

To understand how crafters communicate collaborative processes with craft materials and how learning is produced, we engaged data that was collected based on humanist assumptions (i.e., interviews of people rather than material expressions) from a posthumanist point of view, assuming that materials take on an active role (Jackson & Mazzei, 2011). Thus, we first summarized any evidence that we could find in which crafters referenced collaborative engagement with materials, that is engagement in which it was unclear whether it was the crafter or the material that produced a project or a new technique. To ensure a shared understanding of what aspects to include in the summaries, Author 1 and 2 each created independent summaries of 13 interviews and then discussed them toward a shared summary. The shared understanding of the guided the writing of the remainder of the summaries. Then, Author 1 and 2 discussed the summaries in depth and iteratively coded them for emergent themes that captured how crafters talked about collaborating with materials. The themes included: 1) *sensory collaboration* with the sub-themes of *collaborating with textures* and *collaborating with visual elements* as well as 2) *genealogical collaboration* with the sub-themes of *collaborating with the past, collaborating with the present,* and *collaborating across craft forms.* We define and discuss the themes in depth in the findings section. To analyze the learning that is being produced during the material collaborative processes, we used the photographs to reconstruct the crafting processes that led to the production of new techniques and projects, evidence for the outcomes and learning that the craft experienced as it expanded itself.

Findings

Across the interviews, we identified two human-material collaborative learning processes: 1) *sensory collaboration*, which refers to how the people respond to the characteristics of materials through their senses and the kind of characteristics that the material makes available to people and 2) *genealogical collaboration*, which refers to the expansion of a material practice, like craft, across past, present, and unrelated materials. Collectively, the themes contributed to understanding how crafters reference collaborative learning processes in discussing their engagement with materials and how learning takes place in such moments of collaboration. We discuss the themes and, in the interest of space, focus more closely on one example of genealogical collaboration.

Sensory collaboration

Sensory collaboration included two sub-themes. First, *collaborating with textures* captured how crafters used touch to respond to the tactile qualities of the materials while the materials made these tactile qualities feelable to the crafter. Across the data, 65% (n=37) of the interviews included collaborating with textures. For instance, crafters explained how textured yarn or fabric drove new techniques, projects, and shapes as materials produced particular effects. For example, a 34-year-old crafter said: "Wool fiber has barbs, so it catches on each other as they twist. Silk doesn't do that". The doing and catching of the textured matter was presented as an active contributor to how the project unfolded–a presentation not uncommon in the data.

Second, *collaborating with visual elements* referred to how the crafters used sight to respond to the visual qualities of the materials while the materials made these qualities available. Across the data, 37% (n=24) of the interviews included this theme. Examples involved aesthetics and decoration (e.g., color, general) as the driving force behind the production of new and widely circulating techniques. For instance, Diana, a 53-year-old quilter explained how the *scribbling* technique came about in collaboration among crafter and material: "Free-motion, you're actually (...) scribbling–that's the best way I can describe it–with your needle. (...) The trick is to do it in a way so that the stitches are even (...). If I were going over lines, back and forth, it would look kind of messy, but instead, I'm drawing patterns." This excerpt presents how the visual and decorative elements of the craft materials led to a new technique within quilting, one that could be named and pointed to.

Genealogical collaboration

Genealogical collaboration included three sub-themes that observed multiple lineages of technological development. First, *collaborating with the past* converted how crafters and materials engaged across generations of people and materials. Of all interviews, 22% included mentions that were coded with genealogical, including how working with heritage material made it possible to honor, connect to, and expand (family) history by physically exploring and expanding material markers of the past. The materials made it possible to engage with family members and events that had long passed and to continue them.

Figure 1. Die and diamond-shaped pieces of fabric (left), die cutter (right).

Second, *collaborating with the present* captured how the integration of current technologies and techniques informed the becoming of craft and crafter as the crafter, by virtue of using the new technique, and the new technologies, by virtue of blending with the craft, collaboratively advanced what was possible with craft. Of the interviews, 52% (n=34) included this theme. One example was the integration of a sewing machine with quilting, where the machine produced possibilities for quilting to become more rapid and for new techniques to flourish, including machine appliqué. Where the craft expanded to become a machine integrated technique, the crafters became gift-givers, creative producers, and, in some cases, published authors.

Third, *collaborating across craft forms* captured collaboration of crafters with materials that were not initially part of the craft as well as how the original craft materials accepted these materials to produce new techniques expanding the craft. Of the interviews, 38% (n=25) included this theme, including integrating electronics into traditional fiber craft, dyeing fabric with non-traditional materials, painting fabric and printing photographs on fabric. For example, Jenna, a 61-year-old quilter, described that she had not used her die cutter (Figure 1) for a long time before it "called out to her" for cutting fabric. Die cutters are intended for paper crafts. In quilting, they speed up cutting fabric shapes that can be sewn together. Jenna and the die cutter worked together to produce the shapes. The crafter applied pressure, by turning the handle. The die was open to be used with different materials, including fabric, and the fabric accepted to be cut. This physically expanded the craft of quilting, growing lineages into paper crafts, as well as possibilities for new crafting techniques and designs.

Discussion and implications

This study showed human-material collaborative processes in which materials take on active roles. Learning outcomes–the physical expansion of the (craft) technologies of/in production–were contingent on the crafter *and* the craft materials. This was evidenced through physical technological development but did not belong to the human or the material. *Sensory collaboration* highlights surfaces in collaborative learning technologies and took serious aspects of human-material doing that would typically be considered secondary or unnecessary for learning (e.g., decorating). *Genealogical collaboration* made it possible to see an expansion of craft as a result of collaboration across craft forms and time. Although similarities are present to design in use, in which the purpose of materials can shift over time and with use (Ritella & Hakkarainen, 2012), the posthumanist lens highlighted the becoming of the craft and the crafter as physical evidence of learning through collaborative.

We cannot affirmatively say whether it was the crafter's insight or the materials presence that produced the expansion of the craft. In and of itself, this calls into question whether we can explain the repurposing of tools solely as human purpose. The study showed that the technology of quilting as much as the crafters developed over time. The study has implications for collaborative learning, presenting early evidence to consider including human-material collaborations as a form of collaboration. By seeing human-material collaboration as a driver of technological development and this development as evidence of human learning, a broader understanding of what counts as collaborative learning becomes possible. This is important because our study suggests that considering human-material collaborations make it possible to value practices and sensitivities that are otherwise at risk of being dismissed as secondary (e.g., decoration, feeling the material). For research this means closely attending to which materials to design with and what human-material processes are at play when capturing collaborative learning. More work is needed to consider collaborating with materials beyond adult crafts.

References

Keune, A. (2020). Fabric-based computing – (Re)examining the materiality of computer science learning through fiber crafts. [Doctoral dissertation, Indiana University]. ProQuest Dissertations Publishing.

Keune, A. & Peppler, K. (2019). Child-material computing: Material collaboration in fiber crafts. In Proceedings of *CSCL 2019*. Lyon, France: ISLS.

Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of Women in Culture and Society,* 28(3), 801-831.

Essinger, J. (2004). *Jacquard's web: How a hand-loom led to the birth of the information age.* Oxford University Press.

Gore, N. (2004). Craft and innovation: Serious play and the direct experience of the real. *Journal of Architectural Education*, 58(1), 39-44

Jackson, A. Y., & Mazzei, L. (2011). T*hinking with theory in qualitative research: Viewing data across multiple perspectives.* Routledge.

Kuby, C. R. (2017). Why a paradigm shift of 'more than human ontologies' is needed: Putting to work poststructural and posthuman theories in writers' studio. *International Journal of Qualitative Studies in Education*, 30(9), 877-896.

Mehto, V., Riikonen, S., Hakkarainen, K., Kangas, K., & Seitamaa-Hakkarainen, P. (2020). Epistemic roles of materiality within a collaborative invention project at a secondary school. *British Journal of Educational Technology.*

Ritella, G., & Hakkarainen, K. (2012). Instrumental genesis in technology-mediated learning: From double stimulation to expansive knowledge practices. i 7(2), 239-258.

Stahl, G. (2013). Theories of cognition in collaborative learning. In C. Hmelo-Silver, A. O'Donnell, C. Chan & C. Chinn (Eds.), *International Handbook of Collaborative Learning.* (pp. 74-90). Taylor & Francis.

Yrjönsuuri, V., Kangas, K., Hakkarainen, K., & Seitamaa-Hakkarainen, P. (2019). The roles of material prototyping in collaborative design process at an elementary school. *Design and Technology Education: An International Journal,* 24(2), 141-162.

Acknowledgments

This work was supported by a collaborative grant from the National Science Foundation (#1420303) awarded to Kylie Peppler. Any opinions, findings, and conclusions or recommendations expressed are not those of the National Science Foundation.

International Society of
the Learning Sciences

Negotiating Accountability and Epistemic Stances in Middle-School Collaborative Discourse

Asmalina Saleh, Chen Feng, Haesol Bae, Cindy E. Hmelo-Silver, Krista D. Glazewski
asmsaleh@iu.edu, carrfeng@iu.edu, haebae@iu.edu, chmelosi@indiana.edu, glaze@indiana.edu
Indiana University

Seung Lee, Bradford Mott, James Lester
sylee@ncsu.edu, bwmott@ncsu.edu, lester@ncsu.edu
North Carolina State University

Abstract: Group discourse has often been examined in computer-supported collaborative learning environments to develop adaptive scaffolds for students and intelligent cognitive assistants for teachers. This study leverages work in accountable talk and transactivity to focus on epistemics, or how participants establish, negotiate, and reproduce knowledge claims in collaborative discourse. We examined online chat data from 7 groups of middle school students who engaged in a game-based learning environment. Findings indicate that the framing of the game-based activities influenced how students interacted with one another. When attending to turns at talk, a combination of 1) epistemic status, 2) epistemic gradient, or the relationship between more and less knowledgeable persons, and 3) attention to norms of accountable talk generated more transactive discussions among students.

Background

Computer-supported collaborative learning environments (CSCL) are effective in supporting inquiry processes and associated learning outcomes (Chen et al., 2018; Jeong et al., 2019). Central to the success of collaborative learning is transactivity, which refers to the actions of students building on the ideas presented by others in discourse (Berkowitz & Gibbs, 1983). In the CSCL literature, transactive reasoning has been used effectively to support automated collaborative learning, conservational agents, and scripting (Dyke et al., 2013). Transactivity has also been adopted in other contexts, such as classrooms. Accountable talk for instance, shares similar concepts of reciprocity, holding students responsible for their reasoning, knowledge, and to the learning community (O'Connor & Michaels, 2019). Our work draws on these traditions to understand how knowledge is distributed in small group conversations as students engage in a collaborative game-based learning environment.

Accountable talk, transactivity, and epistemics

Accountable talk is a form of discursive classroom practice intended to support students in engaging in argumentation (O'Connor & Michaels, 2019). Although students can engage in discussion, they might not always participate equally in conversations. Accountable talk distributes responsibility among students by attending to accountability to knowledge, reasoning, and community. Accountability to knowledge refers to holding student responsible for grounding their knowledge claim whereas accountability to reasoning emphasizes the process of argumentation. Finally, accountability to the learning community includes valuing each student's contribution, distributing responsibility, and structuring group thinking. Similarly, Berkowitz and Gibbs (1983) define transacts as transactive reasoning that builds on prior actions from others. According to the authors, there are three types of transacts, representational, operational, and hybrid transacts. Representational transacts utterances that re-cast the prior contribution but does not modify it. This includes request for feedback or paraphrasing prior utterances. Operational transacts on the other hand, refine, or extend the prior contributions. Hybrid transacts are transacts that are in-between representational and operational transacts.

Taken together, the transactivity and accountable talk frameworks are concerned with epistemics, or how participants assert and negotiate knowledge claims through interactions with others (Heritage, 2012). In conversation, participants often demonstrate varying degrees of knowledge. The relative position between a less knowledgeable person (K-) and a more knowledge (K+) partner is referred to as *epistemic status*. This status is a jointly recognized among participants and establishes who has access to information, ability, and rights to the information. *Epistemic stance* refers to how a participant positions themselves through the design of turns at talk (Heritage, 2012). Although epistemic stance is influenced by one's epistemic status, they are not always the same. As a conversation progresses, one's relative epistemic stance shifts moment to moment because of the dynamics of interaction. For example, a teacher, with typically high epistemic status, might choose to start the discussion by taking a less knowledgeable epistemic stance, and ask questions about a concept, "Can someone explain what dissolved oxygen means?" This epistemic stance elicits participation from students, allowing students to adopt a more knowledgeable epistemic stance when responding. To explore knowledge generation in a game-based learning environment, we address the question, how do participants establish, negotiate, and reproduce knowledge claims in collaborative discourse?

Methods

Participants

Participants were 28 students (6th and 7th grade, 10 females, 18 males) in a public school in the United States. Of those who participated, 3 students identified as African/Americans, 4 as multi-racial, 2 as Asians/Pacific Islanders, 1 as Hispanic/Latinx, 1 as Native American/American Indian, and 17 as White. Students worked in groups of four and engaged with a game-based learning environment, CRYSTAL ISLAND: ECOJOURNEYS in two 120-minute sessions. Each group of students was assigned a facilitator who scaffolded collaborative inquiry discussions using accountable talk, such as marking information, eliciting participation, and revoicing student contributions.

Setting

In CRYSTAL ISLAND: ECOJOURNEYS, students arrive on a fictional island in the Philippines where they meet a fish technician, Jasmine, who has a problem: her tilapia are falling sick at an alarming rate. Students engage in two major inquiry activities, collecting information, and then brainstorming ideas about why the problem was happening. To share and negotiate data with their peers, students use an in-game collaborative brainstorming board (Saleh et al., 2020). At the board, students move notes to the columns that align to components that the tilapia fish needed to survive. Students then vote on whether the notes are relevant to the associated component. The votes on each note are represented visually: the note turns green when all students in the group agree or the note turns red otherwise. Throughout the activities, students use an in-game chat to discuss ideas with their peers.

Data sources and analysis

Data for this analysis was derived from all the group's in-game chat data. To analyze the data, we first segmented the chat according to the two major tasks presented to the students, collecting data and brainstorming. We segmented the chat data based on these inquiry phases and determined the topic of discussion in each segment. Drawing on interaction analysis (Jordan & Henderson, 1995), we focused on the temporal sequence of activities and turns at talk to determine how participants demonstrated their knowledge stances during their conversational turns. Based on our analysis, we found that one-minute intervals between the last utterance and the next utterance established the relevance between speakers. To understand how knowledge is generated in collaborative inquiry, we examine the relationship between less and more knowledgeable persons.

Results

Table 1 highlights the topic or framing of the discussion for each task in the inquiry cycle. In general, most of the discussion that occurred was centered around the brainstorming board tasks.

Table 1: Overview of topics categorized by group in each inquiry tasks

Tasks	Main topic of discussion	Groups
Exploration 1	Off-topic discussion	A, B, D
	Accountability to community: Describe tasks and content	C, E, F, G
Brainstorming session 1	*Accountability to reasoning, knowledge, and community*: Vote on relevance of notes to abiotic factors and eliminate factors that may not cause fish illness	All groups
Exploration 2	Off-topic discussion	A, B
	Accountability to knowledge and community: Describe tasks and content	C, E, F, G
	Limited chat	D
Brainstorming session 2	*Accountability to knowledge and community*: Discuss abiotic factors	A
	Accountability to reasoning, knowledge, and community: Discuss the abiotic factors and how they are related to one another	B, C, D, E
	Accountability to reasoning, knowledge, and community: Discuss the abiotic factors and how they are related to one another and the problem	F, G

Across all the discussions, we observed two ways that accountable talk framing influenced how participants established, negotiated, and reproduced knowledge claims in collaborative inquiry discourse. First, the group's orientation to accountable talk shaped how students approached their own knowledge across all the inquiry sessions. In the first exploration session, four groups demonstrated accountability to their learning community, by discussing the tasks that needed to be done and describing their in-game interactions with their peers. These groups continued to be focused on task knowledge in the first brainstorming session. At the start of each brainstorming session, all groups focused on the goal of the activity and the task that needed to be completed. The facilitator typically started conversational sequences by presenting the groups with information about what to do at the start of each phase. This meant that students' less knowledgeable (K-) stances usually involved task-oriented questions about the placement of notes on the board and who had the notes on the location of the notes. Facilitators often reframed these task queries into requests about what the content of the note is. From an accountable talk perspective, this reframing signals to students that they need to reflect on the information that is provided and justify their reasoning before moving the notes. In the second brainstorming session, groups discussed information associated with abiotic factors such as organic matter, dissolved oxygen, and cyanobacteria. With the exception of group A, most groups were able to discuss the relationships among these factors. Group A's progress may be related to the fact that the students engaged in off-topic discussions during both exploration phases.

Second, tracking the knowledge relations among participants, or the epistemic gradient was critical to understand collaborative inquiry and accountability in the group. For example, requests from teachers may appear to be a known answer or initiation-reply-evaluation format (Mehan, 1979), but our data reveals that these knowledge relations may be more nuanced. Specifically, when student requests for information are re-voiced by the facilitator, these requests bolster students' epistemic status, suggesting that answers are more likely to be answered. Additionally, collaborative inquiry involves sharing information that contributes to the knowledge community, even if it appeared as if students were minimally responding to one another. Although it may seem crucial for student to respond to one another, students sometimes pick up threads of conversations later in their discourse. We observe the following, when a student provides information, peers responded (1) with less knowledgeable stances (e.g., I didn't know that), (2) corroborate the information being shared (e.g., I found that out too, I agree with what they said), or 3) build on this information (e.g., I agree because …, I don't think so because …). By sharing information or indexing a more knowledgeable position in their collaborative discourse, students begin to generate a corpus of knowledge that all members can then access and in turn, support collaborative and individual understanding.

Accountable talk and epistemics in action

Because of space constraints, we present a short excerpt from group F to illustrate how accountable talk shape how participants establish, negotiate, and reproduce their knowledge claims (Table 2). In the first session, Turtle was present at the beginning but was called away for other school activities. As a result, Turtle had missed both exploratory phases. Fortunately, there were notes available for Turtle to review during the brainstorming session.

Table 2: Excerpt - Group F Facilitator (Wizard) requests information

	Time	User	In-game chat
1	09:58:53.1	Wizard	if you have to write an explanation, what would it be?
2	10:01:16.3	Turtle	i don`t (know) that much but isnt it obvious that too much organic matter is the problem
3	10:01:34.8	Jeepney	Yes.
4	10:01:37.0	Wizard	why is it obvious?
5	10:01:42.9	Sun	it just is
6	10:01:50.3	Turtle	also coronavirus isn't funny it is just a overused joke
7	10:01:52.2	Jeepney	But dont forget about the oxygen problem.
8	10:01:53.4	Wizard	explain it to me ;)
9	10:02:07.5	Eagle	Because all of the other problems connect and go back to organic matter
10	10:02:32.8	Sun	not the oxygen one
11	10:02:55.4	Eagle	ok, most of thhe problems
	[…]		
12	10:05:03.2	Wizard	but if she wants to know why, what would you say?

	[…]		
13	10:06:33.1	Turtle	so to much organic bacteria means more cyanide and that means competition for oxygen
14	10:06:47.3	Turtle	cyanobacteria i mean not cyanide

Prior to the excerpt, Turtle had made multiple interrogative requests, “I didn’t do this yesterday”, “I am confused”, “Who is Jasmine?”, “Where is her problem?”, “I am very confused because I do not know anything about Jasmine.” Students in the group responded to his request (accountability to the learning community). This knowledge is then leveraged by Turtle in above in response to the facilitator (accountability to reasoning). Although Turtle’s utterance begins with a hedge, “I don’t (know) that much”, he follows up with “isn’t it obvious” (line 2), students in the group either corroborate or extend his claim (lines 5, 7, 9, and 10). In addition, his multiple requests for clarification allowed Turtle to draw on his peers’ prior contributions to provide a more nuanced response about why the tilapia might be sick highlighting the group’s commitment to accountable talk.

Conclusion and implications

In this work, we examined how epistemics can support our understanding of how collaborative knowledge building occurs in a game-based learning environment. Knowledge positions are similar to transactive actions, but also accounts for the power relations between speakers and how adherence to norms may support group accountability to productive discourse. Groups that took longer to adopt norms associated with accountable talk were less transactive than groups who did attend to such norms. On the other hand, groups that appeared to value all dimensions of accountable talk were more transactive, even when considering off-topic discussions. Ultimately, attention to epistemics in student conversations and how it contributes to group discussion will set the stage for the development of computational models of collaborative learning and support the development of intelligent cognitive assistants for teachers.

References

Berkowitz, M. W., & Gibbs, J. C. (1983). Measuring the Developmental Features of Moral Discussion. *Merrill-Palmer Quarterly, 29*(4), 399-410. http://www.jstor.org/stable/23086309

Chen, J., Wang, M., Kirschner, P. A., & Tsai, C.-C. (2018). The Role of Collaboration, Computer Use, Learning Environments, and Supporting Strategies in CSCL: A Meta-Analysis. *Review of Educational Research, 88*(6), 799-843. https://doi.org/10.3102/0034654318791584

Dyke, G., Howley, I., Adamson, D., Kumar, R., & Rosé, C. P. (2013). Towards academically productive talk supported by conversational agents. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 459-476). Springer US. https://doi.org/10.1007/978-1-4614-8960-3_25

Heritage, J. (2012). Epistemics in Action: Action Formation and Territories of Knowledge. *Research on Language and Social Interaction, 45*(1), 1-29. https://doi.org/10.1080/08351813.2012.646684

Jeong, H., Hmelo-Silver, C. E., & Jo, K. (2019). Ten years of Computer-Supported Collaborative Learning: A meta-analysis of CSCL in STEM education during 2005–2014. *Educational Research Review, 28*. https://doi.org/https://doi.org/10.1016/j.edurev.2019.100284

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundations and Practice. *Journal of the Learning Sciences, 4*(1), 39 - 103. http://www.informaworld.com/10.1207/s15327809jls0401_2

Mehan, H. (1979). ‘What time is it, Denise?”: Asking known information questions in classroom discourse. *Theory into Practice, 18*(4), 285-294.

O’Connor, C., & Michaels, S. (2019). Supporting teachers in taking up productive talk moves: The long road to professional learning at scale. *International Journal of Educational Research, 97*, 166-175. https://doi.org/https://doi.org/10.1016/j.ijer.2017.11.003

Saleh, A., Yuxin, C., Hmelo-Silver, C. E., Glazewski, K. D., Mott, B. W., & Lester, J. C. (2020). Coordinating scaffolds for collaborative inquiry in a game-based learning environment. *Journal of Research in Science Teaching*(57), 1490-1518. https://doi.org/10.1002/tea.21656

Acknowledgments

This work is supported by the National Science Foundation through grants IIS-1839966, and SES-1840120.
Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Do all roads lead to Rome? An Expert Study to Assess the Immediacy of Strategies to Regulate Collaborative Learning

Kerstin Fett, Martin Greisel, Markus Dresel, Ingo Kollar
kerstin.fett@uni-a.de, martin.greisel@uni-a.de, markus.dresel@uni-a.de, ingo.kollar@uni-a.de
University of Augsburg, Germany

Abstract: When students work collaboratively, a wide range of different comprehension-related, motivational-affective, coordination- and resource-related problems may arise. To learn and collaborate effectively, these problems need to be regulated with the help of appropriate strategies. Different regulation strategies can however be differently immediate for the solution of different problems. By aid of an online questionnaire, we therefore asked *N* = 71 international experts from research on Computer-Supported Collaborative Learning (CSCL) to assess for individual problem types to what extent they felt that different regulatory strategies would be immediately effective or not. As a result of the analysis of the respective median for the individual strategies, it became apparent that, according to the experts, primarily but not exclusively comprehension problems should be regulated with cognitive strategies, coordination problems with metacognitive strategies, motivation problems with motivational strategies and resource problems with resource-oriented strategies. This has important implications for future interventions designed to support groups in effectively regulating their collaborative learning processes.

Introduction

At university, many students consciously decide to form self-organized small study groups, e.g. to jointly prepare for exams. This decision is easy to understand in view of the well-documented positive effects of collaborative learning on knowledge acquisition (Springer et al., 1999). However, research on collaborative learning shows that students unfortunately do not always make full use of the potential of collaborative learning (Weinberger et al., 2012). According to previous research (e.g., Järvenoja et al., 2019), self-organized collaborative learning may evoke various problems that can be divided into at least four categories: comprehension problems, coordination problems, motivational-affective problems and resource-related problems.

For self-organized collaborative learning to be successful, groups must be able to cope with such problems by activating appropriate strategies. Theoretical models of self-regulated learning (e.g., Panadero, 2017) assume that the choice of a strategy that fits the learning goal is crucial for regulatory success. Accordingly, it is assumed that not every strategy is equally well suited to solve a particular regulation problem (e.g., Engelschalk et al., 2016). The question of the fit between problems and strategies is particularly important because learners should be instructed to adjust their approach as adequately as possible depending on the situation.

To specify this fit more precisely, we developed the concept of "immediacy": A strategy can be said to be immediately effective for a problem if the strategy is generally suitable for the problem to disappear completely after applying the strategy correctly and with sufficient intensity. E.g, in case of unstructured learning material that consists of a large number of unconnected individual texts which do not seem to have any common thread, an organizational strategy would be an immediate strategy. A non-immediate, yet possibly still helpful strategy might be to motivate each other to continue working by offering a reward. In this case, although the group would continue studying, the strategy can not be considered as immediate because the actual problem—the lack of structure of the learning content—would not be solved. Previous studies indicate that comprehension problems should mainly be regulated with cognitive strategies, coordination problems mainly with metacognitive strategies, motivational-affective problems with mainly motivational strategies and resource problems with mainly resource-oriented strategies (Melzner et al., 2020). The present study however takes a closer look and asks what kinds of specific strategies (even within the broad categories just mentioned) are more immediate than others to solve a given regulation problem. We approach this question by aid of an expert survey.

Method

Sample

For this expert study, we approached all *N* = 2324 authors and co-authors who published papers in the CSCL conference proceedings from 2019, 2017, 2015 and 2011 via e-mail and invited them to participate in an online survey. Of the contacted persons, *N* = 71 experts rated at least one problem. Age was distributed as follows:

26–34 years = 25.4%, 36–45 years = 32.4%, 46–55 years = 28.2%, 56–65 years = 8.5%, and ≥ 66 years = 5.6%. About 47% of the study participants were male, and the average time participants had been working in science was M = 17.29 years (SD = 10.94). More than one fifth of the respondents were employed as professors (22.7%), almost one third (29.6%) as associate or assistant professors, and 16.9% were PhD students. The experts included in this study were all researchers who had first-authored at least one contribution in the field of collaborative learning, regulation in collaborative learning settings and/or individual self-regulated learning.

Procedure

After measuring socio-demographic information, participants received a short explanation of the concept of immediacy. Then, the survey presented 33 problems—based on previous problem typologies in the literature (Melzner et al., 2020)—that may occur during self-organized collaborative learning. For each problem, participants were asked to rate 27 strategies on a scale from 1 (*not immediately effective at all*) to 5 (*very much immediately effective*) which were based on strategy typologies from the literature (Melzner et al., 2020). Since participants were not expected to complete the whole questionnaire because of its length, problems were presented in randomized order to balance the number of responses for each problem. The number of respondents per problem varied between N = 12 and N = 20.

Analysis

We assumed that a strategy can be regarded as immediately addressing the respective problem if half of those who assessed it rated a strategy as at least somewhat immediately effective, i.e. strategies that had a median of 3 or greater (theoretical midpoint of the scale) were classified as immediate strategies for the problem at hand.

Results

An overview of the allocation of immediately effective strategies for the individual problems can be seen in Table 1. According to the experts, a percentage of M = 33.44% (SD = 25.93%) of all assessed strategies was at least “somewhat immediately effective” (≥ 3); a share of M = 20% (SD = 17.16%) were rated as ≥ 4 and thus seen as at least “rather immediately effective” and M = 8.53% (SD = 11.16%) as “very much immediately effective” (= 5). Further, Table 1 shows that problems from one kind (e.g., motivational problems) were mostly regarded to best be regulated by strategies from the one category of strategies that best apply to them (in this case: motivational strategies). Yet, for all problem categories, also strategies from other categories were listed to be immediate (this applied in particular to comprehension and coordination problems).

Table 1. Possible problems in collaborative learning settings and regulation strategies immediately addressing these problems according to participants (median of experts’ ratings in parentheses).

Problem	Regulation strategies rated as at least „immediately effective“			
	Cognitive	Metacognitive	Motivational	Resource-oriented
Comprehension Problems				
Unclear Task Definition		PRL (3)		
Unclear Procedure		PRL (4)		
Deficits in Prior Knowledge	CGP (5)			
Difficult Learning Content	OS (3), SIC (5), CGP (3), RDU (4)	PRL (3), REO (3)		ERM (3)
Too Complex Learning Content	OS (4), SIC (4), CGP (3)	PRL (3,5)		ERM (3)
Unstructured Learning Content	OS (4), SIC (3)	PRL (3)		
Coordination Problems				
Inefficient Use of Time		PRL (4)		TMC (5), AM (3), EM (3)
Unfair Distribution of Work Load		PRL (4)	HUG (3)	TMC (3), EM (3), FSA (3)
Lacking Procedural Fairness		PRL (3)	HUG (4)	FSA (3,5)

	Cognitive	Metacognitive	Motivational	Resource-oriented
Coordination Problems				
Differing Technical Understanding	SIC (3), RDU (5)			
Differing Goals	RDU (3)	PRL (4)	HUG (4)	FSA (3)
Incompatible Working Methods		PRL (4)		
Communication Problems		PRL (3), REO (3)	HUG (3)	TMC (3), FSA (4)
Poor Relationship Quality			HUG (4)	FSA (5)
Lack of Information Exchange		PRL (3.5)	HUG (4)	EM (3), FSA (3.5)
Motivation Problems				
Low Value of Learning Method			RS (3), SIT (3), HUG (4)	
Low Usefulness of Learning Content			SIT (3), IPS (4)	
High Costs of Learning Content			SIT (3), IPS (4)	
Low Intrinsic Value of Learning Content			RS (3), SIT (4), IPS (5)	
Low Personal Meaning of Learning Content			SIT (3), IPS (5)	
Procrastination		PRL (3)	RS (3), SIT (3) IPS (3)	TMC (3), EM (3)
Low Self-efficacy Expectation			MPS (3), AST (4)	
Resource Problems				
Lack of Time		PRL (4)		TMC (5), EM (3)
Unfavorable Surrounding Environment				EC (5)
Distraction				EC (4), AM (4)
Undesirable Private Conversations			SIT (3), HUG (4)	AM (5)
Lack of Learning Materials				KIM (4), ERM (4)
Physical Problems				TMC (3), EM (3), FSA (3)
Negative Emotions			SIT (3), MEC (4)	FSA (3)
Insufficient Technical Equipment				ERM (3), UAT (4), RTK (4)
Weak Technical Performance				UAT (5), RTK (4.5)
Lack of Technical Functionality				UAT (5), RTK (3), ATK (3)
Lack of Technical Skills				UAT (3), RTK (4), ATK (5)

Note. *Cognitive*: Organizational Strategies (OS), Strategies for Improving Comprehension (SIC), Strategies for Closing Gaps in Prior Knowledge (CGP), Strategies to Resolve Differences in Understanding (RDU); *Metacognitive*: Planning and Regulation of the Learning Process (PRL), Reflection and Evaluation of the Learning Outcomes (REO); *Motivational*: Reward Strategies (RS), Increasing Situational Interest (SIT), Increasing Personal Significance (IPS), Mastery and Performance-Related Self-Talk (approach and avoidance) (MPS), Ability-Related Self-Talk (AST), Highlighting Group Utility as a Goal (HUG), Management of Emotional Contagion (MEC); *Resource-oriented*: Time Management and Coordination (TMC), Environment Control (EC), Knowledge and Information Management (KIM), Attention Management (AM), Effort Management (EM), External Resource Management (ERM), Fostering a Positive Social Climate (FSA), Use of Alternative Tools (UAT), Recourse to Technical Knowledge for Handling Work Equipment (RTK), Acquisition of Technical Knowledge (ATK); *Only those regulatory strategies that were considered to be immediately effective for a problem at least once are listed*.

Discussion, limitations and conclusions

The purpose of the present study was to examine which regulatory strategies for different problems can be classified as immediately effective, according to experts from the CSCL community. By and large, we found that the common assumption (Melzner et al., 2020) that comprehension problems should best be regulated by employing cognitive strategies, motivational problems by motivational strategies, coordination-related problems by coordinative strategies, and resource-related problems by resource-oriented strategies was supported by the expert ratings. The theoretical assignments to the fit of problems and regulation strategies (e.g., Engelschalk et al., 2016), especially in collaborative learning settings (Melzner et al., 2020), were thus supported by the present results. Yet, we also obtained evidence that previous general classifications might be too simple in some cases. For example, in case of a procrastination problem, experts agreed that it would not be sufficient to use any motivational regulation strategy. Here, it may make sense to also resort to resource-oriented strategies such as Effort Management (EM) or the Planning and Regulation of the Learning Process (PRL) (metacognitive strategy).

In addition, the different size of the median for the individual assessments provides an indicator not only of whether, but also of the extent to which the strategies for the individual problems are—according to the experts' judgements—immediately effective or not, with higher medians indicating greater immediacy than lower medians. Consequently, our data indicates that for example the problem "Lack of Technical Skills" might be best regulated by the Acquisition of Technical Knowledge (ATK) ($Mdn = 5$), while the Recourse to Technical Knowledge for Handling Work Equipment (RTK) would be somewhat less immediately effective ($Mdn = 4$), and the Use of Alternative Tools (UAT) would still be usable, but least preferable ($Mdn = 3$).

In the present study, two decisions we made for the data evaluation might well be criticized: On the one hand, the median of 3 was used as a threshold value to distinguish immediate from non-immediate strategies. On the other hand, the expertise of the participants was solely tied to their first authorships. It would be conceivable to set a higher median for the examination of immediacy and to use the assessments of other indicators of expertise in the corresponding fields for the selection of suitable experts. Moreover, experts' self-reported judgements only provide hints to but are not equivalent to the actual effectiveness of regulation behavior in real learning settings.

Despite these limitations, our results bear important implications for the design of scaffolds to support self-organized collaborative learning. Based on our results, it would seem promising to closely monitor the kinds of problems groups encounter during collaboration and to then prompt students to apply strategies that our study identified as immediate. It is likely that the growing field of Learning Analytics (e.g., Ferguson, 2012) might, in the future, develop algorithms to diagnose current problems and to fade appropriate support in and out as needed.

References

Engelschalk, T., Steuer, G., & Dresel, M. (2016). Effectiveness of motivational regulation: Dependence on specific motivational problems. *Learning and Individual Differences*, *52*, 72–78.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning, 4*(5/6), 304–317.

Järvenoja, H., Näykki, P., & Törmänen, T. (2019). Emotional regulation in collaborative learning: When do higher education students activate group level regulation in the face of challenges? *Studies in Higher Education, 44*(10), 1747–1757.

Melzner, N., Greisel, M., Dresel, M., & Kollar, I. (2020). Regulating self-organized collaborative learning: The importance of homogeneous problem perception, immediacy and intensity of strategy use. *International Journal of Computer-Supported Collaborative Learning, 15*(2), 149–177.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422.

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.

Weinberger, A., Stegmann, K., & Fischer, F. (2012). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior*, *26*(4), 506–515.

PearProgram: Towards Fruitful Collaboration in Computing Education Research and the Learning Sciences

Jorge E. Garcia, Maxwell Bigman, Ethan Roy, Miroslav Suzara, Roy Pea
jorgeedu@stanford.edu, mbigman@stanford.edu, ethanroy@stanford.edu, msuzara@stanford.edu, roypea@stanford.edu
Stanford Graduate School of Education

Abstract. Computer Education Research has not made widespread use of relevant research in the Learning Sciences. Our research seeks to understand pair programming learning dynamics using established Learning Sciences frames to improve pedagogical methods and practices in Computer Science learning. We have developed *PearProgram* to work as both a learning tool and a data-gathering tool. Students who used the tool engaged more actively in the courses and attained higher pass rates. Additionally, analysis of qualitative data points to the importance of *Dual-Space Model* and *Positioning* frameworks in understanding the dynamics of a generative pair programming session. These findings suggest careful attention by the facilitator to offer structure and explicit decision making about roles can shape pair success in assignments and augment implicit learning.

Introduction

The emergence of the *Computer Science for All* movement brought research questions from Computing Education Research (CER)—the study of how people learn and teach computing—into adjacent education fields, including Learning Sciences (LS). Despite the emergence of best practices in computer science (CS) education addressing key research questions of how people can learn computing more effectively and equitably, the CER field may benefit from Learning Sciences research. Pair programming, a practice where two programmers work together on a single program, is an example of a CS education pedagogy with distinct roots in the collaborative learning literature, yet has not previously drawn from LS research, nor from computer supported collaborative learning (CSCL). Pair programming is considered a best practice in K-12 and university CS education, yet we have not found evidence of widespread utilization of relevant insights from the learning sciences used to inform how pair programming might leverage collaborative learning research.

In this paper, we present *PearProgram*, an online collaborative learning and research tool to help introductory CS students learn how to pair-program based on theory (for more, see www.pearprogram.com/#/about). We describe the theoretical foundation of the tool, promising results from early usage, and future research directions. We specifically highlight the potential importance of considering Positioning Theory when students work in (CS) collaborative learning settings. We aim to address the field between CER and LS research so that both may benefit from key findings on these discipline-specific collaborative learning practices that might advance new understandings for LS and CSCL researchers.

Pathways to pair programming insights

Pair programming is a K-12 and university best practice for CS education developed by professional software engineers in the ‘90s (Beck, 2000); its documented benefits for programmers of all levels has a long history (Williams et al., 2000). Twenty years of CS education research, including two rigorous meta-analyses, has demonstrated that “if implemented properly” pair programming has positive effects on students’ programming assignment grades, exam scores, and persistence, especially in introductory CS classes (de Lima Salge & Berente, 2016; Umapathy & Ritzhaupt, 2017). Other benefits of pair programming include greater confidence (Hanks et al., 2004), deeper conceptual understanding (McDowell et al., 2006), and continued interest in the topic (Littleton & Howe, 2010; Werner et al., 2004). However, studies have documented that the benefits of pair programming are not always realized, suggesting unexplored phenomena may play a role.

Grounded by numerous studies on academic achievement showing students tend to learn better collaboratively than alone (Abrami & Chambers, 1996; O’Donnell & Dansereau, 1992), we nevertheless concur with Dillenbourg et al. (1996), who aver that: “collaboration is in itself neither efficient nor inefficient. Collaboration works under some conditions, and it is the aim of research to determine the conditions under which collaborative learning is efficient”. Our aim in developing the *PearProgram* tool is to better understand the conditions that lead to pairs engaging in “efficient” (or generative) collaborative learning when pair programming, and how we might create the most fertile learning environments for all students learning CS.

Sparked by the global pandemic, learning experiences increasingly rely on computing tools and open up possibilities for understanding distributed learning. Our research, and the associated tool, *PearProgram*, is beginning to yield insights into how learning can unfold between pairs of distanced student learners. We were surprised to learn that our tool supported student engagement despite the isolation wrought by the pandemic. Our three aims in developing the *PearProgram* tool are: 1) To help introductory CS students learn how to pair program effectively; 2) To better understand how students pair program by capturing quantitative and qualitative data; 3) To research what factors lead to generative learning outcomes in different contexts.

PearProgram: A theory driven learning tool

Our understanding of pair programming is rooted in socio-constructivist theories of learning (Piaget 1926; Vygotsky 1978), underscoring the importance of social aspects of learning, with collaboration a primary means for thinking and learning. Our design choices stem from Barron's (2003) illumination of the complexities of collaborative work, theorized in terms of a "dual-space model" required for participants' collaborative problem solving, in which students need "to clarify how the content of the problem and the relational context are interdependent aspects of the collaborative situation." Prior CER work studied pair programming via elements of problem-solving processes (Williams et al., 2008), but has yet to focus on the corresponding and interdependent relational context. Our tool uses Barron's dual-space model to establish the importance of having students generate, confirm, document and reflect on one another's proposals so that pairs may more effectively collaborate. Research using our tool helps us understand how to support students navigating these complex spaces. To our knowledge, learning tools have not previously been designed to help students navigate these dual spaces in CS education. Since pair programming is a clear example of collaborative learning, our research aims to contribute to understanding what roles tools, behavioral nudges, and LS research can contribute to the complex task of such dual problem space problem solving.

Drawing on these theoretical foundations, the *PearProgram* tool allows us to establish turn-taking norms by designating specific roles for students (as pilot and co-pilot) with embedded role-changing prompts within a collaborative text editor. The tool is designed to be a student's first exposure to pair programming, inculcating collaborative practices that help students navigate the dual problem space of learning how to program collaboratively. Students are given equitable access to problem-solving processes by explicitly stating in the tool that "mutual understanding is *the* learning goal." While the implicit learning goal for students when pair programming typically is that they complete the programming task, individually learning the key concepts related to the task at hand, in *PearProgram*, we layer a new learning goal: that students help one another understand every line of code. Moreover, we draw on research suggesting that effective collaborations require that students view their success as mutually dependent; this leads to dialog that fosters new, co-created ideas and works to establish a shared social reality (Rogoff, 1990; Roschelle & Teasley, 1995). Accordingly, the tool provides embedded tips which emphasize that success *and* learning of the pair is a mutually dependent process.

Additionally, our tool emphasizes questions and annotations of the code as key mechanisms for driving learning. *PearProgram* has buttons for questions and comments to encourage students to ask questions and prompts students to establish joint attention (Tomasello, 1995). Building on Schneider and Pea's (2013) understanding of joint attention as both physically and verbally constructed, the *PearProgram* tool makes use of the programming environment to visually highlight lines of code when students ask their partner a question. Utilizing visual synchronization in the tool, we theorize that there will be better coordinated joint attention, which will increase the quality of collaboration and learning. Moreover, by actively engaging in learning-by-teaching (Goodland & Hirst, 1989), we anticipate mutual benefit for students using the tool.

Our design similarly draws on the "dual space model" as a mechanism for research analysis as we capture data that illuminates joint meaning-making activity during pair programming learning sessions. In particular, we draw on the dimensions of Meier, Spada and Rummel's (2007) framework for assessing the quality of computer-supported collaboration processes as a way to analyze our pair programmers; these include "sustaining mutual understanding, dialogue management, information pooling, reaching consensus, task division, time management, technical coordination, reciprocal interaction, and individual task orientation" (p. 63). As we analyze our data, these dimensions offer key metrics for understanding what activities effective pairs engage in that less effective dyads do not. Based on our preliminary analysis, these nine important dimensions nonetheless elide an important aspect of collaborative interaction: *Positioning*.

Positioning in pair programming

Davies and Harré's (1990) Positioning Theory offers an analytic lens and an explanatory theory for how discourse interacts with learning and identity development: "as an explanatory theory, Positioning Theory serves as a set of guiding principles for investigating the consequences of the discourse and the interactions of, and with, particular students and groups of students as they assume or reject particular positions or acts of positioning" (Green et. al , 2020). As a lens for understanding pair programming interactions, Positioning Theory offers a conceptually-driven approach to understand positions, acts, and storylines (Davies & Harré, 1990; Harré & van Langenhove, 1999). Incorporating Positioning Theory as an analytic tool in the "dual space model" of collaborative problem solving may reveal new insights into how students navigate the relational elements of collaborative work. Positioning Theory is useful as we center questions of power and privilege in the Learning Sciences and in CS (Annamma & Booker, 2020; Vakil, 2018). It is essential to move toward justice-centered approaches to understanding equity in CS and collaborative learning.

Preliminary findings

PearProgram for in-class activities in three remote introductory CS classes has led to results suggesting the tool has purchase for advancing student learning and research insights. Our post-use survey for instructors indicated that all instructors felt the tool positively contributed to the course; likewise students' post-use survey indicated they had improved as coders and wanted to engage further with *PearProgram*. Further, students who used the tool engaged more actively in the online class discussion forum and attained higher pass rates. While these findings are correlational and cannot ground causal inference, they suggest merits for further investigating the conditions contributing to efficacious uses of *PearProgram*. Moreover, our initial qualitative data illuminates a fertile research topic: positioning in pair programming and collaborative learning. Video analyses of pair programming interactions and student interviews surface findings that contribute to both CER and CSCL fields.

Beyond the quantitative data described, we video-recorded pair work involving lists in Python and conducted some post-use interviews. We focused on one pair who finished their task in ⅓ the time of the others and engaged in unique behaviours compared to other groups. This pair centered the conversation of roles (pilot/co-pilot) in their first ½ minute of interaction. In this interaction, we conjecture that relational work was being accomplished in positioning one another as peer learners, and accruing early practices in joint decision making revealed that after this initial decision, talk time was evenly distributed between the learners throughout the rest of the session compared to other pairs, suggesting the value of this interaction:

Student 1:	So do you want to be pilot or co-pilot? I'm cool with whatever.
Student 2:	Um, right now it says I'm co-pilot. Does it say you're pilot?
Student 1:	Yeah, but if you wanna switch that's like cool with me too. It's up to you.
Student 2:	Um, maybe we can switch like halfway or something? I don't know.
Student 1:	Ok. Yeah that works. Ok. um.

Future study designs

As we design future studies, we use the insights of Positioning Theory to investigate unexamined dimensions of collaborative learning that may prove crucial for understanding the nature of computer-supported collaborative learning environments for students, particularly in the CS domain, fraught with positioning challenges (e.g., stereotype threat). As we progress with the *PearProgram* project, we are conducting contrasting case analyses, comparing successes and difficulties with in-person, remote and hybrid pair programming, as we look for 'positive deviance'—"intentional behaviors that depart from the norms of a referent group in honorable ways" (Spreitzer & Sonenshein, 2004). As learning environments are increasingly technology-mediated and span remote settings in the Covid-19 era, careful attention to the learning tools, the metrics of success, the complexities of navigating collaborative learning environments, including relational spaces, takes on added significance. For educational tools, we adopt the approach that technology should create experiences that go "beyond being there" (Hollan & Stornetta, 1992)—establishing learning opportunities uniquely possible in online settings. To that end, we believe we are also creating research opportunities that can go "beyond being there".

References

Abrami, P. C., & Chambers, B. (1996). Research on cooperative learning and achievement: Comments on Slavin. *Contemporary Educational Psychology*, 21(1), 70-79.

Annamma, S.A. & Booker, A. (2020). Integrating intersectionality into the study of learning. In N. Nasir, C. Lee, R. Pea, & M.M. De Royston. (Eds.). *Handbook of the cultural foundations of learning* (pp. 297-313). New York: Routledge.

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, 12(3), 307–359.

Beck, K. (2000). *Extreme programming explained: embrace change*. New York: Addison-Wesley Professional.

Bigman, M., Roy, E., Garcia, J. E., Suzara, M., Wang, K., Piech, C. (2021, in press). PearProgram: A more fruitful approach to pair programming. In Proceedings of the *52nd ACM Technical Symposium on Computer Science Education* (SIGCSE '21).

Davies, B., & Harré, R. (1990). Positioning: The discursive production of selves. *Journal for the Theory of Social Behaviour*, 20(1), 43–63.

de Lima Salge, C. A., & Berente, N. (2016). Pair programming vs. solo programming: what do we know after 15 years of research? In 2016 *49th Hawaii International Conference on System Sciences* (HICSS) (pp. 5398-5406). Piscataway, NJ: IEEE.

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 machine: Towards an interdisciplinary learning science* (pp. 189–211). Oxford: Elsevier.

Green, J.L., Brock, C., Baker, W.D., & Harris, P. (2020). Positioning theory and discourse analysis: an explanatory theory and analytic lens. In N. Nasir, C. Lee, R. Pea, & M.M. De Royston. (Eds.). *Handbook of the cultural foundations of learning* (pp. 119-140). New York: Routledge.

Goodlad, S., & Hirst, B. (1989). *Peer tutoring. A guide to learning by teaching*. New York: Nichols Publishing.

Hanks, B., McDowell, C., Draper, D., & Krnjajic, M. (2004). Program quality with pair programming in CS1. In Proceedings of the *9th annual SIGCSE conference on innovation and technology in computer science education* (pp. 176-180).

Harré, R., & van Langenhove, L. (1999). *Positioning theory*. Oxford: Blackwell.

Hollan, J., & Stornetta, S. (1992). Beyond being there. *SIGCHI Proceedings* (pp. 119-125).

Littleton, K., & Howe, C. (Eds.). (2010). *Educational dialogues: Understanding and promoting productive interaction*. New York: Routledge.

McDowell, C., Werner, L., Bullock, H. E., & Fernald, J. (2006). Pair programming improves student retention, confidence, and program quality. *Communications of the ACM*, 49(8), 90-95.

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.

O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In N. Miller & R. Hertz-Lazarowitz (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 120-141). New York: Cambridge University Press.

Piaget, J. (1926). *The language and thought of the child*. London: Routledge & Kagan Paul.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford: Oxford University Press.

Roschelle, J, & Teasley, S.D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), NATO ASI Series,Vol. 128: *Computer supported collaborative learning*. Berlin, Heidelberg: Springer.

Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-supported collaborative learning*, 8(4), 375-397.

Spreitzer, G. M., & Sonenshein, S. (2004). Toward the construct definition of positive deviance. *American Behavioral Scientist*, 47(6), 828-847.

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: Erlbaum Associates.

Umapathy, K., & Ritzhaupt, A. D. (2017). A meta-analysis of pair-programming in computer programming courses: Implications for educational practice. *ACM Transactions on Computing Education* (TOCE), 17(4), 1-13.

Vakil, S. (2018). Ethics, identity, and political vision: Toward a justice-centered approach to equity in computer science education. *Harvard Educational Review*, 88(1), 26–52.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.

Werner, L. L., Hanks, B., & McDowell, C. (2004). Pair-programming helps female computer science students. *Journal on Educational Resources in Computing* (JERIC), 4(1), 4-es.

Williams, L., Kessler, R. R., Cunningham, W., & Jeffries, R. (2000). Strengthening the case for pair programming. *IEEE software*, 17(4), 19-25.

Williams, L., McCrickard, D. S., Layman, L., & Hussein, K. (2008). Eleven guidelines for implementing pair programming in the classroom. In *Agile 2008 Conference* (pp. 445-452). IEEE.

Online Design Thinking Faculty Development Workshops: A Design-Based Research Study

Jonan Phillip Donaldson, Daniel Choi, Jean Layne
jonandonaldson@tamu.edu, danielchoi858@tamu.edu, j-layne@tamu.edu
Texas A&M University

Abstract: The Design Thinking for Engaged Learning (DTEL) is a framework for educators to engage students in collaborative project-based learning. The model integrates strategies for designerly ways of knowing into the design thinking process. In the past, the DTEL has been taught as a synchronous, in-person workshop; however, to comply with Covid-19 safety regulations, the DTEL workshop has been modified to allow for dissemination through a synchronous, online modality. The present study examines the efficacy of the virtual DTEL workshop, provides a comprehensive explanation of the theoretical grounding and methodology, and interprets the results through the lens of Design-Based Research (DBR). Limitations of the current study and suggestions for future researchers are discussed.

Background

In the researchers' experiences working with educators in major research universities and other higher education settings, we have found that one of the most pervasive barriers to the sustained implementation of evidence-based, high-impact practices is a perceived failure of initial attempts. These failures appear to result from a lack of structure through which to scaffold learning in high-impact practices such as project-based learning and collaborative learning. Traditional "low-impact" practices (e.g., lecture-dominated courses, memorization, exams) continue to represent the dominant modality (Finelli, Daly, & Richardson, 2014; Hora, 2014; Hora & Ferrare, 2013; Oleson & Hora, 2014); however, researchers acknowledge that the low rate of adoption of high-impact practices is a concern in undergraduate (Campbell, Cabrera, Michel, & Patel, 2017) as well as K-12 education (Coldwell et al., 2017; Hazelkorn et al., 2015). This paper reports on a design-based research (DBR) study investigating the design and implementation of an online workshop for faculty in higher education, which helps them develop expertise in using design thinking as a way to structure project-based and collaborative learning in their courses.

Design Thinking for Engaged Learning (DTEL, Donaldson & Smith, 2017) is a model developed to serve as a framework for collaborative project-based learning (Donaldson, Barany, & Smith, 2020). This model incorporates designerly ways of knowing into the five sequential phases of the design thinking process: 1) Name and Frame, 2) Diverge and Converge, 3) Prepare and Share, 4) Analyze and Revise, and 5) Deploy. The designerly ways of knowing emphasized in the DTEL model are cognitive strategies used by expert designers such as wicked problem framing, abductive reasoning, and reflection-in-action. These strategies are integrated into particular stages of the DTEL process.

In the first phase, Name and Frame, students form groups and identify real-world problems that could be addressed through the design thinking process. The problem identification process constitutes gathering information about the problem through human-centered design methodologies such as interviews and observations, and is followed by a period of re-framing the problem as a wicked problem – a problem that is difficult or impossible to resolve because of incomplete, contradictory, or changing underlying factors that are difficult to recognize. In the next phase, Diverge and Converge, students uninhibitedly brainstorm as many possible solution ideas for the selected wicked problem as possible. Traditionally, this is done by writing or drawing each idea on a sticky note and placing them on a wide-open surface such as a wall. Once the group members have individually generated a sufficiently large quantity of unique ideas, they silently rearrange the ideas into clusters, informed by the contents written on the post-it notes. Afterwards, group members discuss the contents and overarching themes of each cluster and select the one to which they are drawn. Then they undergo a process of refining and synthesizing the ideas within the cluster, until it is possible to craft an initial solution statement. After a period of project planning, each member creates a low-fidelity prototype and presents their unique solution with the other group members. The group then collaboratively decides upon one low-fidelity prototype for which to develop a high-fidelity prototype – a mockup that is as close to a true representation of the final product as possible. Finally, the high-fidelity prototype is deployed and tested in a real-world context and data about the users' experiences are collected for analysis. Subsequently, the findings are translated into design moves, which are used to revise the prototype until a final version is created and ready to be deployed.

We have implemented, reiterated, and investigated the DTEL model in a variety of face-to-face settings, ranging from undergraduate and graduate courses (Donaldson, Barany, & Smith, 2020), pre-conference workshops (Donaldson & Barany, 2017; Wheeler, Trausan-Matu, Donaldson, & Barany, 2019), corporate trainings, and faculty development workshops with positive results. However, in March of 2020, the Covid-19 pandemic made it necessary to redesign the workshop to be administered through an online modality. Prior to this, the researchers had conducted a number of studies to investigate the ideation (divergent and convergent thinking) stage of the DTEL process in digitally-mediated contexts, but results were not promising (Donaldson & Barany, 2020). The urgency of the circumstances of the current study provided a unique opportunity to develop innovative solutions to the model and critically evaluate these solutions to inform design moves for future iterations.

Methods

We used design-based research (DBR) methodology (Sandoval & Bell, 2004) to develop the online DTEL workshop, make improvements for future iterations, and speak back to the theoretical foundations in which it is grounded. For the design of our online DTEL workshop, the researchers determined that it was appropriate to ground the program in constructivism (Piaget, 1952; Vygotsky, 1978), constructionism (Kafai, 2006), and situated learning (Lave & Wenger, 1991) as the theoretical foundations. Furthermore, we translated the model, informed by these theories, into a set of principles which were to construct the first version of the workshop.

The first online iteration of the workshop was structured the same as the face-to-face workshop taught by the researchers in the past, with the only change being the digital tools that were used. There were four synchronous online 90-minute sessions, administered over a four-week period, during which participants worked through the DTEL process in groups of 7-10 members. The sessions were conducted via Zoom, a web conferencing tool, and participants spent the majority of the time in breakout rooms with their group members, working synchronously in the group's Google Document or Miro board, which they were assigned to by the researchers. At the opening of each session, the facilitators spent approximately 15 minutes introducing participants to one of the stages of the DTEL process. The conceptual aspects and activity instructions were accompanied by slides.

After each synchronous session, participants completed a reflection assignment and the researchers wrote memos. Furthermore, after the completion of the 4-week workshop, participants were interviewed. The participants' reflections, interview transcripts, researcher memos, and design elements (e.g., slides, Google documents, Miro Boards) were collected and anonymized to be used as data, intended to inform design moves for future iterations of the workshop. All data artifacts were entered into the MAXQDA Pro analysis software and coded using in-vivo coding strategies with an emphasis on participant reporting of problems, frustrations, and descriptions of their collaborative work with their teams. Analysis of the data was conducted to determine what aspects of the workshop design were effective and what aspects needed improvement. Finally, the findings were translated into design moves, which were then implemented in subsequent iterations of the online DTEL workshop for a total of three iterations.

Findings

Analysis of the data collected from the first iteration of the online DTEL workshop resulted in several issues. First, there was an issue regarding group members not collaborating effectively. This led to the design modification of adding a role-negotiation step to the first session of the workshop. Another problem, based on the researchers' observations was that participants were confused about what exactly they were supposed to be doing during the high-fidelity prototyping stage. Specifically, the researchers noticed that during this phase, individuals were trying to create their own prototypes, rather than collaboratively developing a single prototype as a group. This led to the design move of adding clarification to the instructions for this stage. Finally, the most concerning problem pertained to the issue of figuring out how to use the digital technologies – Zoom video conferencing platform, the Google documents, and the Miro boards. The researchers attempted to address this by providing participants with informative resources explaining how to use these technologies, prior to the beginning workshop; however, we anticipated that this solution would only resolve this issue for some of the participants.

We ran the workshop again, implementing the design moves based on the findings from the first iteration, and collected further data from the second iteration with new participants. Analysis of this data indicated that although some of the design moves – modifications to the workshop – had been successful, new issues had emerged. The addition of the role negotiation component significantly improved group members' collaborative work. Furthermore, clarifying the instructions regarding the expectation for the high-fidelity prototyping stage was also effective, and all of the groups clearly understood the objective of this phase in the workshop. Providing the participants with resources about how to use the technologies, prior to beginning in the workshop, seemed to be more effective than we had anticipated, as the participants were able engage in the activities without any significant issues or delays. However, this may have also been due, in part, to the addition of the role negotiation between group members because some participants reported that their team members helped them resolve these technological issues.

Although the design moves made in the second iteration of the workshop were promising, new issues were found during the analysis of the data from the second iteration as well. First, participants overemphasized using the design thinking process to develop solutions for their own classroom problems, and thus overlooked the primary purpose of the DTEL model as a structuring framework to facilitate students' collaborative project-based learning. The researchers incorporated frequent reminders and discussions throughout the workshop to address this issue. Second, the participants were not practicing human-centered design skills (empathy, perspective-taking), which may have been an issue that the researchers failed to identify, but was present in the first iteration, as well. To address this issue, we created weekly homework assignments between the synchronous workshop sessions in which participants were asked to engage in empathy exercises, such as conducting interviews with members of the particular population for which their group was designing a high-fidelity solution.

We made changes to the design and conducted the third iteration of the workshop and collected more data. Analysis found that the design changes in the second and third iterations were having the intended effects. The frequent reminders and discussion of the purpose of the DTEL framework was effective in keeping participants more focused on having their own students use the process. Furthermore, the empathy exercises between live sessions increased the human-centered approach for those participants who completed them, but the majority of participants reported not being able to find the time.

Analysis of the third iteration data illuminated more fine-grain issues. These included the need for specific examples—especially images—of work prior to each stage of the process, need for clarification of some design-specific terms such as "prototype," and the need to provide optional readings and resources for participants who wished to explore particular concepts in greater depth on their own. More broadly, we found that a few participants were feeling rushed or overwhelmed. The three online iterations of the four-week workshop were completely synchronous, with all activities conducted during four 90-minute sessions. In future iterations, we will leverage our findings to create an online workshop with a mixture of synchronous and asynchronous work spread over five weeks.

The final step in a design-based research project is to speak back to the theory in which the design is grounded. Constructivist theory suggests that part of the learning process involves transforming information into knowledge. Many of our design moves involved providing information that was previously lacking, reframing information, using new information modalities, and providing opportunities for collective construction of knowledge. This suggests that some forms of learning require information as the raw materials from which to construct knowledge, and that the most effective modalities and framing of information is context-dependent. Constructionist theory argues that learning is more effective when learners construct things, which participants spent the majority of their time doing in this workshop. Because this is a creative process, our design moves indicate that learners need clear and frequent guidance in moving from a production mindset to a generative mode of working. Situated learning theory frames learning as changes in patterns of engagement within a community of practice through mentoring and collaborative activity. Our design moves indicate that scaffolding for collaboration is needed, even in contexts in which we would assume that learners have advanced skills in collaboration.

Implications and conclusion

We found that it is indeed possible to engage learners in the design thinking process without the benefit of a face-to-face context. This was an important finding because there are several benefits to online design thinking workshops. First, online design thinking is cheaper because we used only free tools and didn't have to buy design thinking materials such as sticky notes and dry-erase boards. Second, online design thinking may have better retention. In our experience with face-to-face design thinking workshops spread over multiple weeks, participant attendance always drops dramatically in the final weeks, but in the three iterations of the online design thinking workshop the majority of participants continued to engage in the process through all of the weeks. Finally, online design thinking workshops allow for more participants to engage in each workshop. In face-to-face workshops, we are limited by the number of people who can comfortably work collaboratively in a given space. In the online workshop modality, the number of participants is, theoretically, unlimited. Finally, the Miro boards are persistent — they remain available as long as we want them to be. However, in physical spaces we usually have to remove the sticky notes from the wall because others use the spaces.

We also found some drawbacks to the online design thinking processes. In face-to-face contexts we (subjectively) find that the quality of work is higher. This may be due to a number of reasons, but one of the major reasons appears to be that during the convergent thinking stage, the ability to see hundreds of ideas on the wall at one glance and then rearrange the ideas is only imperfectly approximated in a digital environment. Rather than leverage the affordances of digital technologies to develop a completely new convergent thinking process, we remained fixated on replicating a process ideal to the physical space (working with sticky notes on a wall) in the digital space by using digital sticky notes on a digital wall, which is limited by the screen sizes of the devices that the participants used. The finding that it is possible to engage in the design thinking process using only digital technologies is promising. There are still a number of issues that need to be addressed, and we hope to do so in future iterations of this ongoing design-based research project. We would like to see other researchers find ways to leverage the affordances of digital technologies without attempting to simply replicate the physical affordances of sticky notes on walls.

References

Barany, A. & Donaldson, J. P. (2020). *Digitally-mediated design thinking ideation*. Paper presented at the 2020 American Educational Research Association Annual Meeting, April 17 – 21, 2020, San Francisco, California.

Campbell, C. M., Cabrera, A. F., Michel, J. O., & Patel, S. (2017). From comprehensive to singular: A latent class analysis of college teaching practices. *Research in Higher Education, 58*(6), 581-604. doi:10.1007/s11162-016-9440-0

Coldwell, M., Greany, T., Higgins, S., Brown, C., Maxwell, B., Stiell, B., ... & Burns, H. (2017). *Evidence-informed teaching: an evaluation of progress in England.* Research Report. Department for Education.

Donaldson, J. P., & Barany, A. (2017, June 18-22, 2017). *Digitally-mediated design thinking in CSCL environments*. Paper presented at the Computer Supported Collaborative Learning, Philadelphia, PA.

Donaldson, J. P., & Smith, B. K. (2017). Design thinking, designerly ways of knowing, and engaged learning. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy* (pp. 1-24). Cham: Springer International Publishing.

Donaldson, J. P., Barany, A., & Smith, B. K. (2020). Situated Learning Through Situating Learners as Designers. In M. Bishop, E. Boling, J. Elen, & V. Svihla (Eds.), *Handbook of Research in Educational Communications and Technology: Learning Design* (pp. 819-836). Cham, Switzerland: Springer International Publishing.

Finelli, C., Daly, S., & Richardson, K. M. (2014). Bridging the research-to-practice gap: Designing an institutional change plan using local evidence. *Journal of Engineering Education, 103*(2).

Hazelkorn, E., Ryan, C., Beernaert, Y., Constantinou, C. P., Deca, L., Grangeat, M., . . . Welzel-Breuer, M. (2015). *Science education for responsible citizenship*. Report to the European Commission of the expert group on science education.

Hora, M. T. (2014). Exploring faculty beliefs about student learning and their role in instructional decision-making. *The Review of Higher Education, 38*(1), 37-70. doi:10.1353/rhe.2014.0047

Hora, M. T., & Ferrare, J. J. (2013). Instructional systems of practice: A multidimensional analysis of math and science undergraduate course planning and classroom teaching. *Journal of the Learning Sciences, 22*(2), 212-257. doi:10.1080/10508406.2012.729767

Kafai, Y. B. (2006). Constructionism. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 35-46). New York: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University press.

Oleson, A., & Hora, M. T. (2014). Teaching the way they were taught? Revisiting the sources of teaching knowledge and the role of prior experience in shaping faculty teaching practices. *Higher Education, 68*(1), 29.

Piaget, J. (1952). *The origins of intelligence in children*. New York: W W Norton & Co.

Sandoval, W. A., & Bell, P. (2004). Design-based research methods for studying learning in context: Introduction. *Educational psychologist, 39*(4), 199-201.

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*. Cambridge, MA: Harvard University Press.

Wheeler, P., Trausan-Matu, S., Donaldson, J. P., & Barany, A. (2019). *Experiencing and analysing embodied design thinking in CSCL*. Paper presented at the International Conference on Computer Supported Collaborative Learning, Lyon, France.

Acknowledgments

This work was supported by the Center for Teaching Excellence and the Dean of Faculties at Texas A&M University.

Comparing Example-Based Collaborative Reflection to Problem-Solving Practice for Learning during Team-Based Software Engineering Projects

Sreecharan Sankaranarayanan, Siddharth Reddy Kandimalla, Christopher Bogart, R. Charles Murray, Haokang An, Michael Hilton, Majd Sakr, Carolyn Rosé
sree@cmu.edu, skandima@andrew.cmu.edu, cbogart@andrew.cmu.edu, rcmurray@andrew.cmu.edu, haokanga@andrew.cmu.edu, mhilton@cmu.edu, msakr@andrew.cmu.edu, cprose@andrew.cmu.edu
Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Abstract: Contributing to the literature on aptitude-treatment interactions between worked examples and problem-solving, this paper addresses *differential learning from the two approaches when students are positioned as domain experts learning new concepts*. Our evaluation is situated in a team project that is part of an advanced software engineering course. In this course, students who possess foundational domain knowledge but are learning new concepts engage alternatively in programming followed by worked example-based reflection. They are either allowed to finish programming or are curtailed after a pre-specified time to participate in a longer worked example-based reflection. We find significant pre- to post-test learning gains in both conditions. Then, we not only find significantly more learning when students participated in longer worked example-based reflections but also a significant performance improvement on a problem-solving transfer task. These findings suggest that domain experts learning new concepts benefit more from worked example-based reflections than from problem-solving.

Introduction

The trade-off between problem-solving practice and worked example study has not been deeply investigated in the software engineering context, especially for students who have moved beyond basic syntactic and semantic knowledge about programming and on to advanced topics such as Cloud Computing. In this context, the concepts and skills that students are learning are new, but they have acquired substantial foundational knowledge from their prior learning experiences. In domains where this trade-off between problem-solving practice and worked example study has been thoroughly investigated, extensive problem-solving practice is generally considered inferior for positively impacting student learning (Renkl, 2014). Contrary to what this might suggest for software engineering also, problem-solving practice (i.e., computer programming) has remained the predominant form of pedagogy. This may be because the literature on computer science education does not provide a definitive answer about this trade-off or that the findings are thought to apply especially to conceptual knowledge, and not to performance on more authentic, complex problem-solving tasks. Studies adjacent to the worked example literature relying still on cognitive load theory have variously found positive effects (Margulieux et al., 2012) as well as mixed effects (Morrison et al., 2015) in the software engineering context warranting further study about this important comparison. We seek, therefore, to contribute to the literature on aptitude-treatment interactions between worked examples and problem-solving by addressing the fundamental question of *how students learn differentially from the two when they are positioned as domain experts learning new concepts and skills*. We position our investigation in the software engineering context. We are also especially interested in the follow-up question of whether they might gain more conceptual knowledge from worked example study but be left less able to engage effectively in subsequent problem-solving.

To answer these questions, we conduct our study in a synchronous, online, team-based software engineering course on Cloud Computing for graduate students and advanced undergraduate students. We assign students, in their teams, to two conditions. In the first, which we call the *maximize learning from problem-solving (MLPS)* condition, teams are tasked to complete problem-solving and then engage in a brief collaborative reflection based on a worked example in the remaining time. In the second, called the *maximize learning from reflection (MLR) condition*, teams are curtailed from problem-solving after a pre-specified amount of time regardless of whether they reach a completed solution. They subsequently engage in a full-length reflection based on the worked example. The difference is where the time boundary is placed between problem-solving and collaborative reflection. The results challenge deeply held assumptions in computer science education about the extensive computer programming practice being an activity necessary for student learning.

Method

Course Context

This study was conducted in a graduate-level project-based online software engineering course on Cloud Computing offered to graduate and advanced undergraduate students at Carnegie Mellon University and its branch campuses. The course is structured around five project-based units. Each unit has several sub-units and culminates in a large individual project that has assessment components to evaluate achievement in each sub-unit. Our experiment is situated within unit 3.3 that focuses on *"multi-threaded programming and consistency"*. In this sub-unit, students, in groups of 4, work with our synchronous collaborative programming activity, called the Online Programming Exercise (OPE). A summary of the course structure and the location of the study within it is shown in Figure 1. A total of 74 students completed the exercise and the subsequent project. Enrollment numbers were about half the usual as a result of the COVID-19 pandemic. No other substantial changes in course content or structure were needed since the course had been offered online for over 10 prior semesters.

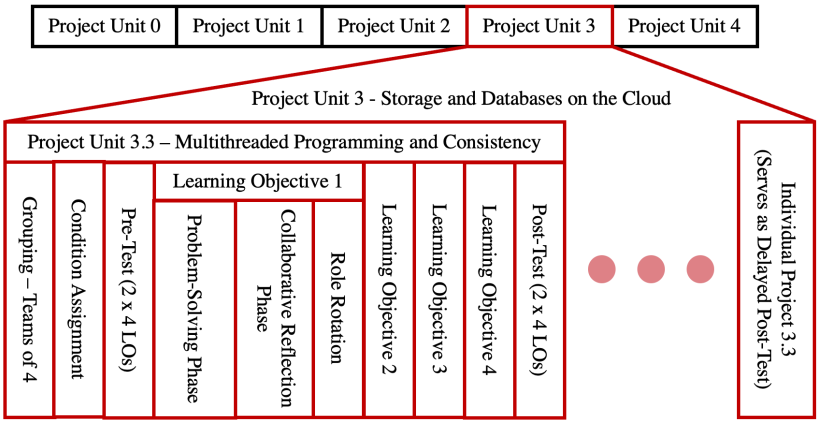


Figure 1: Course structure, pre-test, post-test, and delayed post-test alignment.

Design of the Online Programming Exercise (OPE)

The collaborative programming exercise is divided into four tasks, each targeting a learning objective (LO). Each task is divided into a *problem-solving phase* and a *collaborative reflection phase.* During the problem-solving phase, students are assigned to four independent roles (Driver, Navigator, Researcher, Project Manager) based on an instructional adaptation of the industry practice of Mob Programming (Sankaranarayanan, 2019; Sankaranarayanan et al., 2019). In the subsequent collaborative reflection phase, they are guided by conversational agent-based prompts to reflect based on a presented worked example. The prompting infrastructure is based on the open-source Bazaar conversational agent framework (Adamson et al., 2014). The roles that students are assigned to rotates after each task. In the pre- and post-tests that students complete immediately before and after the task, respectively, they attempt two multiple-choice questions per LO. Performance improvement from pre- to post-test averaged per LO is used as a measure of students' conceptual learning from the task. The individual programming project that serves as a procedural and conceptual delayed post-test is graded by the instructor on a rubric with 12 quality scores, each of which ranges between 0 and 1. Table 1 shows the learning objectives, examples of pre- and post-test questions, and conversational agent-based collaborative reflection prompts corresponding to each task, while Figure 1 shows the position of pre-, post-, and delayed post-tests within the course.

Table 1: Learning Objectives, Corresponding Pre/Post Test Questions (Examples), Information and Transactivity Prompts

	Learning Objective	Example Pre/Post Test Question - Multiple-Choice	Example Collaborative Reflection Prompts
1	Building blocks of multithreading.	Which of the following statements about multithreading in Java is INCORRECT?	Was your approach similar to the reference solution? What Thread class functions did you use? Take turns explaining the logic.
2	Diagnosing and fixing deadlocks.	The usage of notify() will never result in a deadlock in which of the following multithreaded scenarios?	In an ideal scenario, can you think of a built-in Java thread-safe class that could replace the priority queue? Take turns explaining.

3	Diagnosing and preventing a race condition.	Examining the following code snippets, identify the one that will NEVER lead to a race condition.	Comparing your approach to the reference solution, how did you avoid the race condition here? Take turns explaining the logic.
4	Ensuring strong consistency in data stores.	How would you acquire a lock on a critical resource shared by multiple threads to ensure consistent runtime behavior?	Can you put what you are learning in all these tasks together to think about ensuring strong consistency? Take turns explaining.

Experimental Design

Two weeks before the experimental manipulation, students participated in a training OPE session in randomly formed teams of 4 based on their time availability. In preparation, students were provided with videos and text explaining the OPE and motivating its use for collaborative team projects. The exercise was relatively simple data processing using the 'pandas' library in Python. While still a meaningful component of the course, it was meant as an opportunity for students to familiarize themselves with role-taking, role-rotation, collaborative reflection, and the interface of the Cloud9 IDE used for the task. Each exercise session lasted for a total of 80 minutes.

For the experimental manipulation, students were again grouped randomly into teams of 4 based on their time availability while ensuring that they weren't placed into teams with students they had done the training with. The activity, once again, lasted a total of 80 minutes. A total of 74 students were assigned to 19 teams of which 17 were 4-member teams and 2 were 3-member teams. In the 3 member teams, the student assigned to the project manager role also acted as the researcher. 9 teams were assigned to the *maximize learning from problem-solving (MLPS) condition*, where for each task, teams complete the problem-solving and then enter into a reflection phase for the remaining time, and 10 teams were assigned to the *maximize learning from reflection (MLR)* condition, where problem-solving was curtailed after a pre-specified amount of time, and they enter the reflection regardless of whether they completed the problem-solving or not.

Hypotheses, analysis, and results

Hypothesis 1: The Online Programming Exercises (OPEs) results in pre- to post-test learning gains

To evaluate the general value for learning of the activity regardless of condition, we compared pre- and post-test scores per learning objective, role, and condition, where pre- and post-test scores vary between 0 and 1 per learning objective. For this analysis, we build an ANOVA model with test score as the dependent variable, and time-point (pre- vs post-test), condition (MLPS vs MLR), role (Driver, Navigator, Researcher, Project Manager), and learning objective (listed in Table 1) as independent variables. We also included pairwise interaction terms between time-point and each of the other three independent variables. There was a significant effect of time-point $F(1,410) = 3.77$, $p < .0001$, effect size .38 s.d., with an average pre-test score of .55 (.37 s.d.) and average post-test score of .69 (.37 s.d.). None of the pairwise interactions were significant. Thus, we confirmed that students in both conditions learned based on the significant difference between pre- and post-test scores across the two conditions regardless of role or learning objective. Thus, the first hypothesis is confirmed.

Hypothesis 2: *The MLR condition will result in better pre- to post-test learning gains.*

In order to test the effect of condition on the magnitude of learning we compared post-test scores between conditions controlling for pre-test scores. In particular, we computed an ANCOVA model with post-test score as dependent variable, pre-test score as the covariate, and condition (MLPS vs MLR), and learning objective as independent variables. We found a significant effect of condition such that students in the MLR condition learned more $F(1,254) = 6.0$, $p < .05$, effect size .24 s.d.. For the MLR condition, the average pre-test score was .53 (.36 s.d.) and post-test score was .72 (.34 s.d.), and for the MLPS condition, average pre-test score was .57 (.38 s.d.) and post-test score was .65 (.37 s.d.) We find significantly higher pre- to post-test gains in the MLR condition in comparison with the MLPS condition ($p < 0.05$), which suggests that students with domain expertise benefit more from the worked example-based reflection than the problem-solving for acquisition of new conceptual knowledge. Thus, the second hypothesis is confirmed.

Hypothesis 3: *The MLPS condition will result in better performance on the delayed post-test*

In order to test the effect of condition on achievement on the transfer task, we considered each of the 12 quality ratings assigned by the instructor within a single model in order to control for multiple comparisons. In particular, we computed a single ANCOVA model with numeric quality rating as the dependent variable, total pre-test score across learning objectives as a covariate, and the condition and the name of the quality rating as independent variables. We also include the pairwise interaction term between the two independent variables. There was a significant effect of condition such that students in the MLR condition scored higher than students in the MLPS condition, $F(1,707) = 4.36$, $p < .05$, effect size .15 s.d., which is a weak effect. The average score was 4.7 (2.7 s.d.) for the MLPS condition and 5.0 (2.5 s.d.) for the MLR condition. There was no significant interaction between condition and quality rating name. Thus, although the effect is weak, students in the MLR condition performed better than students in the MLPS condition across the 12 quality ratings. The third hypothesis is rejected, and in fact, the opposite is supported.

Discussion and conclusion

In this paper we presented a study in which we contrasted two conditions: *maximize learning from problem-solving* and *maximize learning from reflection*.

First, we find that the team project exercises lead to significant pre- to post-test learning in both conditions. This indicates that both worked example study, and problem-solving practice are potentially valuable for learning. While we did not compare the sequential presentation of the problem-solving, worked examples and the collaboration scaffolds with either of the scaffolds provided on their own, the lack of a detrimental effect in either condition means that the role of the scaffolds was not so redundant as to draw student attention away from the relevant problem states.

When comparing across conditions, we see that the condition where students spent more time on worked example-based reflection resulted in significantly more pre- to post-test learning. Based on cognitive load theory, we could surmise that it is indeed the case that extensive problem-solving consists of production steps that are superfluous to the learning here. While problem-solving practice was not detrimental to student learning, we can more efficiently use student time and impact their learning more if we use worked examples as well, with an emphasis on time spent on reflecting rather than the completion of the problem-solving.

One concern among educators has been that while students' conceptual learning can be positively impacted by the use of worked examples, they may not perform as well when asked to problem-solve on a transfer task because they received less practice. We started with the hypothesis that this might be the case, and we would have not been surprised to have found that. However, what we found was the opposite. Students who reflected longer also performed better on a subsequent authentic programming task, though the effect size was small. We can conclude that students, at this point in the course, had already acquired the procedural knowledge of programming enough to not need the practice i.e., given a schema, they were able to translate that into a solution to the problem. The positive impact of the worked example condition on the conceptual process of schema acquisition and induction then led to a positive impact on student performance on the subsequent project also.

References

Adamson, D., Dyke, G., Jang, H., & Rosé, C. P. (2014). Towards an agile approach to adapting dynamic collaboration support to student needs. *International Journal of Artificial Intelligence in Education*, *24*(1), 92-124.

Margulieux, L. E., Guzdial, M., & Catrambone, R. (2012, September). Subgoal-labeled instructional material improves performance and transfer in learning to develop mobile applications. In *Proceedings of the ninth annual international conference on International computing education research* (pp. 71-78).

Morrison, B. B., Margulieux, L. E., & Guzdial, M. (2015, August). Subgoals, context, and worked examples in learning computing problem-solving. In *Proceedings of the eleventh annual international conference on international computing education research* (pp. 21-29).

Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive science*, *38*(1), 1-37.

Sankaranarayanan, S. (2019, February). Online Mob Programming: Effective Collaborative Project-Based Learning. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (pp. 1296-1296).

Sankaranarayanan, S., Wang, X., Dashti, C., An, M., Ngoh, C., Hilton, M., Sakr, M., & Rosé, C. (2019, June). An Intelligent-Agent Facilitated Scaffold for Fostering Reflection in a Team-Based Project Course. In *International Conference on Artificial Intelligence in Education* (pp. 252-256). Springer, Cham.

Acknowledgements

This work was funded in part by NSF grants IIS 1822831, IIS 1917955 and funding from Microsoft.

International Society of the Learning Sciences

Designing Online Teaching for Equitable Distribution of Student Engagement in Collaborative Small Groups: The Effects of Group Building and Reciprocal Feedback

Manfred Kaiser
Technical Writing and Technical Communication, Munich University of Applied Sciences, Germany
Faculty of Psychology and Educational Sciences, Ludwig-Maximilians-Universität München, Germany
manfred.kaiser@campus.lmu.de
Thomas Lerche
Faculty of Psychology and Educational Sciences, Ludwig-Maximilians-Universität München, Germany
thomas.lerche@edu.lmu.de
Nicolae Nistor
Faculty of Psychology and Educational Sciences, Ludwig-Maximilians-Universität München, Germany
Richard W. Riley College of Education and Leadership, Walden University, USA
nic.nistor@uni-muenchen.de

Abstract: The unequal distribution of workload in groups is a phenomenon that affects all areas of life. At universities, group work constantly leads to dissatisfaction among participants in collaborative groups and often influences the outcome of their work. In this study, we tested different settings, which can influence the contribution of members in blended learning courses in higher education. We used a questionnaire, asking for the assessment of one's own and that of the group members contribution, conducted at two points of the course. In addition, the members of certain groups were asked to give reciprocal feedback in the middle of the course. Additionally we used log files for the evaluation of the activities in the Learning Management System. Groups with reciprocal feedback initially rated the participation of members higher and reduced their rating in the end. Without reciprocal feedback, participants rated the contribution of their group members higher at the end than during the course. The perceived own contribution increased over time across all groups. The actual activities, measured through log data from the Learning Management System showed that self built groups with reciprocal feedback exhibit lower participation in the online environment. The grounding effect is evident across all groups; the activities in the Learning Management System have strongly decreased during the courses.

Introduction

Online courses are becoming increasingly important in the academic environment, especially in higher education. The current pandemic has given a further boost to the development and raised the need for collaboration in online environments. The phenomenon of social loafing, meaning that individuals reduce their effort being part of a group is a well-evaluated effect (Latané, Williams, & Harkins, 1979). This phenomenon is also evident in online communities. Various studies have identified factors that can encourage or reduce social loafing. For example group evaluation and rewards and increased task difficulties can reduce social loafing (Mefoh & Nwanosike, 2012). Highly performance-oriented characters tend less to social loafing in group settings (Hart, Karau, Stasson, & Kerr, 2004). Individuals with a high degree of autonomy, on the other hand, are more disposed to social laziness (Huguet, Charbonnier, & Monteil, 1999, p. 124). Each individual can influence the motivation and contribution of the other group members through his or her own characteristics and contribution to the common task. Early (1989) examined the differences in social loafing tendencies between Chinese and Americans and found that Chinese perform better in group conditions than as individuals, when they remain among themselves and groups are not mixed.

Three approaches for assigning students to teams have been investigated: self-selection, random assignment, and teacher assignment (Decker, 1995). Each method has advantages and disadvantages. In self-selection, students could choose each other because they have had positive experiences with their fellow students before, or they choose students who are consciously experienced in order to get through the seminar with little effort. Another disadvantage is that teams can possibly be too homogeneous and suffer from a lack of diversity. There are different opinions about randomly composed groups. Some randomly selected teams may happen to receive a desirable combination of students, but others may not (Bacon et al., 1998, p. 69). However, this is also the case with self-assembled groups. The third system for assigning team members is the teacher assignment method. Criteria for this method vary widely and can be difficult to implement. In our study we use the two

 © ISLS

methods self-selection and team assignment by the instructor. Since no special criteria were applied in the second case, this procedure is equivalent to a random division.

It is generally accepted that reciprocal feedback is a good measure to reduce social loafing. The knowledge about the observation from outside increases the pressure to provide better performance. But feedback can also have the opposite effect. Negative feedback can affect performance and motivation and even lead to a lower rating of other group members (DeNisi, Randolph, & Blencoe, 1983). It matters when feedback is given during a task. Feedback can also encourage undesirable behavior if it happens at the end of a course. Without mutual confrontation with these habits, students tolerate the behavior and think that their poor performance at the end of the semester has no relevance in the peer evaluations. (Bacon, Stewart, & Silver, 1999).

These results in mind, the composition of groups can have a great influence on the cooperation and the outcome of a task. In this study we want to investigate to what extent it makes a difference if groups are formed by the teacher or if the participants are allowed to form groups themselves. As a second factor, we want to investigate to what extent reciprocal feedback during group work influences the participation and the perception of the contribution of the group members. We used two questionnaires that have been established in earlier studies. Høigaard (2006; 2010) developed the Self-Reported Social Loafing Questionnaire (SRSLQ) for the assessment of the own level of contribution and effort in a handball team. Brooks (2003) used a questionnaire to evaluate the team members contribution of a group in a business course. We applied both questionnaires in a German translation.

Learning Management Systems (LMS) offer comfortable possibilities for online collaboration. The access of each user is documented with the help of log files (Lerche & Kiel, 2018). We used these logfiles for our study to find out if there is a correlation between real activity in online collaboration and settings in the design of courses, in our case group formation and reciprocal feedback. We assume that grounding will take place, as it is a typical phenomenon in communication and is evident in all collective activities (Clark & Brennan, 1991). A new situation or a new environment requires an increased need for communication. Over time, this demand decreases, as basic information is available, to which only new input needs to be added. (Kraus & Fussel, 1991). The need for knowledge is intensive at the beginning, decreasing along the learning process (Kraus & Fussel, 1991). If students have already worked together in previous academic activities, they may need less grounding than if they meet for the first time in a collaborative group.

Research questions

Within the theoretical framework outlined above, in this study we investigate two research questions (RQs):

What are the effects of group building (by instructors vs. by students) and reciprocal feedback (absent vs. present) on students' changes of the perceived engagement (RQ1) and of the actual engagement (RQ2) in online groups during a collaborative learning process?

As suggested in the introduction, we expect both the perceived and the actual engagement to decrease in time along with grounding, and to be stronger in self-built groups and with reciprocal feedback.

Methodology

The present quantitative study has a 2x2 factorial design. The participants were graduate students of educational sciences, N = 114 (N = 110 female , N = 4 male) at a large German university. N = 62 (N = 59 female , N = 3 male) agreed to participate in the study and completed all surveys. The students attended the graduate course "Training Methods". This course covered the development of a training concept using realistic cases from various areas of life. Each case had a different problem as a starting point. The students had to use current scientific methods and concepts for the analysis and development of the training. They acted as potential contractors in the form of a fictitious start-up. The tasks were solved within groups of 3 to 5 students. Each group received a different case or a different training proposal. The course covered the areas of need/task analysis, training objectives, instructional concept, media concept and evaluation concept. The students were encouraged to realize the mutual exchange via the forum of the Learning Management System (LMS), in this case Moodle, and had to deliver the tasks at certain deadlines. The course was offered four times, but the setting was changed with two independent variables: composition of groups and reciprocal feedback. There were following settings: Group building by students with reciprocal Feedback (GSt_FB), Group building by students, no reciprocal Feedback (GSt_NoFB), Group building by instructor with reciprocal Feedback (GIN_FB), Group building by instructor, no reciprocal Feedback (GIN_NoFB).

In one of each of these two courses, the students had to assess each other individually within the LMS parallel to the first survey (reciprocal feedback). All groups had to complete a questionnaire in the middle and at the end of the course in which they assessed their own participation on the task with 7 items (SRSLQ) (Høigaard & Ingvaldsen, 2006), α = 0.75 and that of the group members with 6 items (PSLQ) (Brooks & Ammons, 2003), α = 0.93, which represents the dependent variable. The possible answers were based on a 7-point Likert scale. The scale ranged from 1 (Not applicable at all) to 7 (Fully applicable). To get comparable results we reverse scored items in the Self Evaluation, in order to receive the value of the contribution and not the value of the Social Loafing for both questionnaires (see Table 1). In addition, the log files of the LMS were used to determine the actual participation of each student within the online-forum of the LMS. We measured the contribution to the result of the work by using the datasets ForumThemeView and ForumPostCreate from the log files. They were split into data until the first evaluation and data from the first to the second evaluation.

Table 1: Questionnaires for self-evaluation and group member evaluation

Self Evaluation	Team Member Evaluation
In a team,…	Group members…
1. I am not indispensable	1. prompt in attendance at team meetings.
2. I will try as hard as I can	2. delivered agreed-upon parts of project in a complete fashion.
3. I will contribute less than I should[a]	3. met deadlines.
4. I will actively participate in the discussion and contribute ideas	4. volunteered appropriately during team meetings when tasks needed to be accomplished.
5. it is okay even if I do not do my share[a]	5. pulled fair share with regard to overall workload.
6. it does not matter whether or not I try my best[a]	6. showed enthusiastic and positive attitude about team activities and fellow team members.
7. given my abilities, I will do the best I can	
[a] Items reverse scored.	

Findings

All dependent variables were tested for normal distribution. A transformation with a square root was necessary to receive satisfying results ($p > .05$). Skewness and kurtosis were within ±3. The Z-standardized value was smaller than ±1.96 for all variables ($p < .05$ for $n < 200$). The Levene Test showed homoscedasticity for all questionnaires. We found few large residuals, that we decided to keep and replace the value through the mean.

RQ1. To measure the influence of the independent variables group building and reciprocal feedback on the evaluation of one's own contribution and that of the group members, we conducted two seperate repeated measure ANOVA.

For the own assessment there was a significant main effect, $F(1, 58) = 8.81$, $p = .004$, partial $\eta^2 = .13$. A significant interaction could not be observed with group building, reciprocal feedback, nor with group building * reciprocal feedback. The ANOVA delivered no significant between-subject effects. For the assessment of the group members, no significant main effect was found. There was significant interaction with feedback, $F(1, 58) = 8.15$, $p = .006$, partial $\eta^2 = .12$, but no effect with group building, or group building * reciprocal feedback. Feedback had a significant between-subjects effect, $F(1, 58) = 9.16$, $p = .004$, partial $\eta^2 = .13$.

RQ 2. There was no homoscedasticity in the log data, so a non-parametric test was used. The Kruskal-Wallis H test showed that there was a statistically significant difference in the course settings until the first survey, $H(3) = 20.58$, $p = 0.001$, with a mean rank activity of 42.81 for GIN_noFB, 34.12 for GST_noFB, 33.88 for GIN_FB and 11.50 for GST_FB. The pairwise comparison showed a significant effect for GSt_FB - GIN_FB ($p = 0.027$), GSt_FB - GSt_NoFB (p=.001), GSt_FB - GIN_NoFB (p=.001). No significant differences ($H(3) = 7.69$, $p = .053$) were found among the four group settings (GSt_FB, GSt_NoFB , GIN_FB, GIN_NoFB) in the log data from the first to the second survey.

Conclusions

This study had the aim to show that the way of group building and mutual feedback can influence the perceived collaboration of students in group settings. The study revealed that, regardless of these factors, students rated their own performance during the course relatively similarly and gave themselves an even slightly higher rating at the end of the course. The evaluation of the group members, however, differed depending on the reciprocal feedback. If there was no reciprocal feedback in the course, the students assessed the performance of their group members similarly to their own performance at both measurement points. The score for the performance increased at the end of the course. However, if the students had to give reciprocal feedback at the first survey, they even attested the group members a higher contribution than themselves. At the end of the course they corrected this assessment and adjusted it to their own performance. Overall, the evaluation of the group members is rated higher with feedback than without feedback. The phenomenon could be explained by the fact that, due to the awareness that one is being evaluated oneself and the fear of receiving a bad evaluation, the group members' performance is considered very high as a precaution. At the end of the course, the evaluation no longer matters and has no influence on the result of their work. They can therefore return to a more realistic assessment.

Participation in the LMS fluctuated widely between the course settings. In the GSt_FB course, online activity was significantly lower than in the other constellations. This explains the statistical significance of the pairwise comparison at the first measurement interval. It is quite possible that students used other communication channels in addition to the LMS, or met privately for collaboration. Between the first and the second part of the courses the log data evaluation of the LMS showed a strong difference in the amount of data created. The grounding effect was evident across all settings. Activity was approximately halved in all courses.

The results of our study provide important recommendations for educational practice. The impact on performance in collaborative communities is highest when the instructor assigns the students to groups, and when there is reciprocal feedback during the course.

There are some limitations in this study. The relatively small sample size does not allow a generalization of the result. The study does not examine real Social loafing, as it only considers the perceived contribution to group work and not the actual outcome. The log data evaluation does not provide an exploitable result. Furthermore, social loafing tendencies can diminish if mean values of all participants, the hard working and the lazy ones, are considered. In a further investigation, it should be considered to filter the values that clearly indicate social loafing and to examine them separately regarding group settings and reciprocal feedback.

References

Bacon, D. R., Stewart, K. A., & Stewart-Belle, S. (1998). Exploring predictors of student team project performance. Journal of Marketing Education, 20(1), 63–71. https://doi.org/10.1177/027347539802000108

Bacon, D. R., & Stewart, K. A. (2019). “Lessons from the best and worst team experiences: How a teacher can make the difference”: Reflections and recommendations for student teams researchers. Journal of Management Education, 43(5), 543–549. https://doi.org/10.1177/1052562919849670

Brooks, C. M., & Ammons, J. L. (2003). Free riding in group projects and the effects of timing, frequency, and specificity of criteria in peer assessments. Journal of Education for Business, 78(5), 268–272. https://doi.org/10.1080/08832320309598613

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.) Perspectives on socially shared cognition (pp. 127-150). Washington, DC: American Psychological Association.

DeNisi, A. S., Randolph, W. A., & Blencoe, A. G. (1983). Potential problems with peer ratings. Academy of Management Journal, 26(3), 457–464. https://doi.org/10.5465/256256

Earley, P. C. (1989). Social loafing and collectivism: A comparison of the United States and the People’s Republic of China. Administrative Science Quarterly, 34(4), 565. https://doi.org/10.2307/2393567

Efraty, D., & Stratton, W. E. (1995). Developments in business simulation & experiential exercises, Volume 22 , 1995 Developments In Business Simulation & Experiential Exercises , Volume 22 , 1995. 22, 268–269.

Hart, J. W., Karau, S. J., Stasson, M. F., & Kerr, N. A. (2004). Achievement motivation, expected coworker performance, and collective task motivation: Working hard or hardly working? Journal of Applied Social Psychology, 34(5), 984–1000. https://doi.org/10.1111/j.1559-1816.2004.tb02580.x

Hoigaard, R., Säfvenbom, R., & Tonnessen, F. E. (2006). Group cohesion, group norms, and perceived social loafing in soccer teams. Small Group Research, 37(3), 217–232.

Høigaard, R., Fuglestad, S., Peters, D. M., Cuyper, B. De, Backer, M. De, & Boen, F. (2010). Role satisfaction mediates the relation between role ambiguity and social loafing among elite women handball players. Journal of Applied Sport Psychology, 22(4), 408–419. https://doi.org/10.1080/10413200.2010.495326

Huguet, P., Charbonnier, E., & Monteil, J. (1999). Productivity loss in performance groups: People who see themselves as average do not engage in social loafing, 3(2), 118–131.

Krauss, R. M., & Fussel, S. R. (1991). Constructing shared communicative environments. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), Perspectives on socially shared cognition (pp. 172-200). Washington, DC: American Psychological Association.

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. Journal of Personality and Social Psychology, 297–308. https://doi.org/10.1037/0022-3514.37.6.822

Lerche, T., & Kiel, E. (2018). Predicting student achievement in learning management systems by log data analysis. Computers in Human Behavior, 89, 367-372. https://doi.org/10.1016/j.chb.2018.06.015

Mefoh, P. C., & Nwanosike, C. L. (2012). Effects of group size and expectancy of reward on social loafing. IFE Psychologia: An International Journal, 20(1), 229–239. https://journals.co.za/doi/10.10520/EJC38908

Analyzing debugging processes during collaborative, computational modeling in science

Nicole M. Hutchins, Vanderbilt University, nicole.m.hutchins@vanderbilt.edu
Caitlin Snyder, Vanderbilt University, caitlin.r.snyder@vanderbilt.edu
Mona Emara, Damanhour University, mona.emara@damanhour.edu.eg
Shuchi Grover, Looking Glass Ventures, shuchig@cs.stanford.edu
Gautam Biswas, Vanderbilt University, gautam.biswas@vanderbilt.edu

Abstract: This paper develops a systematic approach to identifying and analyzing high school students' debugging strategies when they work together to construct computational models of scientific processes in a block-based programming environment. We combine Markov models derived from students' activity logs with epistemic network analysis of their collaborative discourse to interpret and analyze their model building and debugging processes. We present a contrasting case study that illustrates the differences in debugging strategies between two groups of students and its impact on their model-building effectiveness.

Introduction

Research has demonstrated debugging's pivotal role in learning computational thinking (CT; Rich et al., 2019) and learning by computational modeling in science domains (Sengupta et al., 2018), thus warranting further research to enhance our understanding of debugging processes that support this learning. Analyzing students' performance purely on the basis of the artifacts (model, program) they produce may not provide sufficient information of the processes they apply for learning computational concepts and practices, and, therefore, the difficulties they face in their model-building tasks (Grover et al., 2017). In addition, analyzing activity data from log files may not provide sufficient evidence on the processes students employ to integrate domain-specific knowledge with CT concepts and practices to build and debug models. To overcome this problem, we develop a multimodal learning analytics (MMLA) approach combining quantitative activity data and collaborative discourse analysis to study students' debugging processes during collaborative, computational modeling activities in a high school physics class. We focus on the following research questions: (RQ1) *How can we identify and characterize students' model-building and problem-solving processes from the log data collected as they worked on modeling science phenomena?* and (RQ2) *What additional information can we derive from students' collaborative discourse on the processes they employ in integrating science and CT concepts and practices as they build and debug their computational science models in a block-based programming environment (BBPE)?*

MMLA framework to evaluate collaborative debugging

A review of recent national standards in science, technology, engineering, and math (STEM) and computer science demonstrates that common problem-solving practices targeted across the two domains include the creation of artifacts and using CT (NGSS, 2013; Collegeboard, 2017). Students' application of these practices offers a unique context to examine their learning processes as they build computational science models. We hypothesize that analyzing instances of shared practices will provide an informative context for examining mutually beneficial instances of students' synergistic STEM and CT learning (Snyder et al., 2019). We focus on two factors that are indicative of synergistic learning: (1) shared practices between STEM and CT and (2) combining concepts through domain-specific block structures provided in our environment (e.g., contextualizing a conditional block when considering the stopping condition for a truck to stop at a stop sign) (Hutchins et al., 2020).

Our analysis centers on the examination of students' model-building and debugging strategies in C2STEM, a web-based, BBPE (Hutchins et al., 2020). In this study, 14 high school students worked in groups of 2 or 3 to build multiple kinematic models of 1-D and 2-D motion of objects. The curriculum included instructional tasks, designed to help students develop specific physics and CT concepts, followed by model-building tasks where students modeled the motion of a truck in three phases: (1) a speed up phase; (2) a cruise phase with constant velocity motion; and (3) a slow-down phase where the truck comes to a stop at a pre-designated STOP sign. Students were encouraged to work together to understand the problem, and build a kinematics model that produced the right behaviors.

To answer RQ1, we applied Markov Chain analysis (Craig & Sendi, 2002) to logged activity data for each student group and evaluated the occurrence of high probability sequences of actions and represented their temporal occurrence using CORDTRA graphs. Student actions were recorded in log files with timestamps, and analyzed sequentially along with the context in which they were being applied to derive action patterns that we then mapped into model-building and debugging tasks. To increase interpretability, we developed a hierarchical task-oriented structure (Emara et al., 2021) to map student actions to a level of abstraction so that the derived action patterns could be interpreted as code construction and code assessment activities (cf. Rich et al., 2019). Code assessment actions were labeled as either data analysis (DATA; this involved opening the graph tool) or visual feedback (PLAY; i.e., running the simulation). Code construction actions were labeled as either building the model (BUILD; i.e., constructing new elements of the program code) or adjusting the existing model (ADJUST; i.e., debugging code after identifying an issue).

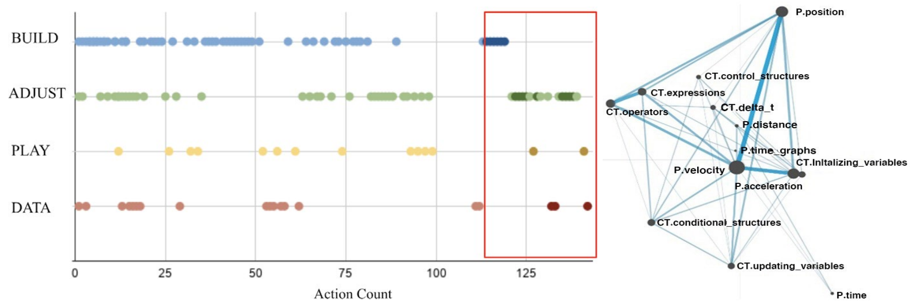
To answer RQ2 and track students' ability to understand and combine CT and STEM concepts and practices, we developed a MMLA approach that combined our log-file based behavioral pattern analysis with Epistemic Network Analysis (ENA; Shaffer et al., 2009) of students' discourse. As a result, our MMLA analysis helped us better interpret temporal patterns derived from coded discourse data on STEM and CT topics. Our coding scheme included physics concepts: P.position, P.velocity, P.acceleration, P.time, P.distance, P,displacement and P.time_graphs. Similarly, the coding scheme included the following CT concepts: CT.delta_t, CT.block_ordering, CT.control_structure, CT.initalizing_variables, CT.updating_variables, CT.operators_expressions, CT.conditional_structures, CT.data_collection and CT.data_visualization. Two researchers transcribed the audio for the two case study groups selected and coded 40 utterances (approximately 15% of all segments) using this coding scheme, resulting in good IRR agreement (k = 0.76). One researcher coded the remaining segments. We present a mixed-method case study approach to illustrate two contrasting cases, where students adopt different approaches resulting in varying degrees of success in their model building tasks. Groups were scored utilizing our integrated STEM+CT computational modeling rubric (Emara et al., 2021) and the two groups selected represent high performing and low performing groups based on a median split of the overall, final model scores. We hypothesize that students who are successful in integrating STEM and CT concepts also show greater learning in both domains.

Results and discussion

Results of the Markov Chain (MC) analysis revealed students used key modeling and debugging strategies, including: (1) **Depth-First** (DF) for multiple, simultaneous code construction actions (e.g., a sequence of BUILDs) without intervening code assessment actions (e.g., PLAY actions), (2) **Tinkering** (TIN), or trying different parameters or block changes in a trial and error fashion to gain some understanding of how to get to a solution (e.g., sequences of "ADJUST→PLAY" (0.20) and "PLAY→ADJUST (0.39) actions), (3) **Multi-Visual Feedback** (MV), represented by sequence of PLAY activities with high probability (0.19), and (4) **Simulation-based Assessment** (EVAL), or evaluating the model using the data tools, e.g., PLAY→DATA (0.29) and DATA→PLAY (0.19). We believe this may represent a build and test strategy. A more refined version of build and test, the modeling-in-parts strategy, has also been shown to support successful model creation (c.f., Grover et al., 2017). This strategy involves working on the model in small segments (i.e., BUILD↔ADJUST) combined with a sequence of execution actions (PLAY↔DATA) to understand the behaviors generated by that segment of the model, and debugging the model, as needed. We will examine identified processes in our case study.

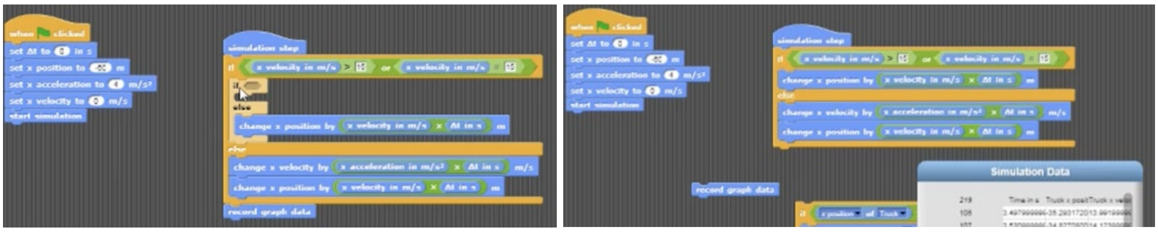
We apply our MMLA framework to examine the reasoning processes of two groups of students that had difficulties in constructing the conditional expression for the third phase of the truck's motion. Group 1 is characterized by their DF approach to model construction and debugging. Figure 1(a) shows that before action #50, most of the model building happened in a DF manner. The DF approach resulted in complex conditionals (Figure 2(a)), which increased the complexity of their debugging tasks. Rather than using feedback from the tools provided to analyze their model behaviors (i.e., PLAY or DATA actions), they used their physics knowledge to analyze the model (via pen-and-paper calculations), and a DF approach (e.g., the darkened blue and green circles in the red box of Figure 1(a)) in an attempt to make their code correspond to the correct physics solution they had derived on paper. These results match the literature on debugging that has shown novices tend to use depth-first construction and debugging strategies because they lack the insight for breaking down a complex task into its sub-parts (Grover et al., 2017). The group's ENA graph (Figure 1(b)) indicates a separation of physics-focused and CT-focused discourse with the top three connections between concepts in the ENA graph include "position-velocity" (edge weight = 0.78), "velocity-acceleration" (edge weight = 0.70), and "operators-expressions" (edge weight = 0.62). The inability to think of the physics and CT concepts as one integrated unit, may indicate difficulties in translating their pen-and-paper physics calculations into computational form (Basu et al., 2017).

The group was unable to build a correct model. Instead of working on the conditional expressions, the group used the data tools and analyzed parts of the model (Figure 2(b)) to calculate the time required from start to cruise (the speed up phase), and then used that information to calculate on paper the distance at which to slow down. This further confirmed their strength in domain knowledge, but weakness in applying CT-related concepts.



(a) (b)

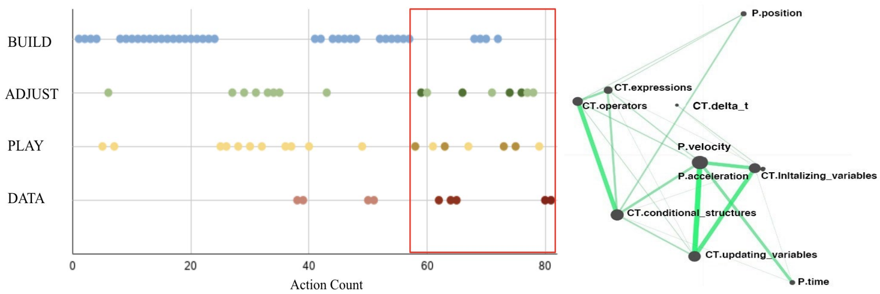
Figure 1. Group 1 CORDTRA graph (a) and ENA graph (b).



(a) (b)

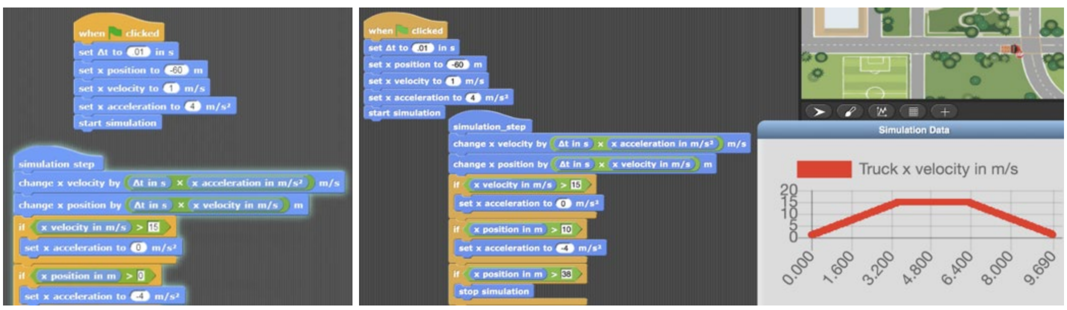
Figure 2. Group 1 code snippet prior to debugging (a) and after debugging (b).

Group 2 used a modeling-in-parts strategy to construct their model. Their debugging processes started around action 60 (indicated by the red box in Figure 3(a)), and it included multiple PLAY and DATA actions, indicating the use of TIN and EVAL strategies identified previously. Simultaneously, their ENA graph (Figure 3(b)) indicates increased integration of STEM and CT discourse elements compared to Group 1. The top three discourse connections were “updating variables - velocity” (edge weight = 0.87), “updating variables - acceleration” (edge weight = 0.73), and “operators - conditional structures” (edge weight = 0.73).



(a) (b)

Figure 3. Group 2 CORDTRA graph (a) and ENA graph (b).



(a) (b)

Figure 4. Group 2 code snippet prior to debugging (a) and after debugging (b).

Unlike Group 1, Group 2 elected to use the visualization tool to adjust the segment of code they were working on. For instance, during their debugging process a student noted that they were not sure if the conditions they created for switching phases would overlap, but they should run the newly created condition with 0 "for now," just to check. Once they determined how that part of their code worked, they used the data tools to determine the approximate location to begin the slowing down phase, recalling that the slowing down process was a complement of the speeding up process that they had implemented. Once they got that condition working, using the data tools to estimate the x-position in which to begin slowing down, they proceeded to implement the final stopping condition, and again used the TIN and EVAL strategies to complete their model (Figure 4(b)).

Conclusions

Our MMLA analyses provides initial evidence that: (1) the analysis of students' modeling and debugging processes by combining activity logs and ENA for discourse analysis provides a deeper understanding of students' reasoning with STEM and CT concepts and practices, and (2) students use multiple approaches and different strategies to develop and debug solutions, and this provides them the opportunities for developing higher order thinking and problem-solving skills. However, the difficulties students face in working on these complex problems also suggest that it may be helpful to provide scaffolding to students when they cannot progress in their problem-solving tasks (Basu et al., 2017). These findings could be used to provide feedback or support for students to make it easier to decompose a complex task into sub-parts and to understand how to employ visualization and data to support debugging. Overall, this approach supports learning-by-modeling approaches for synergistic STEM and CT learning. Finally, these findings can be extended to CS and introductory programming classrooms as well, to guide teaching and learning of programming and debugging.

References

Basu, S., Biswas, G. & Kinnebrew, J.S. (2017). Learner modeling for adaptive scaffolding in a Computational Thinking-based science learning environment. *User Model User-Adap Inter*, *27*, 5–53.

CollegeBoard. Advanced placement computer science principles course guide. 2017.

Craig, B.A., & Sendi, P.P. (2002). Estimation of the transition matrix of a discrete-time markov chain. *Health economics*, *11*(1), 33–42.

Emara, M., Hutchins, N.M., Grover, S., Snyder, C., & Biswas, G. (2021, in press). Examining Students' Regulation of Collaborative, Computational, Problem-Solving Processes in Open-Ended Learning Environments. *Journal of Learning Analytics*. SOLAR.

Grover, S., Basu, S., Bienkowski, M., Eagle, M., Diana, N., & Stamper, J. (2017) A framework for using hypothesis-driven approaches to support data-driven learning analytics in measuring computational thinking in block-based programming environments. *ACM Trans. Comput. Educ*., *17*(3).

Hutchins, N.M., Biswas, G., Maróti, M., Lédeczi, Á., Grover, S., Wolf, R., … & McElhaney, K. (2020a). C2STEM: a System for Synergistic Learning of Physics and Computational Thinking. *Journal of Science Education and Technology*, 29(1), 83-100.

NGSS Lead States. (2013). Next Generation Science Standards: For states, by states. Washington, DC: National Academies Press.

Rich, K. M., Strickland, C., Binkowski, T. A., & Franklin, D. A K-8 debugging learning trajectory derived from research literature. *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 745–751, New York, NY, USA. Association for Computing Machinery.

Sengupta, P., Dickes, A., & Farris, A. (2018). Toward a phenomenology of computational thinking in stem education. In *Computational thinking in the STEM disciplines*, 49–72. Springer.

Shaffer, D. W., Hatfield, D., Svarovsky, G., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A., & Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st century assessment of learning. *International Journal of Learning and Media*, *1*(2), 33–53

Snyder, C., Hutchins, N., Biswas, G., Emara, M., Grover, S., Conlin, L. (2019). Analyzing Students' Synergistic Learning Processes in Physics and CT by Collaborative Discourse Analysis. *Proceedings of 13th International Conference on Computer Supported Collaborative Learning 2019* (360-367), Lyon, France. International Society of the Learning Sciences.

Acknowledgments

This research is sponsored by the National Science Foundation Award DRL-1640199. We also appreciate the contributions of our C2STEM research team members and the suggestions of the anonymous reviewers.

Young Students' Experience of Analytics-supported CSCL and the Influence of Parental Attitudes

Xiao Hu, The University of Hong Kong Shenzhen Institute of Research and Innovation, China, xiaoxhu@hku.hk
Jeremy Ng, The University of Hong Kong, Hong Kong S.A.R., jntd@connect.hku.hk,
Xuemei Jiang, The 5th Huaguoyuan Primary School, Guizhou, China, 948639234@qq.com

Abstract: Being recognized as a research priority in the computer-supported collaborative learning (CSCL) community, learning analytics (LA) help harvest and make sense of empirical evidence of students' collaborative learning. Given that prior investigations on LA-supported CSCL largely focused on university classrooms, implementation of LA for facilitating wiki-based CSCL environments in primary education is rarely explored. The focus on school-aged children also brings about the consideration of their parents' attitudes. This study aims to examine primary school students' experience of LA-supported CSCL and how their self-perceived parents' attitudes towards their use of digital technology influence their learning experience. Survey, wiki logs and interview data were collected from 46 students and three teachers involved in LA-supported wiki-based group inquiry projects. Results show that after receiving LA support, students' perceptions of inquiry-based learning and those of wiki-supported learning improved while they were less positive towards collaborative learning. In the meantime, it is found that, from students' perspective, their parents' awareness of the purpose of their mobile usage was negatively correlated with students' participation on wiki. Discussion and implications were drawn.

Introduction

Computer-supported collaborative learning (CSCL) creates an online collaborative learning environment that prompts social construction of knowledge and harnesses students' interactive and collaboration skills (Kwon et al., 2014). The online collaborative writing tool, wiki, offers a space for dynamic co-authorship and has often been used to support CSCL in various disciplines (Chu et al., 2013). Despite that wikis record data of student activities during their learning process (e.g., page revision history), these data emerge in large amounts that could intimidate users and thus remain under-used (Kear et al., 2014). In recent years, the CSCL community has recognized learning analytics (LA) for understanding and supporting collaboration as a "top research priority" (Wise & Schwarz, 2017: p. 444). LA is defined as the analysis and visualization of data from learners and the learning environments for optimizing teaching and learning (Vieira et al., 2018). While LA has been shown effective in monitoring and assessing CSCL (e.g., Ng et al., 2019), prior investigations were largely conducted in university and secondary school classrooms or targeting teachers, while few studies have focused on primary school students' experience of LA-supported CSCL. The lack of research on LA in primary education could be attributed to the "cold-start problem", which refers to the insufficient scale of learning data for realizing the potential of LA (Donaldson et al., 2017: p. 6). Given the need to consider contextual factors when examining CSCL (Jeong & Hmelo-Silver, 2015), parents stand as an expected external influence on school-aged students' experience. In particular, parents of young learners might hold conflicting opinions on the use of digital devices for teaching and learning (Plowman et al., 2012). Nonetheless, the possible effects of parental attitudes on the implementation of LA for CSCL have remained largely unexplored. To bridge these gaps, this study sets out to answer the following two research questions:

RQ1: How do primary school students perceive their experience of analytics-supported CSCL?
RQ2: How do primary school students' self-perceived parental attitudes towards their use of digital technology influence their experience of analytics-supported CSCL?

Findings to the first research question will shed light on how LA-supported CSCL can be implemented to facilitate young students' collaborative learning experience, whereas the second research question can yield implications on whether parental attitudes need to be changed for optimizing students' CSCL experience.

Related work

Learning analytics in CSCL environment

Making sense of students' learning behaviour data serves as an overarching goal of LA tools such as dashboards (Vieira et al., 2018). Using wiki has been shown to improve the effectiveness of inquiry-based learning, where students co-construct knowledge through collaborative inquiry on a topic and jointly composing and editing content on wiki (Chu et al., 2013). In a wiki-based learning environment, teachers offer scaffolding on students' collaboration and monitor their progress on wiki (Jeong & Hmelo-Silver, 2015). However, the multi-level complex data buried in wikis present unique challenges for teachers in regulating and assessing CSCL (Kear et al., 2014). To mitigate these challenges, LA offers a robust and efficient approach to harvesting and presenting empirical evidence of students' collaborative learning (Wise & Schwarz, 2017). Existing LA tools designed for CSCL environments generate visualizations of statistics such as students' participation frequency and summaries of their collaboration, on both group and individual levels (Hu et al., 2016). In wiki-based CSCL environments, these LA tools help teachers orchestrate students' collaboration and monitor their progress more effectively and efficiently (Ng et al., 2019). They also provide formative feedback to students, supporting self and peer assessment of wiki contributions (Vieira et al., 2018). While few studies have demonstrated the benefits of LA-supported CSCL in the context of secondary education (e.g., Ng et al., 2019), there seems a lack of direct studies on primary school students' experience of LA-supported CSCL, particularly in a wiki-based environment.

Parental attitudes towards technology

A meta-review of 37 studies conducted in kindergartens and K-12 schools showed correlations between parents' general supervision of students' learning activities and their academic achievement (Castro et al., 2015). While parents have shown supportive attitudes towards students' literacy development, they tend to be hesitant towards the use of digital technology by young learners, due to concerns on children's physical and mental development (Anastasiades et al., 2008). Plowman and colleagues (2012: p. 32) reported case studies of 14 families with young learners on domestic technology-supported learning, concluding that parents held a "guarded" attitude and tended to limit their children's access to technology (e.g., reduced screen time). Given the important role played by parental attitudes in technology-supported education (Kostyrka-Allchorne et al., 2017), this study aims to explore how parental attitudes towards technology would influence students' experience of LA-supported CSCL.

Research context and methodology

This study was conducted in a primary school in a capital city in southwestern China in the spring semester of 2017. 46 Grade 5 students participated in this study as a non-compulsory after-class non-formal learning activity on improving their collaboration and inquiry-based learning skills. In groups of four to five, they worked on group inquiry projects on assigned topics for Chinese (e.g., local cuisines) and Mathematics subjects (e.g., distance and time measurements). Under the guidance of three teachers and two researchers, students wrote their project reports on a customized BlueSpice MediaWiki platform in a largely self-directed manner. In other words, besides initial introductions of the learning tasks, the teachers offered pedagogical assistance only when students raised questions. Each group worked on their own wikis that were not accessible by other groups, where each wiki consisted of pages based on the rationale of inquiry-based learning, including *Introduction*, *Information Collection and Analysis, Reflections*, etc. Adopting an interrupted time-series research design, an LA tool named Wikiglass (Hu et al., 2016) was introduced mid-way through the study such that half of the inquiry-based learning topics were carried out without LA support (i.e., pre-intervention) and the latter half had LA support (i.e., post-intervention). This design was to ensure fairness in terms of having all participants receive the intervention of LA support. Wikiglass fetches data from student wikis and visualizes statistics of students' contribution (e.g., number of words inputted on each wiki page) and participation (e.g., number of revisions) on class, group, and individual levels, for between-group and within-group comparisons of student progress. Both teachers and students could view visualizations of all groups and individual students involved. Figure 1 shows screenshots of Wikiglass displaying students' revision counts on wikis. The interface was in Chinese, the native language of the students and teachers. The researchers provided on-site training and produced instructional videos for using wiki and Wikiglass. Since computers in the computer room of this school were out of order and not repaired throughout the study, students used their own digital devices to access wiki and Wikiglass. In particular, 88.9% of the students owned a smartphone while only 20% had a computer at home.

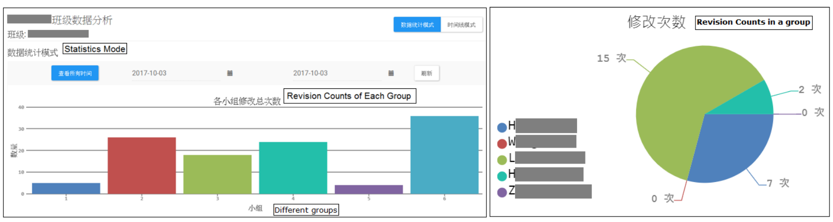


Figure 1. Screenshots of wiki statistics on class and group levels (with English translations) on Wikiglass

This study adopted a mixed method approach, collecting both quantitative (survey, system logs) and qualitative (interview) data. Ethical consent was sought from the principal, teachers, students, and parents prior to data collection. To explore student perceptions of their experience, a pre-intervention and a post-intervention questionnaires were administered on their perceptions towards inquiry-based learning (3 items), collaborative learning (5 items), wiki (3 items), and writing (5 items), as well as their self-perceived parents' attitudes towards their use of digital technology (6 items). Definitions of specialized items (e.g., inquiry-based learning) were provided in the questionnaires. A 6-point Likert scale (from 1 = Strongly Disagree to 6 = Strongly Agree) was adopted on the multiple-choice questionnaire responses. Internal consistency of items in these various constructs was measured by Cronbach alpha, ranging from 0.797 to 0.820. Semi-structured individual interviews were conducted with ten students and three teachers to elicit detailed elaboration of their experience and opinions on LA-supported CSCL. We also calculated statistics of students' behaviors from their wiki logs, such as number of words amended (i.e., added and deleted) and number of revisions. Non-parametric paired-samples Wilcoxon signed-rank tests and correlation analysis were conducted on the quantitative data collected, where statistical significance is indicated by * at $p < 0.05$ and ** at $p < 0.01$.

Findings and discussion

Table 1 shows the statistics of students' perceptions on various aspects of their learning experience (RQ1). Students' perceptions towards inquiry-based learning (5.15 vs. 5.48; p = .001**) and wiki (4.94 vs. 5.60; p = .000**) improved after their experience in LA-supported CSCL. As part of the skillset for inquiry-based learning, students deemed the CSCL experience useful for developing their information literacy (Chu et al., 2013), including "learning how to collect information" (Student interviewee S5). While students' inquiry-based learning in a CSCL environment necessitates teachers' monitoring, teachers expressed awareness of the "automatic computation and visualization of learning data" (Teacher interviewee T1) from Wikiglass that reduced their workload and, in turn, allowed them more time on teaching (Ng et al., 2019). Learning "how to use wiki" (S4) was also perceived as useful in cultivating students' digital literacy (Dede, 2010). Meanwhile, the information about students' progress on Wikiglass enabled them to "remind each other" (S9) during their collaborative work, indicating social cohesion and co-regulation among group members (Kwon et al., 2014). Nevertheless, students became less positive towards collaborative learning after their LA-supported CSCL experience (5.66 vs. 5.07; p = .000**). They had "rarely collaborated with each other" in their learning before this CSCL experience (S3) and thus might have had higher expectations on CSCL. In addition, it turned out that they did not always welcome "negative comments" from peers (T1) and felt uncomfortable when their own inputs were "removed due to different viewpoints among group members" (S7). While the inclination towards individualistic orientation in modernized society would influence how CSCL takes effect (Zhu, 2013), collaborating and compromising strategies are fundamental societal skills (Dede, 2010). This implies that when orchestrating collaboration, teachers should encourage and guide students to be more accepting towards each other's opinions.

Table 1: Statistics of students' perceptions of their analytics-supported CSCL experience (N=46)

Aspect of CSCL experience	Pre-intervention: Mean (SD)	Post-intervention: Mean (SD)	Sig.
Inquiry learning	5.15 (0.55)	5.48 (0.54)	.001**
Collaborative learning	5.66 (0.44)	5.07 (0.89)	.001**
Wiki	4.94 (0.71)	5.60 (0.58)	.000**

In relation to the second research question, a teacher remarked that parents would "limit their children's computer usage" because students might "abuse the e-learning time to play computer games" (T1). From students' perspective, their parents were accepting towards their use of mobile devices (e.g., smartphones) for learning purposes (mean = 5.09; SD = 1.31). Results of correlation analysis showed a significantly moderate positive correlation (Pearson's r = 0.311; $p < 0.05$*) between parents' allowance for students' use of mobile for learning and students' enjoyment of inquiry-based learning with wiki. Considering that these young learners usually own a mobile phone instead of a computer, parents' positive attitudes towards students' mobile usage demonstrates the potential of implementing mobile CSCL which also befits the trend of mobile learning analytics (Vieira et al., 2018). Interestingly, the more students perceived their parents were aware of their purpose of using their mobile phones, the lower the total number of words amended on wiki (r = -0.441; p = 0.017*). This implies that students might feel their technology usage being monitored by their parents and thereby being under pressure and limiting their online learning activities (Castro et al., 2015). Though how this occurred warrants further qualitative investigation, this can be seen as a possible parental influence on students' reduced participation in CSCL. An implication is that, in educational settings where technology is under-utilized, parents could be more informed of the pedagogical benefits of CSCL and how they could provide guidance to their children through more school-parent communication (Plowman et al., 2012).

Limitations and future work

A major limitation of this study is that parental attitudes were elicited from students' perceptions and teachers' observations rather than directly from parents. Future endeavours on analytics-supported CSCL will also collect parents' perceptions and opinions, and further examine the relationships between student perceptions and the quantities of their contributions and participation in the CSCL environment.

References

Anastasiades, P. S., Vitalaki, E., & Gertzakis, N. (2008). Collaborative learning activities at a distance via interactive videoconferencing in elementary schools: Parents' attitudes. *Computers & Education, 50*(4), 1527-1539.

Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro-Asencio, E., & Gaviria, J. L. (2015). Parental involvement on student academic achievement: A meta-analysis. *Educational research review, 14*, 33-46.

Chu, S. K. W, Siu, F., Liang, M., Capio, C. M., & Wu, W. W. (2013). Users' experiences and perceptions on using two wiki platforms for collaborative learning and knowledge management. *Online Information Review, 37*(2), 304-325.

Cole, D., Rengasamy, E., Batchelor, S., Pope, C., Riley, S., & Cunningham, A. M. (2017). Using social media to support small group learning. *BMC medical education, 17*(1), 201.

Dede, C. (2010). Comparing frameworks for 21st century skills. 21st century skills: Rethinking how students learn, 20, 51-76.

Donaldson, P., Ntarmos, N., & Portelli, K. (2017). *A Systematic Review of the Potential of Machine Learning and Data Science in Primary and Secondary Education*.

Hu, X., Ip, J., Sadaful, K., Lui, G., & Chu, S. (2016, April). Wikiglass: A learning analytic tool for visualizing collaborative wikis of secondary school students. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 550-551.

Jeong, H., & Hmelo-Silver, C. E. (2015). Research Questions and Research Methods in CSCL Research. *International Society of the Learning Sciences, Inc. [ISLS]*.

Kear, K., Donelan, H., & Williams, J. (2014). Using wikis for online group projects: Student and tutor perspectives. *The International Review of Research in Open and Distributed Learning, 15*(4).

Kostyrka-Allchorne, K., Cooper, N. R., & Simpson, A. (2017). Touchscreen generation: children's current media use, parental supervision methods and attitudes towards contemporary media. *Acta Paediatrica, 106*(4), 654-662.

Kwon, K., Liu, Y. H., & Johnson, L. P. (2014). Group regulation and social-emotional interactions observed in computer supported collaborative learning: Comparison between good vs. poor collaborators. *Computers & Education, 78*, 185-200.

Ng, J., Hu, X., Luo, M., & Chu, S. K. (2019). Relations among participation, fairness and performance in collaborative learning with Wiki-based analytics. *Proceedings of the Association for Information Science and Technology, 56*(1), 463-467.

Plowman, L., Stevenson, O., Stephen, C., & McPake, J. (2012). Preschool children's learning with technology at home. *Computers & Education, 59*(1), 30-37.

Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education, 122*, 119-135.

Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning, 12*(4), 423-467.

Zhu, C. (2013). The effect of cultural and school factors on the implementation of CSCL. *British Journal of Educational Technology, 44*(3), 484-501.

Acknowledgments

This study is partially supported by a grant (Project No. 61703357) from the National Natural Science Foundation of China and an Early Career Scheme grant (Project No. HKU 27401114) from the Research Grants Council of the Hong Kong S. A. R., China.

Augmented Reality in Collaborative Problem Solving: A Qualitative Study of Challenges and Solutions

Eduardo Hoppenstedt, Iulian Radu, Tugce Joy, Bertrand Schneider
ehoppenstedt@gse.harvard.edu, iulian_radu@gse.harvard.edu, joytugce@gmail.com, bertrand_schneider@gse.harvard.edu
Harvard Graduate School of Education

Abstract: Augmented reality (AR) can help learners solve problems more effectively during collaborative problem solving (CPS) experiences. In this study we investigate what kinds of difficulties and challenges were encountered by pairs of collaborators as they interacted with a tangible learning activity, with and without AR visualizations; and we suggest AR features to address those difficulties. We qualitatively analyze the behaviors of 8 groups, selected from a larger study in which participants interacted with an augmented speaker. We identify episodes where collaborators show struggles during CPS, discuss types of difficulties encountered, and use existing literature to suggest AR features that address the identified issues.

Introduction and related work

Collaborative problem solving (CPS), defined as the capacity of individuals to solve problems by joining their knowledge, skills, and efforts with others, is a fundamental competency to modern workspace and societal needs (OECD, 2017). Appropriate design of CPS learning experiences allows learners to develop the skills needed to successfully build knowledge and solve problems in collaborative settings. Existing research on CPS has been focused on measuring specific constructs to predict collaboration quality, by measuring interpersonal competencies (Oliveri et al., 2017). Augmented Reality (AR), which is the mixing and incorporation of virtual and real content (Milgram et al. 1995), has the potential to help collaborators during CPS activities. When investigating how new technologies such as augmented reality can be designed to improve collaborative learning, educational designers need to understand the specific difficulties that learners encounter in traditional CPS activities and what new technological features could be designed to address them. In this research we study how this can be achieved when designing augmented reality to improve CPS learning experiences.

Augmented reality has been shown to improve learning processes by reducing cognitive load, connecting intangible phenomena to observable graphic representations, or improve learning by making abstract concepts more intuitive, thus improving the retention of the learned materials (Radu, 2014). The use of AR media has been shown to impact dyads' behavior and communication when compared to non-AR groups (Unahalekhaka et al., 2019; Hung-Yuan, 2014), and research in higher education settings shows that the use of AR during CPS activities influences the time spent at different points in the CPS process such as organizing and interpreting data (Hung-Yuan, 2014). In this work, we contribute to research by comparing what difficulties are encountered as collaborators engage with an existing learning experience under AR and Non-AR conditions, with the goal of identifying new AR features that can be implemented to improve the CPS experience. We apply a granular analysis of learner behaviors to answer the following research questions: RQ1: What difficulties are encountered during the collaborative problem-solving learning activity, and how do they differ between AR vs. non-AR contexts? RQ2: What AR features can be used to reduce or prevent these difficulties from arising?

Methods

We use data collected from a previous study (Radu & Schneider, 2019) where 60 dyad pairs interacted with a loudspeaker system (Figure 1) designed for encouraging learner explorations of how sound, electric current and magnetic fields are related. Participants could interact with the system by playing music through a headphone connection from a smartphone, pushing buttons to change the amplification and direction of electricity, and moving the speaker membrane (a cup) to explore its effects on sound output. Study participants were split into two groups, either seeing augmented reality educational visualizations ("AR" group), or not ("Non-AR" group). Participants in the AR group received the same system, physical tools, and printed poster information as Non-AR group, but also could see the poster information as represented in augmented reality visualizations (Figure 1) shown as component labels, and visualizations of audio sound waves, electric currents, and magnetic fields. These representations were provided in static form for the Non-AR condition. For this paper we qualitatively analyzed a subset of 8 groups, where half saw educational AR (AR group) and half did not (Non-AR groups). Half the groups were selected having strong collaboration scores and half low scores, using the rubric from (Meier et al., 2007), done to ensure variety of collaborations and balance within conditions, and was not analyzed separately.

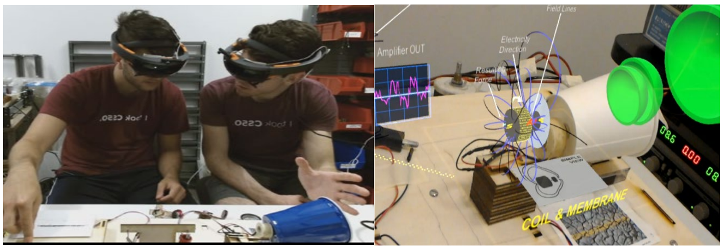


Figure 1. Two users interacting with the system (left). Physical model with educational AR overlays (right) showing electricity (yellow electrons, blue bar charts), magnetic fields (blue lines), audio visualizations (green).

To answer the research questions, we analyzed data using a 2-phase process, which identified types of collaborative problem solving difficulties and determined what AR features may address those difficulties. The first phase involved analyzing videos to detect possible indicators of difficulties in collaboration or problem solving. To detect such moments, we created a coding scheme to identify instances where participants may be having difficulties in the categories of searching for information, communicating information, or leading the activity. The coding scheme was applied to videos from the 8 groups (total 5.36 hours) in 15-second intervals. Each code was tagged if it was present during a 15-second time interval, and multiple codes could be tagged for the same time interval. For inter-rater reliability (IRR), two raters coded 20% of the videos and reached a Cohen-Kappa of .85, which implies substantial agreement. After the IRR test, one researcher coded the remaining 80% of the videos. We then reviewed the tagged time intervals to identify episodes of active collaboration, defined as time periods where at least 3 of the behaviors in the coding scheme were marked during 2 intervals (30 seconds) or more. In the second phase, the video recordings were reviewed by one researcher during those time intervals of active collaboration, to determine the specific difficulties participants encountered during those episodes, and using thematic coding to yield the results presented below. After identifying these categories, we generated augmented reality features that may prevent such problems, by reflecting on existing AR literature.

Results

In this section we discuss the difficulties identified during CPS learning, highlighting differences between AR and Non-AR groups, and we present features that AR systems may provide to address these difficulties. When examining the 8 group videos for this study, we marked a total of 710 time intervals where at least one behavior in the coding scheme was present, yielding 40 episodes of collaboration difficulties. 18 episodes arised in AR groups, and 22 in Non-AR groups. The 40 episodes were organized into the following 5 problem categories:

Lacking Understanding of Representations: This category happened to 7 of the 8 groups (3 AR, 4 Non-AR) and includes moments where participants struggle to reach consensus about how a phenomenon should be defined or represented. This includes situations where reaching consensus is difficult due to imbalanced knowledge, or due to participants misinterpreting the representations from the AR system, or struggling to represent imaginary shapes with available tools. The source of the problem is the difficulty of representing intangible phenomena when the concept has not been understood by both participants. While Non-AR had difficulties discussing the static information accessible to them, the AR groups encountered some issues interpreting 3D dynamic information. AR environments can provide capabilities for participants to generate other representations of invisible phenomena, such as through drawings and gestures. AR features can allow participants to create 3D structures (Kaufmann, 2003) or draw in 3D space to enhance mutual understanding (He et al., 2019).

Lacking Perceptual Information and Causal Relationships: This category happened to 7 of the 8 groups (3 AR, 4 Non-AR), and contains difficulties originated by one or both participants expressing inability to perceive a phenomenon because of perceptual issues such as their vision being blocked, or inability to hear the speaker. Blocked access to sensory information limits participants' capability to understand the problem at hand and to engage in collaborative learning, because participants are unable to discover the connections and ideas necessary to build knowledge and understand the explanations of the studied phenomena. AR features such as show-through techniques can help users to see the objects that are occluded by others (Argelaguet et al., 2010). The users could also be alerted when there is an important change in the system (García et al., 2008), and when in doubt of their decisions, they can use reviewing tools to compare their options (Xia et al., 2018).

Lack of Awareness of Other Person: This category occurred in 4 of 8 groups (1 AR, 3 Non-AR), and includes episodes where a participant is not aware of what their peer is focusing on. This includes cases when participants do not know each other's actions, or they cannot confirm if the peer is aware of a situation, impacting the problem-solving process or slowing the dyad's progress when solving tasks. When participants have poor awareness this limits their engagement with each other and with materials and resources and impairs participant's opportunities to contribute during discussions and reflections (Gutwin & Greenberg, 2002; Mathieu et al., 2000).

In our study, groups with AR visualizations, such as the virtual text labels and diagrams, encountered fewer difficulties while remaining aware of each other. We suggest AR features to further increase peer awareness and prevent difficulties in this category. Virtual pointers have been shown to help peers understand what a collaborator is referring to (Bauer, 1999). Field of view indicators can help learners stay aware of what objects are inside their peer's vision, and enhance collaboration (Piumsomboon et al., 2019). Or, participants could see from each other's point of view via a small AR video window, such as presented in (He et al., 2019).

Lacking Easy Access to Information and Resources: This category occurred in 4 of 8 groups (1 AR, 3 Non-AR), and refers to episodes where either a participant lacks access to resources, or the person is left out of participating while the peer controls the tools and system, or participants struggle to link information from the wall poster to the physical system. Previous research of the same AR speaker system found that dominant behaviors are less detrimental to collaboration when participants have AR visualizations, likely because both participants have easier access and improved visibility of information (Radu & Schneider, 2019b). To enhance this ability, AR systems can employ sharing one's viewpoint with peers (Szalavári et al., 1998) or use a snapshot to prevent situations where information may be lost over time (Lee et al., 2020). The system can also be designed to encourage participation by requiring both collaborators to take an action to proceed (Piumsomboon et al., 2019).

Lacking Memory or Background knowledge: This category occurred in 4 of 8 groups (1 AR, 3 Non-AR), and refers to events where participants are confused about the name or function of an object or tool or do not have the vocabulary to describe an object or phenomenon, and situations where participants lack the necessary background knowledge to progress in the task or an inability to remember the past actions during an experimentation process. This limits learners' performance during CPS activities and their ability to progress in the task. Of all groups exhibiting this category of events, indicating that AR representations such as text labels and visualizations helped participants to remember the names and functions of system components. Features such as virtually writing or drawing directly on top of the system objects may help reduce problems related to definitions and memory (Aschenbrenner et al., 2018). AR can also help users track what system components have been interacted with (Benko et al., 2004) and a collaborative session could be recorded and replayed (Greenwald et al., 2019). Users can also explore different steps of an AR pre-recorded tutorial, an approach that has shown to encourage learning and experimentation (Kaufmann, 2003).

Discussion

The results show that difficulties occur in both AR and Non-AR settings, but that augmented reality visualizations mediate the kinds of difficulties encountered during CPS processes. For example, in the case of lacking perceptual information, AR participants had problems related to key information being blocked by their position or their peers, thus increasing the difficulty of solving the task, while in the same category the Non-AR dyads had more problems related to limitations in their perception such as an inability to detect sound or movement. Moreover, problems grouped in the categories of lacking memory or background knowledge, as well as lacking easy access to information resources, were predominantly populated by Non-AR dyads. Thus, while both AR and Non-AR dyads experienced problems related to all categories, the nature of the problems did vary, revealing an effect of AR during CPS learning activities.

Designers working with other educational technologies can benefit from using this process to identify problems and needs. Educators aiming to incorporate AR in their classroom might find it useful to use the listed features in order to prevent the detected problems from happening, thus avoiding the identified problems from impairing student learning. It is important to note the limitation that the problems detected in this study are from a small sample (8 groups) and are heavily influenced by the design of the speaker activity. Future work should observe larger datasets and different activities to find commonalities and expand our understanding of issues that arise during CPS learning activities. Additionally, the method used in this research could be useful for designers working with other technologies to make sure their developments address existing problems and improve the quality of learning activities.

In conclusion, we found that AR and Non-AR participants encounter problems during the CPS learning activities, but the nature of problems change, and the saliency of the problem varies according to the AR and Non-AR conditions. Identifying these problems is useful to ideate AR system features that allow learners to spend less time on difficulties and instead focus on exploration, interpretation, reflection, and learning.

References

Argelaguet, F., Kunert, A., Kulik, A., & Froehlich, B. (2010, March). Improving co-located collaboration with show-through techniques. In 2010 IEEE Symposium on 3D User Interfaces (3DUI) (pp. 55-62). IEEE.

Aschenbrenner, D., Rojkov, M., Leutert, F., Verlinden, J., Lukosch, S., Latoschik, M. E., & Schilling, K. (2018, October). Comparing different augmented reality support applications for cooperative repair of an

industrial robot. In 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) (pp. 69-74). IEEE.
Bauer, M., Kortuem, G., & Segall, Z. (1999, October). " Where are you pointing at?" A study of remote collaboration in a wearable videoconference system. In Digest of Papers. Third International Symposium on Wearable Computers (pp. 151-158). IEEE.
Benko, H., Ishak, E. W., & Feiner, S. (2004, November). Collaborative mixed reality visualization of an archaeological excavation. In Third IEEE and ACM International Symposium on Mixed and Augmented Reality (pp. 132-140). IEEE.
García, A. S., Molina, J. P., Martínez, D., & González, P. (2008, December). Enhancing collaborative manipulation through the use of feedback and awareness in CVEs. In Proceedings of the 7th ACM SIGGRAPH international Conference on Virtual-Reality Continuum and Its Applications in industry (pp. 1-5).
Greenwald, S. W., Corning, W., McDowell, G., Maes, P., & Belcher, J. (2019). ElectroVR: An Electrostatic Playground for Collaborative, Simulation-Based Exploratory Learning in Immersive Virtual Reality.
Gutwin, C., Greenberg, S.. (2002). A Descriptive Framework of Workspace Awareness for Real-Time Groupware. Computer Supported Cooperative Work, 11(3-4), 411-446.
Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance.Journal of Applied Psychology, 85(2), 272–283.
Hung-Yuan W., Henry Been-Lirn D., Nai Li, Tzung-Jin L., Chin-Chung T., (2014). An Investigation of University Students' Collaborative Inquiry Learning Behaviors in an Augmented Reality Simulation and a Traditional Simulation. Journal of Science Education and Technology, 23(5), 682-691.
Kaufmann, H. (2003). Collaborative augmented reality in education. Institute of Software Technology and Interactive Systems, Vienna University of Technology.
Lee, G., Kang, H., Lee, J., & Han, J. (2020, March). A User Study on View-sharing Techniques for One-to-Many Mixed Reality Collaborations. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 343-352). IEEE.
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.
Milgram, Paul, Haruo Takemura, Akira Utsumi, and Fumio Kishino. "Augmented reality: A class of displays on the reality-virtuality continuum." In Telemanipulator and telepresence technologies, vol. 2351, pp. 282-292. International Society for Optics and Photonics, 1995.
OECD (2017), PISA 2015 Results (Volume V): Collaborative Problem Solving, PISA, OECD Publishing, Paris.
Oliveri, M. E., Lawless, R., & Molloy, H. (2017). A literature review on collaborative problem solving for college and workforce readiness. ETS Research Report Series, 2017(1), 1-27.
Piumsomboon, T., Dey, A., Ens, B., Lee, G., & Billinghurst, M. (2019). The effects of sharing awareness cues in collaborative mixed reality. Frontiers in Robotics and AI, 6(5), 02.
Radu, I., & Schneider, B. (2019). What Can We Learn from Augmented Reality (AR)? Benefits and Drawbacks of AR for Inquiry-based Learning of Physics. In 2019 CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019).
Radu, I., & Schneider, B. (2019b). Impacts of Augmented Reality on Collaborative Physics Learning, Leadership, and Knowledge Imbalance.
Radu, I., Hv, V., & Schneider, B. (2021). Unequal Impacts of Augmented Reality on Learning and Collaboration During Robot Programming with Peers. Proceedings of the ACM on Human-Computer Interaction, 4(CSCW3), 1-23.
Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. International Journal of Computer-Supported Collaborative Learning, 8(4), 375–397.
Unahalekhaka, A., Radu, I., & Schneider, B. (2019). How Augmented Reality Affects Collaborative Learning of Physics: a Qualitative Analysis.
Xia, H., Herscher, S., Perlin, K., & Wigdor, D. (2018, October). Spacetime: enabling fluid individual and collaborative editing in virtual reality. In Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (pp. 853-866).

Acknowledgements

This material is supported in part by the National Science Foundation grant no. 1917716.

Measuring Disagreement in Learning Groups as a Basis for Identifying and Discussing Controversial Judgements

Nils Malzahn, Farbod Aprin, H. Ulrich Hoppe
nm@rias-institute.eu, fa@rias-institute.eu, uh@rias-institute.eu
Rhine-Ruhr Institute for Applied Systems Innovation (RIAS), Duisburg, Germany
Sabrina C. Eimler, Sarah Moder
Sabrina.Eimler@hs-ruhrwest.de, sarah.moder@hs-ruhrwest.de
University of Applied Sciences Ruhr West (HRW), Bottrop, Germany

Abstract: Learning scenarios that build on socio-cognitive conflict as a trigger of learning constitute an established approach in collaborative learning. The identification of disagreement is an important premise for this approach. We have selected a measure of disagreement based on a comparative mathematical analysis and have applied it in the context of learning about toxic phenomena and discrimination in social media. The data collected in an online study have been used to test the disagreement measure in combination with a game-based tagging tool.

Introduction: Disagreement as a pedagogical opportunity

The practice of computer-supported collaborative learning (CSCL) in classrooms and other settings is largely based on certain ways and patterns of structuring group activities. CSCL scripting (cf. Fischer et al., 2007) is an approach that imposes an *explicit process structure* on the group activity. The “jigsaw” method fosters collaborative knowledge exchange and knowledge building in a group through inducing a certain *distribution of knowledge*. Other approaches, especially in the context of learning driven by argumentation (Andriessen, Baker, & Suthers, 2003), are based on *role models* that guide the interaction between the multiple parties taking part in the argumentation. “ArgueGraph” (Jermann & Dillenbourg, 1999) is an example that combines a scripting approach with roles in an argumentative context between two partners. The interacting dyads are selected based on contrasting prior opinions on the subject to be discussed. ArgueGraph is one of the early examples of CSCL scenarios that use “socio-cognitive conflict” (Mugny & Doise, 1978) as a driver of shared knowledge exchange and knowledge revision. Asterhan et al. (2010) have analyzed the influence of motivational and affective factors in learning scenarios based on “socio-cognitive conflict”. Ideally, the conflict should be a trigger of cognitive activity and engagement, yet not emotionally destructive for the social interaction (cf. Näykki et al., 2014). Many of the relevant factors and options to be considered when implementing learning through conflict and controversy in the classroom, such as individual heterogeneity, availability of information, or perspective taking skills, have already been identified and discussed by Johnson and Johnson (1979). These authors also point to the importance of *disagreement* as an opportunity and trigger for learning. The most obvious and most frequently addressed cases and areas where disagreement arises have to do with opinions and (ethical and other) judgements.

The work reported here aims at using controversy as a driver of learning about discriminatory and toxic effects of social media. There is evidence that these phenomena are widespread in the targeted age group of junior high-school students (Schultze-Krumbholz et al., 2012). The basic “pedagogical workflow” in our intended classroom scenario would start with an individual activity in which the learners classify given items or instances of possibly problematic social media content using different predefined categories (“sexism”, “hate speech” etc.). The individual judgements are stored in a classroom repository that feeds into a teacher dashboard in which the items appear ordered and grouped according to their degree of controversy. The teacher may select examples from the repository for plenary or small-group discussions using the level of controversy or disagreement as a clue.

In the following, we investigate the potential of assessing disagreement in a given group of learners through a systematic, statistical comparison of individual judgements. Based on a mathematical analysis and comparison of several approaches, we have selected a specific measure of disagreement. Due to the restrictions imposed by the COVID-19 pandemic, we have not been able to orchestrate and test our scenario in face-to-face classroom settings. However, we have collected data on an individual level in an online study and checked the practicality of the disagreement measurement in comparison to other indicators such as response times and agreement with expert ratings. Finally, we discuss these findings and their relevance for further applications in offline and online educational settings.

Measures of disagreement

The premises of the situation in which we want to calculate the disagreement between student judgements are these: We have a given set of items (Instagram-style text-decorated images) that are classified (i.e. tagged) by a certain number of raters (the students in the learning group) by selection of one out of a set of pre-defined labels or tags. This means that we must compare multiple raters who rate multiple items. Here, the actual ratings (i.e. tags) are defined on a nominal scale, i.e. without a given inherent order. This excludes the use of most dispersion measures from descriptive statistics and leaves only few options. Among these is the "dispersion index" (Walker, 1999). We have also found one measure of disagreement ("group disagreement") that was genuinely conceived from the perspective of collaboration research (Whitworth, 2007). There is a direct correspondence of measures of disagreement (***D***) with measures of agreement (***A***). If these measures are normalized on a scale ranging from ***0*** to ***1***, this correspondence is expressed by the equation ***D = 1 – A***. This suggests that known measures of agreement, such as those used to calculate inter-rater reliability, could be used in the inverse way. Given that we must deal with multiple raters and a nominal scale, Fleiss' kappa (Fleiss, 1971) is a candidate in this group.

In a mathematical analysis and comparison of these measures, which cannot be reproduced here under the given limitations, we have found that Fleiss' kappa is the A-measure that exactly corresponds to Whitworth's group disagreement (***GD***), i.e. it equals ***1 - GD***. As already noted by Whitworth (2007), for a high number of raters the maximum value of ***GD*** tends to approach ***(K – 1) / K*** with ***K*** representing the number of categories. I.e., this cap is smaller for a smaller number of categories and amounts to ***0.5*** for ***K = 2***. The dispersion index (***DI***) uses another normalization factor that corrects for this cap in the range of values. Interestingly, these measures of very different provenance only differ in the normalization factor and we have chosen ***DI*** as a disagreement measure because of its better scaling behavior. The measure is calculated in the following way:

$$\boldsymbol{DI} = \frac{K\left(N^2 - \sum_{k=1}^{K} {f_k}^2\right)}{N^2\,(K-1)}$$

N: *Number of raters*
K: *Number of categories*
f_k: *Sum of ratings for each category (frequency)*
$\sum f_k^2$: *Sum of squared frequencies/ratings*

Pedagogical goals and tool design

In the on-going project "Courage", a group of European researchers works on building a learning environment to support and train teenage school students in developing strategies to cope with discrimination and toxic content in social media. The targeted phenomena include the distribution of fake news and conspiracy theories, as well as direct discriminatory practices such as hate speech, bullying, and cyber-mobbing. The primary goal is to reduce toxic effects by improving self-protection skills or resilience. The approach is based on fostering understanding instead of avoidance or external protection (e.g., by censoring or filtering).

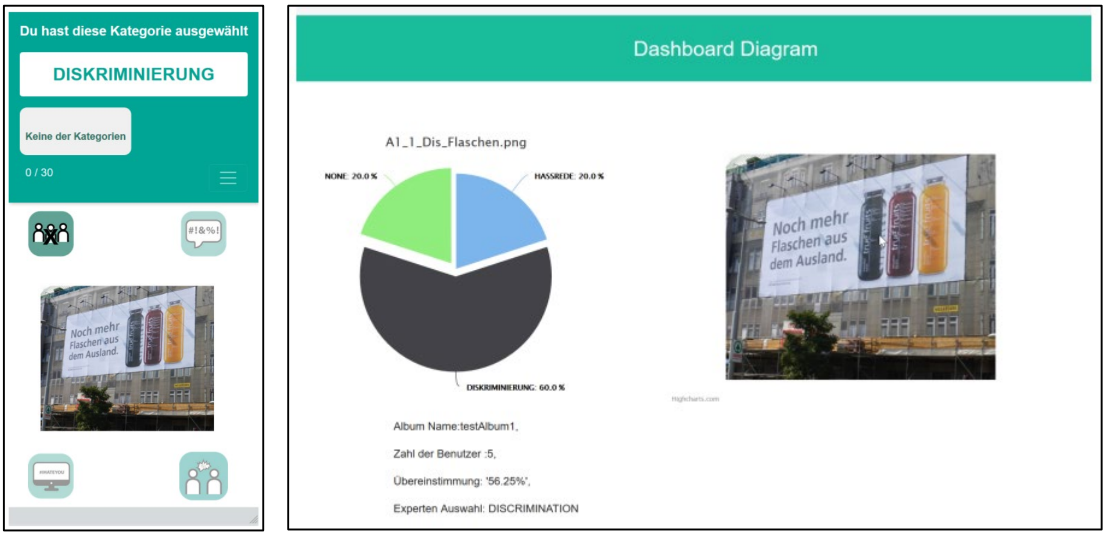


Figure 1. The SwipeIt app (left) with a corresponding view of the teacher dashboard (right).

As a first step, we have designed a serious minigame (the mobile app "SwipeIt") that allows for a playful and controllable interaction of young learners with potentially toxic content items. The app was designed to be used in a classroom context to support classroom discussions around example social media items. The labels or tags to be used in our scenario and study were selected based on semi-structured interviews with teachers and focus groups with teenagers. This was double-checked with terms from literature. The result was a set of four labels (terms): "verbal violence", "hate speech", "discrimination" and "cybermobbing". A set of 30 images representing such phenomena was chosen from a total of 142 images from various social media platforms such as Facebook and Instagram. All images were independently tagged by two experts with a psychology background.

The SwipeIt app displays these 30 images in fixed order. Each user is asked to select the label that best describes the current image. If none of the labels is considered to be adequate, the user may select the option "None of the categories". Figure 1 (left) shows the SwipeIt application. Here, the user has currently selected the tag "discrimination" (German: "Diskriminierung"). The user may swipe the image afterwards in any direction to see the next image as an image cannot be revisited after having decided on a label. Every interaction, including any button-press, together with the final selection for each image, is stored in a database.

Experimental setting

The original scenario was planned to take place in computer classrooms at selected schools. The average duration of each trial should last around 90 minutes subdivided in different phases. In the beginning, the teenagers would be introduced to the overall topic and to the labels used as tags in the SwipeIt app to ensure a common understanding. Then the students would fill in a questionnaire hosted at SoSci Survey. Individual codes generated on beforehand would be distributed giving access to the questionnaire and allowing a match of questionnaire data (across different iterations) with the SwipeIt results. Participants would be redirected from the questionnaire to the game comprising 30 images to be tagged by every student. After tagging the images, a classroom discussion around more or less controversial examples images should take place. To support this, SwipeIt ranks the images according to the level of agreement for an item using the dispersion index ***DI***. A teacher dashboard provides an overview of all images ranked by degree of controversy and allows for inspecting the distribution of labels for each image (see Fig. 1, right).

Due to the COVID-19 pandemic, the experiment had to be modified as an online scenario without classroom interaction. The face-to-face introduction was replaced by the provision of introductory texts in a Moodle environment. To avoid confronting young adults with toxic content without any counsel by a teacher or researcher, the scenario was moved from schools to university students (entry level, all legal age). The 47 participating students (44 males, 3 females) received credits for participation in empirical studies required in their HCI study program. In the experimental online scenario, participants would be transferred from the Moodle environment to the online questionnaire and then to the SwipeIt app. To preserve anonymity, randomly assigned individual tokens were propagated between the Moodle environment, the questionnaire and SwipeIt avoiding separate authentication procedures and allowing to track the students' inputs across the different tools.

Specific research questions

Since we could only run the study without the following classroom and possibly small group interaction, our main goal has been the validation of the agreement measure with respect to its practical usefulness to trigger controversy-based further interactions. We have analyzed the data particularly with these questions in mind:

- Given that the credits would not depend on answer quality, do the participants actually make an effort to meaningfully and adequately characterize the examples? Indicators for this would be response times and agreement rates. It might also be that the participants tend to take the task less seriously towards the end of the completion process. This should lead to a decrease in response time and agreement rates.
- Is answer time (time spent on one image) inter-related with "controversiality" (disagreement)?
- For 25 out of the 30 examples, we have expert ratings that coincide in terms of a unique category assignment. How do the student classifications (agreement levels) compare to these expert ratings?

Data analysis

For the data analysis conducted so far, we have used the database with user ratings, expert ratings (one per image), and time stamps of the user actions. From the time stamps we have calculated answer times per image and user. We have also calculated disagreement (***DI***) and agreement (***1 – DI***) per image. For five items there was not enough agreement between the experts so that no expert rating was assigned. On this basis, we have compared two variables: the agreement of user ratings with the expert tagging (if available), i.e. the fraction of user tags that coincide with the expert tag, and the agreement between the participant ratings measured by ***1 – DI***. We found a Pearson correlation of $r = .71$ ($p < .0001$) between these two parameters. Of course, a high agreement with the expert rating would necessarily go along with a high (yet possibly smaller) agreement between user ratings, yet not necessarily vice versa. Practically, this implies that we may rely on inter-user agreement even in absence of an expert ground truth. It also indicates that the user judgements are not arbitrary. Regarding answer time and disagreement (DI), there was no significant correlation ($r = .14$, $p = .5$). This rules out the option to use answer time as an indicator for controversiality (mediated by individual insecurity). To capture sequence effects in the progression through the images we have correlated the image numbers (steps) with disagreement ($r = -.10$, $p = .62$) and answer time ($r = -.14$, $p = .50$). This is a positive message since it indicates that there is no significant deterioration of the rating behavior when progressing through the sequence of items. A certain reduction of answer time could also be explained by a procedural learning effect (routinization) in handling the app.

Discussion and outlook

We see the work reported here as a first step in providing an analytics-based underpinning for collaborative learning scenarios that rely on detecting and responding to disagreement as an opportunity for discussion and group learning. Our approach can serve as a tool to inform and orchestrate classroom scenarios based on socio-cognitive conflict. The first challenge that we have tackled here is the operationalization of disagreement as an important premise. Although we have not been able to test the fully-fledged approach in collaborative learning settings, we have seen that the interplay between analytical instruments and experimental tools "works" so that it provides a reasonable practical basis for further experimentation. Our specific goal of training young learners to better understand and handle toxic phenomena in social media is particularly susceptible to approaches that build on controversy and disagreement. We are currently replicating our study to enlarge our database, which should allow us to include demographic and personality factors in the analysis.

Given the on-going restrictions related to presence-based classroom experimentation, we intend to extend our online scenario with group interactions. One challenge here is the preservation of anonymity of the individual judgements. So far, our scenario does not require a combination of the internal user IDs with real identities, as long as the point is to identify the controversiality of items or artifacts. This would be different if we wanted to introduce group formation based on the characterization of users. We are currently favoring solutions that would not make use of such information, still focusing on the attribution of controversiality to the artifacts.

References

Andriessen, J., Baker, M., & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. In J. Andriessen, M.J. Baker & D. Suthers (Eds.), Arguing to Learn (pp. 1–25). Springer, Dordrecht.

Asterhan, C., Schwarz, B., Butler, R., Butera, F., Darnon, C., Nokes, T., ... & Sinatra, G. M. (2010). Motivation and affect in peer argumentation and socio-cognitive conflict. In Proceedings of the 9th International Conference of the Learning Sciences (ICLS 2010). Chicago (USA).

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. Psychological Bulletin, 76 (5), 378–382.

Jermann, P. & Dillenbourg, P. (1999). An analysis of learner arguments in a collective learning environment. Proceedings of the 3rd Conference on Computer-Supported Collaborative Learning (CSCL). Stanford (USA). pp. 265–273. hal-00190257

Johnson, D. W., & Johnson, R. T. (1979). Conflict in the classroom: Controversy and learning. Review of Educational Research, 49(1), 51–69.

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. European Journal of Social Psychology, 8(2), 181–192.

Näykki, P., Järvelä, S., Kirschner, P. A., & Järvenoja, H. (2014). Socio-emotional conflict in collaborative learning - A process-oriented case study in a higher education context. International Journal of Educational Research, 68, 1–14.

Schultze-Krumbholz, A., Jäkel, A., Schultze, M., & Scheithauer, H. (2012). Emotional and behavioural problems in the context of cyberbullying: a longitudinal study among German adolescents. Emotional and Behavioural Difficulties, 17 (3-4), 329–345.

Walker, J. (1999). Statistics in Criminal Justice: Analysis and Interpretation. Jones & Bartlett Learning.

Whitworth, B. (2007). Measuring disagreement. In R. A. Reynolds, R. Woods, & J. D. Baker (Eds.), Handbook of Research on Electronic Surveys and Measurements (pp. 174–187). IGI Global.

Acknowledgement

This work has been partially supported by the Volkswagen Foundation in the line "AI and the Society of the Future". We thankfully recognize this support and take the full responsibility for the content of this publication on us. We also thank Clarissa Arlinghaus, Stefan Sommer and Lena Hilbig for their support in material preparation and data collection.

International Society of the Learning Sciences

Symposia

 © ISLS

International Society of
the Learning Sciences

Advancing Technology Environments for Learning Communities

Anja Amundrud, University of Osla, anja.amundrud@iped.uio.no
Ole Smørdal, University of Oslo, ole.smordal@uv.uio.no
Jianwei Zhang, University at Albany, SUNY, jzhang1@albany.edu
Mei-Hwa Chen, University at Albany, SUNY, mchen@albany.edu
Jeremiah (Remi) Kalir, University of Colorado Denver, remi.kalir@ucdenver.edu
Bodong Chen, University of Minnesota, chenbd@umn.edu
Yu-Hui Chang, University of Minnesota, chan1173@umn.edu
David Groos, Minneapolis Public Schools, David.Groos@mpls.k12.mn.us
Joel P. Wiebe (co-chair), University of Toronto, joel.wiebe@mail.utoronto.ca
James D. Slotta (co-chair), University of Toronto, jslotta@gmail.com
Yotam Hod (discussant), University of Haifa, yotamhod24@gmail.com

Abstract: This symposium will examine and envision new possibilities to design next generation technology environments for advancing the study of classroom learning communities as a pedagogical approach. Of the many different kinds of technology environments used in educational research, technology designed for learning communities represents a unique genre in which the environments must support particular epistemic values, modes of learning, and discourse. The talks in this session represent five distinct projects, each focusing on the features of a technological environment and how they support learning communities (e.g., making learners' ideas salient; representing community knowledge; enabling idea interaction across boundaries; fostering a sense of progress). Across the papers, we identify cross-cutting research priorities and common technological elements that characterize this pivotal research area, with implications for future research and the development of community-supporting technology.

Introduction

Within the learning sciences, the 1990s gave rise to a *sociocultural revolution* that produced conceptual shifts whereby learning was conceptualized as inherently social and cultural activities and tools are emphasized (Hod, Bielaczyc, & Ben-Zvi, 2018). A good example of this shift is the so-called "learning community approach" (Bielaczyc & Collins, 1999, 2006) in which students are engaged in a collective effort toward understanding and advancing their community knowledge. In recent decades, a small community of researchers has been actively exploring topics such as collective epistemology (Acosta et al, 2018), knowledge building discourse (Chen & Hong, 2016), collective inquiry designs (Slotta, Quintana & Moher, 2018), microblogging (short and real-time digital posts) combined with 'talk rules' (Rasmussen et al., 2019), and cross-community collaboration (Zhang et al., 2018). Within this research, a new generation of methods and technologies have emerged, including support for automated discourse analysis and learning analytics for live student feedback (Stahl, 2015). These technologies scaffold student interactions, support community knowledge construction, and foster a sense of collective progress and idea improvement (Cress, Stahl, Ludvigsen, & Law, 2015) – all with the ultimate aim of transforming classrooms into learning communities.

Against this background, our field still confronts a significant challenge in designing such *technology environments for learning communities*, as well as supporting their uptakes in classrooms. Our world is facing the serious consequences of an ongoing pandemic, climate change, political polarization, and misinformation. These challenges cannot be fully explained by individual cognition but are deeply embedded in social and cultural norms and practices, which are segregated by attention-grabbing social media platforms. Recently, the co-editors-in-chief of ijCSCL have advocated for the CSCL community to help illuminate how to design technological settings for group collaboration and how people live and learn as we collectively explore what post-COVID education will become (Järvelä & Rosé, 2020). More than ever, researchers face core conceptual and design challenges in motivating and sustaining collaborative inquiry and knowledge-building dialogues, informing students' dynamic and creative interactions with ideas, and supporting knowledge building across social levels (individual, small groups, community, and open networks) and over sustained durations of time.

This symposium brings together five different research groups, each investigating its own particular questions relating to learning communities. Together we offer a suite of technologies for learning communities grounded in core design principles and classroom research: (1) Talkwall is a socially oriented technology for increasing and shaping class participation, (2) Idea Thread Mapper expands student collaborative discourse across multiple classrooms with continual idea build-on across school years, (3) CROWDLAAERS is a dashboard of social analytics for annotations of the web to visualize, make sense of and (re)mediate knowledge construction, (4) IdeaMagnets brings student discourse beyond the classroom by combining in-class knowledge building with public discourse on the web, (5) SCORE is a curriculum authoring, orchestration, and student learning environment for hybrid learning designs. Each tool has been designed based on the distinct theoretical and methodological perspectives of its authors (e.g., knowledge building, social networks, scripting, learning analytics). Yet the environments share commonalities (e.g., student discourse and inquiry) and offer complementarities as well as cross-cutting insights. This session will advance our exploration of these possibilities, including scaffolded discourse, social annotations, social learning analytics, orchestration of complex learning activities and interactions, and visualization of community knowledge.

This session will include two parts. First, totaling 60 minutes, each presenter will introduce their theoretical perspective and the corresponding technology environment within the context of a current research study, with time for questions after each presentation. Second, totaling 30 minutes, our discussant (Yotam Hod, from the University of Haifa, Israel) will facilitate a panel discussion with the presenters to engage in conversations across the different technologies and perspectives, and engage audience members in questions and answers. Together with the audience, we will identify opportunities for connecting these next-generation technologies to advance our field's investigation and support of learning communities.

Talkwall - Microblogging for students' reasoning through sharing ideas

Anja Amundrud and Ole Smørdal

Since 2008, we have explored socially oriented technologies to improve the quality of talk and increase the level of participation in classrooms, using various off-the-shelf platforms such as wikis, chats, and microblogging tools. How students and teachers talk together is often implicit and based on historical precedent, such as teachers asking questions or students should wait for their turn to talk. We have worked with teachers to grow a dialogic classroom culture through the introduction and maintenance of explicit 'talk rules' (Rasmussen et al., 2019), aiming to potentially transform pedagogical practices fostering the 'complex competencies' today's students need (for instance, by critically thinking about one's own ideas and how they relate to ideas of others through elaboration and reasoning). Recently, we have developed Talkwall (see figure 1), a microblogging tool aiming to better support same-time, same-place interactions, in line with the specific research-based understandings of dialogic teaching in the Thinking Together approach (Mercer et al., 1999). Here, we present our design, situated within a case where students (aged 11-12) use Talkwall to separate facts from opinions. The findings show that microblogging can provide new possibilities for peer interactions by systematically enabling students to access more of their peers' opinions, produce and discuss collective ideas and contribute and participate in productive talk. In particular, the creative use of a concept tagging activity proved to be a well suited resource for the facilitation of peer interactions where students practice their reasoning together. Microblogging can increase participation, allow for new perspectives to enter dialogues, and support teachers and students in conveying ideas from small group to whole-class interactions.

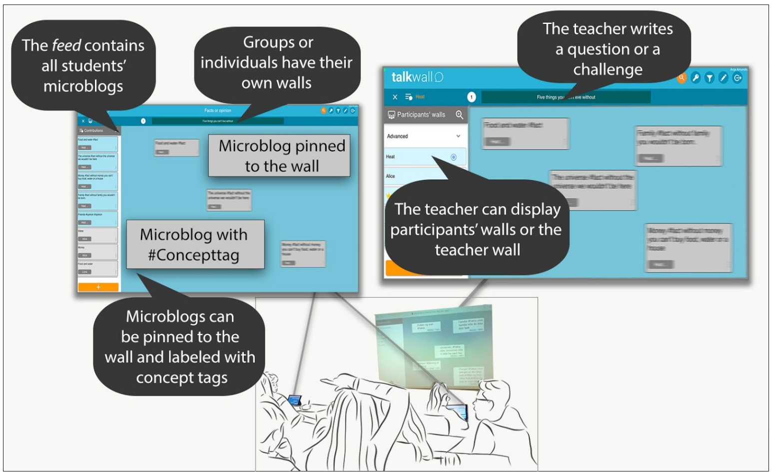


Figure 1. Talkwall being used in a classroom

In order to analyse how ideas travel between group interactions and whole-class conversations, a research tool called "swimlanes" was developed based on digital trace data from Talkwall interactions (such as ‘create’, ‘edit’, ‘pin’ and ‘move’). The ‘swimlane’ visualisation shows groups along an x-axis and their interactions over time along a y-axis, and may show how ideas appear that in one group are taken up and improved in another, or used in a spatial arrangement of ideas by a third group. Such visualisations can provide a powerful complement when combined with verbal transcripts of video recordings, allowing a multi-level coding of the dialogic character of the lesson.

Finally, the consequences of the current pandemic are now bringing education into unfamiliar territory, with hybrid and often incongruent oral and digital practices, we argue that joint negotiation of ‘talk rules’ may help students and teachers to connect diverse discourses (Staarman et al., 2003) across home, school, and diverse technologies.

Idea thread mapper: Expand and sustain collaborative knowledge building across communities

Jianwei Zhang and Mei-Hwa Chen

To prepare students for a new “white water” world featuring extraordinary challenges and rapid changes, research on learning communities needs to embrace more dynamic and transformative forms of collaboration through which students work creatively and continually on emergent challenges and move beyond pre-set frames and boundaries. Aligned with this need, researchers call for pedagogical and technological innovations to extend and expand collaborative learning over longer timescales and across multiple social levels (Stahl, 2013; Wise & Schwarz, 2017; Zhang, Yuan, & Bogouslavsky, 2020). The Idea Thread Mapper (ITM) project (see http://idea-thread.net) aims to address this need by creating technology support that serves to make collective progress in online discourse visible for student reflection and further accessible for cross-community sharing. With such support, students can better monitor emergent advances and directions in their own classroom’s discourse and further build idea connections between different classrooms that investigate related challenges. The design of ITM is guided by a dynamic system view of cross-community knowledge building featuring emergent, multi-level interaction and boundary crossing (Yuan & Zhang, 2019; Zhang et al., 2020). Accordingly, ITM uses a multi-layer framework to organize the collaborative spaces, which include the local discourse space of each classroom where students conduct collaborative discourse to advance their understandings of various problems and a cross-classroom meta-space where students can access and interact with the ideas from the “buddy classrooms” (Figure 2). Valuable ideas and problems developed in each community can be further shared in the cross-community space for higher-level discourse. At the same time, insights developed in the cross-community space are brought back to each community to develop further inquiry and integrated understandings. Learning analytics and visualizations are embedded in the local and the cross-community space to support students’ reflective monitoring of emerging inquiry directions, progress, and idea connections. These include theme- and timeline-based visualization of students’ online discourse as idea threads, reflective supports for co-authoring “super note” (Journey of Thinking) reflection to deliberate idea progress and problems in each thread of inquiry, a cross-classroom meta-space for sharing inquiry area maps and super notes (Journey of Thinking syntheses), and cross-classroom Super Talk to address emergent challenging issues of common interests.

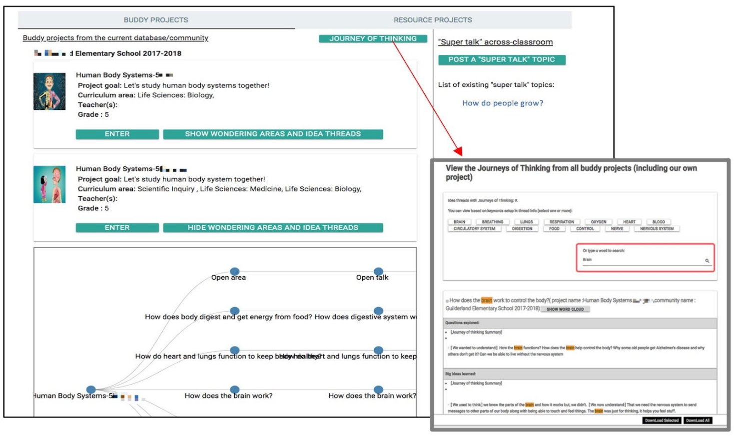


Figure 2. ITM’s meta-space for cross-classroom interaction

We conducted a design-based research on cross-community knowledge building in a network of Grades 5 classrooms over three school years. While members of each classroom worked together to investigate various problems and deepen their understandings in their own discourse space, they composed super notes to synthesize each productive line of inquiry, posted in a meta-space accessible to the partner classrooms. Some of the challenging problems emerging from the cross-classroom sharing resonated with the interests of different classrooms and became the focus of Super Talk across classrooms. Social network analysis revealed intensive connections formed among the students within each classroom, between different classrooms and student cohorts across school years. Mixed methods analyses of the multi-level discourse suggest that dialoguing with the ideas of different communities helped students to enrich and broaden their knowledge, engage in deeper reflection and inquiry, and further rise above distributed expertise to investigate complex issues (Yuan & Zhang, 2019; Zhang, et al., 2020; Yuan et al., 2021).

Co-designing a social learning Aanalytics dashboard to support community collaboration via social annotation

Jeremiah (Remi) Kalir

This study examines dialogic collaboration and knowledge building as jointly mediated by group use of social annotation (SA) and a social learning analytics dashboard. SA is a genre of learning technology that enables the annotation of digital resources for information sharing, social interaction, and knowledge construction (Kalir, 2020; Novak et al., 2012; Zhu et al., 2020). Research suggests SA enables interaction with texts as discursive contexts whereby "anchored discourse" (Gao, 2013) encourages "group-level" (Stahl, 2017) knowledge construction and meaning-making (e.g., Chan & Pow, 2020; Kalir & Garcia, 2019; Plevinski et al., 2017). To augment annotation-enabled discourse, a social learning analytics dashboard has been iteratively developed to help visualize group-level "activity traces" (Shum & Ferguson, 2012) and reveal to a given group their SA discourses, content, and activity contexts (Kalir, in press). Extending methods of participatory design research in the learning sciences (Bang & Vossoughi, 2016; Svihla & Reeve, 2016), this study asks: *How can group-level "meaningful participation" (Espinoza et al., 2020) in joint SA and social learning analytics activity foster collaborative learning?*

Since 2017, a team of researchers has developed a dashboard for Capturing and Reporting Open Web Data for Learning Analytics, Annotation, and Education Researchers (CROWDLAAERS; pronounced "crowd layers"). As a dashboard visualizing collaborative processes (e.g., Martinez-Maldonado et al., 2015), CROWDLAAERS displays social ("crowd") interaction and helps reveal how annotation ("layers") mediates collaboration. The dashboard reports social learning analytics associated with the open-source Hypothesis SA tool (Kalir, in press). We report on a study of the Right to Learn Undergraduate Research Collective (R2L), a university-based group studying case law to understand how concepts of dignity and equality are fundamental to the right of personhood. Working asynchronously and across continents, R2L has used Hypothesis SA and CROWDLAAERS to analyze a corpus of 52 legal documents (over 2,000 pages). As one participant noted, "Hypothesis was our tool of choice because of its capacity to function as a digital historian of our thinking." This study analyzes design artifacts, R2L's SA data, CROWDLAAERS activity traces, and audio recordings of group sessions. We document patterns of meaningful and multimodal group dialogue, conflict, and synthesis across social and technological contexts. Groups like R2L can leverage collaborative SA practices and open-access analytics resources like CROWDLAAERS to help visualize, make sense of, and (re)mediate knowledge construction. This case indicates promising approaches to the enactment of group-level constitutive acts and sociotechnical practices for meaningful participation in community learning.

Connect knowledge building and public discourse with the IdeaMagnets tool

Bodong Chen, Yu-Hui Chang, and David Groos

Our world is facing serious challenges such as an ongoing pandemic, climate change, political polarization, and misinformation. It will be "a great waste" for learners if school is largely disconnected from what is going on in the world. This study attempts to connect knowledge-building classrooms with public discourse by supporting the movement of knowledge artifacts across web spaces. This work builds on the Knowledge Building (KB) tradition that involves classes of students to work as knowledge communities to solve authentic problems through the continual improvement of their ideas (Chen & Hong, 2016; Scardamalia & Bereiter, 2014). To firmly connect the knowledge-building enterprises in schools and society, we posit that new technological infrastructures that embrace openness and bridge different web spaces are needed. To cohere learners' engagement in school and societal issues, we need to create entry points from the open web to sustained knowledge work, and vise versa.

Following the design-based research approach (Collins et al., 2004), we developed the IdeaMagnets tool based on an open-source web annotation technology named Hypothesis. The technological design of IdeaMagnets included two key components: (a) a collaborative web annotation system based on Hypothesis for students to annotate web documents; and (b) a Knowledge Forum add-on that queries Hypothesis annotations and enables learners to import annotations into Knowledge Forum for further discussion. Essentially, IdeaMagnets opens a gate for information from the open web to seamlessly enter Knowledge Forum while also projecting a knowledge-building mindset to learners' engagement with web resources.

To pilot the IdeaMagnets tool, we worked with five science classes (n = 95) in an urban public school in the United States. During this pilot, students studied a science unit on energy in connection with the "Green New Deal" that was trending in the US public discourse. Multiple sources of data were collected including student interviews, researcher fieldnotes, video recordings of classroom activities, and students' online discussions. Results showed that students approached the "Green New Deal" from different disciplinary angles (e.g., energy, economics, population, public health) and incorporated annotated information from public discourse to advance their Knowledge Forum discussions. Figure 3 presents a discussion thread made of 16 notes contributed by seven students. While this thread was launched based on a news article about rare-earth mining, students enriched their work by incorporating ideas captured in other web sources; five students created six notes containing web annotations imported via IdeaMagnets. The discourse pattern illustrated by the example was reflected in all five classes, showing the promise of supporting students to connect in-class knowledge building with public discourse.

Based on the study, we discuss the following opportunities of designing future technology for learning communities: (a) interfacing a community's knowledge processes with the open world, (b) allowing CSCL technologies to converse with general-purpose open-source tools while focusing on epistemic scaffolding, and (c) exploring open data exchanges among learning systems to enrich learning experiences and learning analytics.

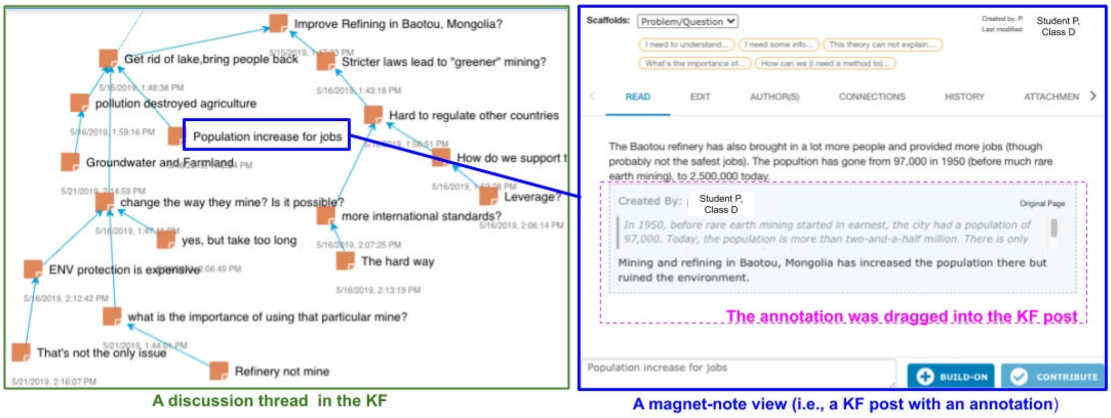


Figure 3. An example of using IdeaMagnets to connect public discourse with Knowledge Forum.

Scripting and orchestration of knowledge community and inquiry

Joel Wiebe and Jim Slotta

The pedagogical model called Knowledge Community and Inquiry (KCI) has been developed to guide our design of learning community curricula that engage students in collective forms of inquiry in which they, develop shared knowledge and practices while engaging in scripted inquiry activities (Slotta & Peters, 2008; Slotta, Quintana & Moher, 2018). KCI employs collaboration scripts (Kollar, Fischer & Hesse, 2006) to coordinate the orchestration of complex collaborative tasks, explicating roles, goals, groupings, sequences, and materials (Slotta, Tissenbaum & Lui, 2013). In KCI, students construct a collective knowledge base, indexed to a specified set of learning goals, that serves as a central resource for subsequent scripted inquiry activities (Slotta & Peters, 2010; Slotta & Najafi, 2013). Over the past decade, KCI has investigated forms of knowledge representation, classroom discourse, and inquiry designs (Moher et al., 2015; Fong & Slotta, 2018) often supported by bespoke technology environments, tailored for particular research contexts, including real-time analytics, intelligent agents, and student and group process modeling (Fong & Slotta, 2018; Slotta & Acosta, 2017).

We are currently developing a next-generation learning environment called SCORE (SCripting and ORchestration Environment) that adds a powerful layer for authoring and run-time orchestration capable of supporting a broad range of KCI inquiry scripts. The SCORE platform was launched in early 2020, coincident with the emergence of the COVID-19 pandemic. Our technology and pedagogy have positioned our research as a means to design for and promote peer-supported learning during these unique times of hybrid and online learning. This paper will begin with a summary of KCI and SCORE, followed by a review of early studies of undergraduate students in Wuhan, China and pre-services teachers in Munich, Germany. We then present a study of intermediate statistics in a graduate program in Toronto, Canada in which we re-envisioned the notion of *flipped classrooms* (Bishop & Verleger, 2013), that included pre-recorded lectures at-home activities, by re-imagining *how* they can be delivered as digestible, interactive activities in the context of cooperative and collaborative scripts. These at-home activities were then developed and orchestrated within SCORE. We describe how our collaboration scripts served to coordinate individual, collaborative, and collective activities that aim to improve peer-support for learning. We present student outcomes in terms of epistemological beliefs about the value of their peers and learning community pedagogy. We close with a discussion of the opportunities for further research in scripting and orchestration that may be afforded by a flexible authoring environment like SCORE. At present, we are developing new features for SCORE, including improved learning analytics, orchestration controls for teachers to monitor and guide activities, components for helping small groups negotiate and socially annotate community knowledge, and intelligent agents to support effective student groupings, alerts, and activity transitions. We also discuss possible intersections with other tools and environments presented within this session.

References

Acosta, A. & Slotta, J.D. (2018). CKBiology: An active learning curriculum design for secondary biology. Frontiers in Education, 3, 1-19. doi: 10.3389/feduc.2018.00052.

Bang, M., & Vossoughi, S. (2016). Participatory Design Research and Educational Justice: Studying Learning and Relations Within Social Change Making. Cognition and Instruction, 34(3), 173-193.

Bishop, J., & Verleger, M. A. (2013). The Flipped Classroom: A Survey of the Research. 23.1200.1-23.1200.18. https://peer.asee.org/the-flipped-classroom-a-survey-of-the-research

Bielaczyc, K., & Collins, A. (1999). Learning communities in classrooms: A reconceptualization of educational practice. Instructional-design Theories and Models: A New Paradigm of Instructional Theory, 2, 269-292.

Bielaczyc, K., & Collins, A. (2006). Fostering knowledge-creating communities. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), Collaborative Learning, Reasoning, and Technology (pp. 37–60). Mahwah, NJ: Lawrence Erlbaum Associates.

Chan, J. W., & Pow, J. W. (2020). The role of social annotation in facilitating collaborative inquiry-based learning. Computers & Education, 147, 103787.

Chen, B., & Hong, H.-Y. (2016). Schools as knowledge-building organizations: Thirty years of design research. Educational Psychologist, 51(2), 266–288. https://doi.org/10.1080/00461520.2016.1175306

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. Journal of the Learning Sciences, 13(1), 15–42. https://doi.org/10.1207/s15327809jls1301_2

Cress, U., Stahl, G., Ludvigsen, S., & Law, N. (2015). The core features of CSCL: Social situation, collaborative knowledge processes and their design. International Journal of Computer-Supported Collaborative Learning, 10(2), 109-116.

Espinoza, M. L., Vossoughi, S., Rose, M., & Poza, L. E. (2020). Matters of participation: Notes on the study of dignity and learning. Mind, Culture, and Activity, 1-23.

Fong, C., & Slotta, J. D. (2018). Supporting communities of learners in the elementary classroom: the common knowledge learning environment. Instructional Science, 46(4), 533-561.

Gao, F. (2013). A case study of using a social annotation tool to support collaboratively learning. The Internet and Higher Education, 17, 76–83.

Hod, Y., Bielaczyc, K., & Ben-Zvi, D. (2018). Revisiting learning communities: innovations in theory and practice. Instructional Science, 46(4), 489-506.

Järvelä, S., & Rosé, C. P. (2020). Advocating for group interaction in the age of COVID-19. International Journal of Computer-Supported Collaborative Learning, 15(2), 143–147. https://doi.org/10.1007/s11412-020-09324-4

Kalir, J. H., & Garcia, A. (2019). Civic writing on digital walls. Journal of Literacy Research, 51(4), 420-443.

Kalir, J. (in press). Designing a social learning analytics tool for open annotation and collaborative learning. In P. Prinsloo, S. Slade, & M. Khalil (Eds.), Learning analytics in open and distributed learning: Potentials and challenges. SpringerOpen.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts–a conceptual analysis. Educational Psychology Review, 18(2), 159-185.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. British educational research journal, 25(1), 95-111.

Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., & Clayphan, A. (2015). LATUX: An iterative workflow for designing, validating, and deploying learning analytics visualizations. Journal of Learning Analytics, 2(3), 9-39.

Moher, T., Slotta, J. D., Acosta, A., Cober, R., Dasgupta, C., Fong, C., ... & Peppler, K. (2015). Knowledge construction in the instrumented classroom: Supporting student investigations of their physical learning environment. International Society of the Learning Sciences, Inc.[ISLS].

Novak, E., Razzouk, R., & Johnson, T. E. (2012). The educational use of social annotation tools in higher education: A literature review. *The Internet and Higher Education*, *15*(1), 39-49.

Plevinski, J., Weible, J. and Deschryver, M. (2017), "Anchored annotations to support collaborative knowledge construction introduction", Making a Difference: Prioritizing Equity and Access in CSCL, 12th International Conference on Computer Supported Collaborative Learning (CSCL), No. 2013, pp. 111-118.

Rasmussen, I., Amundrud, A., & Ludvigsen, S. (2019). Establishing and maintaining joint attention in classroom dialogues, In The Routledge International Handbook of Research on Dialogic Education. Routledge. 28. s 410 – 423.

Roediger, H. L., & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. Journal of Applied Research in Memory and Cognition, 1(4), 242–248. https://doi.org/10.1016/j.jarmac.2012.09.002

Scardamalia, M., & Bereiter, C. (2014). Smart technology for self-organizing processes. *Smart Learning Environments*, 1(1), 1. https://doi.org/10.1186/s40561-014-0001-8

Shum, S. B., & Ferguson, R. (2012). Social learning analytics. Journal of Educational Technology & Society, 15(3), 3-26.

Slotta, J. D., & Acosta, A. (2017). Scripting and orchestrating learning communities: A role for learning analytics. In Proceedings of the 12th International Conference on Computer Supported Collaborative Learning (CSCL) 2017, Volume 1. Philadelphia, PA: International Society of the Learning Sciences.

Slotta, J. D., & Najafi, H. (2013). Supporting collaborative knowledge construction with web 2.0 technologies. In Emerging technologies for the classroom (pp. 93-112). Springer, New York, NY.

Slotta, J., & Peters, V. (2008). A blended model for knowledge communities: Embedding scaffolded inquiry. In Proceedings of the 8th International Conference on International Conference for the Learning Sciences-Volume 2 (pp. 343-350). International Society of the Learning Sciences.

Slotta, J., Quintana, R., & Moher, T. (2018). Collective Inquiry in Communities of Learners. In the International Handbook of the Learning Sciences. (F. Fischer, C. Hmelo-Silver, P. Reimann, & S. Goldman, Eds.). Routledge.

Slotta, J. D., Tissenbaum, M., & Lui, M. (2013). Orchestrating of complex inquiry: Three roles for learning analytics in a smart classroom infrastructure. In Proceedings of the Third International Conference on Learning Analytics and Knowledge (pp. 270-274). ACM.

Staarman, J. K., Aarnoutse, C., & Verhoeven, L. (2003). Connecting discourses: Intertextuality in a primary school CSCL practice. International Journal of Educational Research, 39(8), 807-816.

Stahl, G. (2013). Learning across levels. International Journal of Computer-Supported Collaborative Learning, 8(1), 1-12.

Stahl, G. (2015). A decade of CSCL. International Journal of Computer-Supported Collaborative Learning, 10(4), 337-344.

Stahl, G. (2017). Group practices: A new way of viewing CSCL. International Journal of Computer-Supported Collaborative Learning, 12(1), 113-126.

Svihla, V. & Reeve, R. (2016). Untold stories. In V. Svihla & R. Reeve (Eds.), Design as scholarship: Case studies from the learning sciences (pp. 1-10). New York, NY: Routledge.

Yuan, G., & Zhang, J. (2019). Connecting knowledge spaces: Enabling cross-community Knowledge Building through boundary objects. British Journal of Educational Technology. https://doi.org/10.1111/bjet.12804

Yuan, G., Zhang, J., & Chen, M.-H. (2021). Using Idea Thread Mapper to support cross-classroom "Super Talk" among four Grade 5 knowledge building communities. In *Proceedings of the ISLS Annual Meeting 2021*. Bochum, Germany: International Society of the Learning Sciences.

Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. International Journal of Computer-Supported Collaborative Learning, 12(4), 423-467.

Zhang, J., Yuan, G. & Bogouslavsky, M. (2020). Give student ideas a larger stage: support cross-community interaction for knowledge building. *International Journal of Computer-Supported Collaborative Learning, 15*, 389–410.

Zhang, J., Tao, D., Chen, M.-H., Sun, Y., Judson, D., & Naqvi, S. (2018). Co-organizing the collective journey of inquiry with Idea Thread Mapper. Journal of the Learning Sciences, 27, 390-430.

Posters

 © ISLS

International Society of the Learning Sciences

Empowering Secondary School Students' Argumentative Writing Skills: The Effectiveness of Dialogic Support and Cognitive Strategic Support On Students' Collaborative Writing Processes.

Yana Landrieu, Fien De Smedt, Hilde Van Keer, Bram De Wever
yana.landrieu@ugent.be, fien.desmedt@ugent.be, hilde.vankeer@ugent.be, bram.dewever@ugent.be
Tecolab. Department of Educational Studies. Ghent University.

Abstract: The proposal aims to outline an overview of three consecutive studies: a state-of-the-art study, a first and second-iteration intervention study. The aim of both intervention iterations is to enhance eleventh graders' argumentative, collaborative writing skills by combining two promising approaches, (1) explicit instruction regarding the writing process (i.e., cognitive strategic support) and (2) explicit instruction on collaboration (i.e., conversational support). To our knowledge, only one study (Granado-Peinado, Mateos, Martín & Cuevas, 2019) focusses on the effectiveness of combining both approaches.

Keywords: argumentative writing, secondary education

General overview

Effective writing skills are considered imperative in our 21st century society. Unfortunately, the majority of secondary school students do not develop proficient writing skills automatically. Furthermore, education fails to effectively promote these essential communication skills (NCES, 2012). The situation is even more problematic for argumentative writing: students' argumentative writing skills are often poorly developed and therefore they struggle to write qualitatively strong argumentative texts (Graham & Perin, 2007; NCES, 2012; Song & Ferretti, 2013; Philippakos & MacArthur, 2019).

Notwithstanding students' disappointing argumentative writing performance and their persistent writing difficulties, research indicates that these can be overcome by instruction, with strategic instructional and conversational practices as complementary approaches (Ferretti & Lewis, 2013). More particularly, strategy instruction, involving explicit and systematic teaching of the writing process, appears to have positive effects on the quality of students' writing (Graham, Gillespie, & McKeown, 2013). Second, argumentation is by nature something that only happens in interaction (e.g. with peers or with audience). Providing conversational support by creating opportunities in which students collaborate and assist one another while writing revealed beneficial effects on students' individual writing (Van Steendam, 2016). Research on supporting argumentative writing skills of secondary school students is, however, still scarce as most research focusses on higher education. Therefore, this research project aims to optimize the collaborative writing processes of eleventh grade school students. In Figure 1, the research project and the three different studies within the project are visualized.

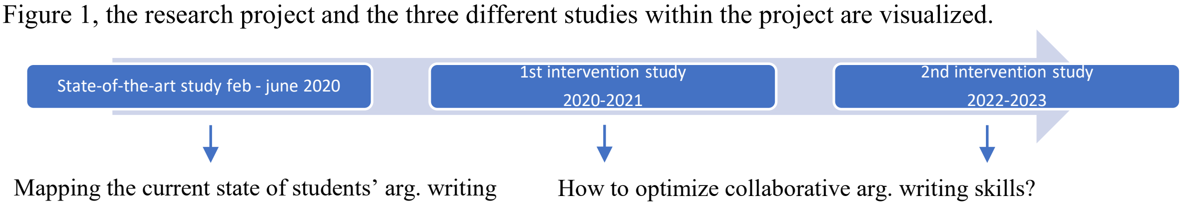


Figure 1. overview of the research project

State-of-the-art study

In 2020 a state-of-the-art study was set up to explore the current state of students' argumentative writing skills and different ways of assessing the text quality. More particularly, secondary school students (n = 159) were asked to individually write an argumentative text based on two digital source texts using a word processor. Originally, +/- 400 students would have written an argumentative text, but due to the Covid-19 pandemic (and the closing of schools in Belgium), this number was reduced to 159 students. In a subsample of participants (n = 50), screen recordings were captured using Screencast-O-Matic and keystroke logging data were collected using Inputlog to collect more in-depth information of the online writing processes of eleventh graders. To assess students' argumentative writing performance, multiple rating procedures were used: a) holistic rating (i.e., rating a text as a whole), b) analytic rating (i.e., rating a text on multiple aspects by assigning subscores) and c) pairwise comparisons (i.e., comparing two texts and deciding which one is the better text). Correlations between the three different rating procedures were calculated to investigate overlap between the rating procedures (Table 1). A distinction is made between argumentative texts (n = 129) and all texts (argumentative + informative, n = 159). Results reveal that not all students succeeded in writing an argumentative text: when no point of view (and therefore also no (counter) arguments) could be distinguished in a student's text, the text was considered as an informative text (instead of an argumentative text). In total 124 texts were categorized as argumentative texts while 35 texts were identified as informative texts. Most students do succeed in defending their point of view, but experience more difficulties taking into account the opposite point of view (i.e., 73 students did not provide any counterargument data). Half of the students did not succeed in refuting the opposite point of view (n = 80).

Table 1: Correlation coefficients between holistic (HOL), analytic (ANA) and pairwise comparisons (PC).

	Argumentative texts (n = 124)			Argumentative + Informative texts (n = 159)		
	HOL	ANA	PC	HOL	ANA	PC
HOL	-	0.587**	0.492**	-	0.291**	0.501**
ANA		-	0.315**		-	0.337**
PC			-			-

** significant on the 0.01-level

First-iteration intervention

Based on the initial state-of-the-art study, an authentic intervention study is being set up that tackles the question on how to foster these argumentative writing skills. This study will start in January 2021. In this study, students will go through the writing process collaboratively by interacting and supporting each other while planning and revising their texts. As a consequence, collaboration processes will be investigated and optimized.

The intervention study will combine explicit instruction of the writing process (i.e., cognitive strategic support) and explicit instruction on collaboration (i.e., conversational support). To our knowledge, only one study (Granado-Peinado, Mateos, Martín,& Cuevas, 2019) focusses on the effectiveness of combining both approaches. During the first-iteration intervention study, the impact of conversational support and cognitive strategic support of teachers on the collaboration of eleventh' graders will be examined. In addition, the impact of collaboration on the quality of argumentation and possibilities of optimizing those collaboration processes will be investigated.
The following student instruments will be administered both at pretest and posttest: a) individual argumentative writing test b) a student survey to measure students' writing processes, writing motivation, collaboration attitudes and self-efficacy, c) a reading comprehension test, d) log data (Inputlog) and e) screen capturing software.
We foresee to present the preliminary results at the poster session of ISLS.

Second-iteration intervention

After completion of the first-iteration intervention study, a follow-up study will be implemented in 2022-2023 in which the collaboration processes during argumentative writing will be refined and optimized.

References

Ferretti, R., & Lewis, W. (2013). Best practices in teaching argumentative writing. In S. Graham, C. A. MacArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (Second Edition ed., pp. 113-140). New York: The Guilford Press.

Graham, S., & Perin, D. (2007). What we know, what we still need to know: Teaching adolescents to write. *Scientific Studies of Reading*, *11*(4), 313-335.

Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: importance, development, and instruction. *Reading and Writing, 26*(1), 1-15. doi:10.1007/s11145-012-9395-2

Granado-Peinado, M., Mateos, M., Martín, E., & Cuevas, I. (2019). Teaching to write collaborative argumentative syntheses in higher education. *Reading and Writing*, *32*(8), 2037–2058. https://doi.org/10.1007/s11145-019-09939-6

National Center for Education Statistics. (2012). *The nation's report card: Writing 2011*. Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.

Song & Ferretti (2013). Teaching critical questions about argumentation through the revising process: Effects of strategy instruction on college students' argumentative essays. *Reading and Writing*, *26*(1), 67-90.

Philippakos, Z. A., & MacArthur, C. A. (2019). Integrating Collaborative Reasoning and Strategy Instruction to Improve Second Graders' Opinion Writing. *Reading & Writing Quarterly*, 1–17. https://doi.org/10.1080/10573569.2019.1650315

Van Steendam, E. (2016). Forms of collaboration in writing. *Journal of Writing Research, 8*(2), 183-204. doi:10.17239/jowr-2016.08.02.01

Introducing a New Approach for Investigating Learning Behavior

Mario Niederhauser, University of Applied Sciences and Arts Northwestern Switzerland (FHNW)
mario.niederhauser@gmail.com
Alessia Ruf, University of Basel, alessia.ruf@fhnw.ch
Joscha Jäger, filmicweb, joscha.jaeger@filmicweb.org
Carmen Zahn, University of Applied Sciences and Arts Northwestern Switzerland (FHNW)
carmen.zahn@fhnw.ch
Klaus Opwis, University of Basel, klaus.opwis@unibas.ch

Abstract: The potential of learners' video interactions to understand learning behavior has been recognized in previous research. However, little research has yet been conducted on enhanced video-based environments using behavior sequence analyses. Hence, we developed *Logible*, a sensitive, web-based tool to detect and analyze meaningful behavior sequences of learners interacting with such environments. The tool is based on an iterative method. With *Logible* we were able to visualize learning behavior and emphasize differences in experimental conditions.

Keywords: learning behavior, method development, sequence analysis, video-based learning

Introduction

Today, *enhanced video-based environments* are able to support conceptual understanding and deep processing as they provide – along with basic video control tools (e.g., play, pause, rewind) – new tools that allow to annotate, discuss, or edit videos alone or in groups (e.g., Zahn, 2017). Research on *learning analytics* emphasizes that much can be learned from learners' interactions with such enhanced videos about their learning behavior by analyzing log files that contain logged (inter-)action data (Mirriahi & Vigentini, 2017). However, while previous research has mostly focused on analyses of frequencies of single actions (e.g., by summarizing play clicks) (Mubarak et al., 2020), Sinha et al. (2014) argued that such analyses make it difficult to trace results back to actual learning behavior as hidden information from behavior patterns get lost. They thus recommended to encode meaningful sequences from log files by grouping single (inter-)actions. Yet, such behavior analyses are still rare in previous research on enhanced video-based environments – not least because of the huge effort that needs to be invested to detect and elaborate meaningful sequences from raw log files (Mubarak et al., 2020). We thus developed an interactive web-based tool (*Logible*) that is based on an iterative method and automatically visualizes and analyzes meaningful behavior sequences from raw log files of students learning either individually or in groups of two (i.e., learning setting) and using either hyperlinks or self-written annotations (i.e., learning task). In the present work, we ask: *can differences in learning behavior be made visible by developing a new method and tool using raw log files?* – and hypothesize that differences in learning settings (H1a) and learning tasks (H1b) can be found.

Methods

Logible (making *log* files leg*ible*) was developed based on an exploratory and iterative method using 92 data sets of totally 134 Swiss university students (75% female, *M* = 24.18 years, *SD* = 6.78) who learned about synaptic plasticity with the enhanced video-based environment *FrameTrail* (https://frametrail.org) either individually (N = 50) or collaboratively in dyads using one shared desktop computer (N = 84 individuals in 42 dyads). Participants could either add self-written annotations based on additional predefined informative texts (N = 65) or hyperlinks including these texts (N = 69) directly at appropriate places into the video and change their display time on the video timeline. *FrameTrail* automatically provided raw log files for each individual or collaborative data set containing learners' interactions in chronological array of occurrence. To find meaningful behavior sequences, two experts manually built groups of actions performed in conjunction with each other and defined rules for each of the detected sequences (e.g., mandatory actions of a sequence). The method was continuously improved and finally resulted in 17 behavior sequences (see Figure 1). At the same time, a first prototype of *Logible* was developed that was able to read the log files (see Figure 2) and considered the factors *priority*, *length* (of sequence), and *homogeneity* to detect the most meaningful sequences to be used for further analyses (see black framed sequence in Figure 2). The sequences were then assigned to five *behavior patterns* based on (1) their *intentional level* (i.e., did learners *search* or *find* appropriate places in the video to add an annotation or hyperlink) and (2) their *level of content creation* (i.e., did learners *add* or *modify* annotations or hyperlinks). We used behavior patterns for further analyses as (1) they show a higher contrast capability than sequences and because (2) not all sequences appear in the hyperlink condition (see Figure 1).

Behavior Patterns	Behavior Sequences: Annotation	Behavior Sequences: Hyperlink	Prioritization
1. Search and add	1.1 Search position and add annotation	1.1 Search position and add annotation	3
	1.2 Search position and add annotation and adjust time	1.2 Search position and add annotation and adjust time	2
	1.3 Search position and create annotation		2
	1.4 Search position and create annotation and adjust time		1
2. Search and modify	2.1 Search to adjust annotation time	2.1 Search to adjust annotation time	5
	2.2 Search to change/complement annotation text		5
	2.3 Search to change/complement annotation and adjust time		4
3. Find and add	3.1 Find position and add new annotation	3.1 Find position and add new annotation	8
	3.2 Find position and add annotation and adjust time	3.2 Find position and add annotation and adjust time	7
	3.3 Find position and create new annotation		7
	3.4 Find position and create annotation and adjust time		6
4. Find and modify	4.1 Find position and adjust annotation time	4.1 Find position and adjust annotation time	10
	4.2 Find position and add further video information		10
	4.3. Find position and add further video information and adjust time		9
5. Search and navigate	5.1 Rewatch	5.1 Rewatch	12
	5.2 Jump Forward	5.2 Jump Forward	12
	5.3 Skipping	5.3 Skipping	11

Figure 1. Overview of behavior sequences and behavior patterns with prioritization

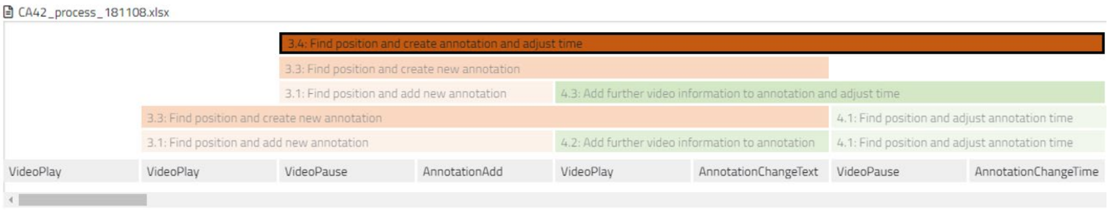


Figure 2. Example of detected sequences of a collaborative group (ID = 42) using annotations in *Logible*.

Findings

A MANOVA with *learning task* and *setting* as between-subject factors and relative values of the five behavior patterns as dependent variables revealed a significant main effect for *learning task* ($F(4, 85) = 4.650$, $p = .002$; Pillai's Trace = .180), as expected (H1b), indicating a difference in the frequencies of behavior patterns between the annotation and hyperlink condition (i.e., *2. Search and modify* ($p = .010$) and *4. Find and modify* ($p = .023$) occurred more often in the annotation condition and *3. Find and add* ($p = .035$) and *5. Search and navigate* ($p = 0.15$) occurred more often in the hyperlink condition). No main effect was found for learning setting ($p > .05$).

Conclusions and implications

This study aimed to develop a new method and tool (based on log files) for gaining insights into learners' behavior when interacting and learning with an enhanced video-based environment. Our approach to detect and visualize meaningful behavior sequences from raw log files was successful and resulted in interesting new insights. Therefore, we conclude that differences in learning behavior can be made visible by applying our method using the newly developed application *Logible*. Future research should increasingly consider sequence behavior analyses when investigating enhanced video-based learning. We encourage researchers to work with our tool.

References

Mirriahi, N., & Vigentini, L. (2017). Analytics of learner video use. In *Handbook of learning analytics* (Vol. 1, pp. 251–267).

Mubarak, A. A., Cao, H., & Ahmed, S. A. M. (2020). Predictive learning analytics using deep learning model in MOOCs' courses videos. *Education and Information Technologies*. https://doi.org/10.1007/s10639-020-10273-6

Sinha, T., Jermann, P., Li, N., & Dillenbourg, P. (2014). Your click decides your fate: Inferring Information Processing and Attrition Behavior from MOOC Video Clickstream Interactions. *Proceedings of the EMNLP'2014 Workshop*, 3–14.

Zahn, C. (2017). Digital design and learning: Cognitive-constructivist perspectives. In S. Schwan & U. Cress (Eds.), *The Psychology of Digital Learning: Constructing, Exchanging and Acquiring Knowledge with Digital Media.* (pp. 147–170). Springer International Publishing A.

Acknowledgements

This work was funded by the Swiss National Science Foundation (SNSF) under Grant 100014_176084.

Building Intercultural Competencies through Virtual Teams in Engineering Education

Anuli Ndubuisi, Jim Slotta, Elham Marzi
Anuli.ndubuisi@mail.utoronto.ca. Jim.slotta@utoronto.ca, Elham.Marzi@utoronto.ca
University of Toronto, Toronto, Canada.

Abstract: In an increasingly complex, interconnected, and globalized economy, engineers require intercultural competencies. This paper describes a unique international experience program composed of students and faculty from multiple universities and engineering firms, in which student teams collaborate on technical projects. We investigate students' understanding of intercultural competencies to ascertain the impact of our program on their perceptions of intercultural values and virtual teamwork. We discovered that integrating intercultural content with virtual international projects was a successful approach for helping these students build intercultural competencies and expand their engineering project knowledge and experiences.

Keywords: Collaborative Learning, Intercultural Competence, Global Virtual Teams, Engineering Education

Introduction: The growing need for intercultural competencies in engineering

Beyond technical skills and domain knowledge, future engineers require intercultural competencies and collaboration skills to succeed in an increasingly diverse and interconnected global economy (Downey et al, 2006). Globalization trends, technological advancements and the recent COVID-19 pandemic have accelerated the rise in global engineering projects with multicultural virtual teams. However, diverse global teams often encounter cultural differences, which can lead to poor social integration, mistrust, and conflict (Han and Beyerlein, 2016). Thus, we must prepare future engineers with intercultural competencies that can help them communicate in culturally respectful ways and make inclusive decisions. Our research takes advantage of a new global engineering program and investigates the development of intercultural competencies by participating engineering students from various nations who collaborate on engineering projects while working in virtual teams. We report early-stage findings from a study of the development of intercultural competencies within this program.

A Virtual International Experience Pilot Program

This study was situated at a large university in Canada, within a new program where virtual teams of engineering students are engaged in international cross-institutional partnerships to conduct technical projects. The study included 7 virtual project teams involving 20 engineering students from 8 universities and 1 industry partner. Each virtual team comprised 3 - 5 students and their supervisors. The students from over 10 countries were self-selected and represented diverse engineering fields.

Design framework

The design of our intercultural curriculum was guided by the Knowledge Community and Inquiry model (KCI) which provides a set of principles for learning community curriculum (Slotta, Quintana & Moher, 2018). Working in a KCI curriculum, students engaged in collaborative knowledge construction related to virtual team working and intercultural communication. Their knowledge artefacts (e.g., presentations, discussion notes or survey responses) are aggregated to form a community knowledge base that the students regularly referenced and, which reflects the community's growing "voice" and resources about intercultural understanding, and perspectives.

Learning activities and materials

The 4-week curriculum was delivered through a blend of social and collaborative project-based learning using Microsoft Teams, and Zoom. Students engaged in asynchronous and synchronous learning, which featured interactive lectures and guided collaborative small group discussions. The goal was to help students reflect on their prior knowledge and experiences, work in groups to co-construct new intercultural understandings and apply their learnings to improve intercultural communication and team effectiveness in their virtual teams.

Data collection and analytic approach

 © ISLS

The study employed a mixed method approach. Data included pre- and post-survey responses, asynchronous discussions, responses to inquiry items, and observation notes. The pre-survey collected information on students' background and cultural orientation. At the end of the learning program, students were surveyed about their overall impressions of the course and addressed specific open-ended questions about their experience in virtual teams. Content analysis examined patterns in students' experiences and their perceptions of intercultural learning and virtual team collaboration.

Findings and discussion

Pre-survey of students' experiences and cultural orientations

Students' responses showed that *no student recognized the importance of intercultural communication, and only one student identified intercultural awareness as a potential challenge for virtual teams.* Hence, before the program, students did not recognize the importance of intercultural competencies to virtual teams.

Post-survey and discussion of students' experience (Likert and open-ended items)

Students were highly satisfied with the learning sessions, with 90% rating the intercultural communication and virtual team sessions as "excellent" or "very good" and saying they would recommend the sessions to others. Student feedback revealed four distinct themes, described in the following sections: (1) intercultural awareness and appreciation, (2) diversity, (3) intercultural communication, and (4) trust and commitment.

Intercultural Awareness and Appreciation. Students identified cultural differences (e.g. *"weekends aren't the same across the world")* and their potential impact on virtual teams (e.g. *"different dedicated religious days and workdays that result in compromise"*). Several approaches for improving intercultural awareness were articulated, such as *"Being aware that there are cultural differences and being accepting and understanding of these differences"*, and *"Compromising, so that work is still completed while being respectful to your teammates"*.

Diversity. Students indicated an appreciation of the similarities (e.g., *"shared university culture, peer-aged"*), the diversity of team members, (e.g., *"There're some team members that are quiet and some that are the complete opposite which makes the team balanced"*) and the benefits of working in multicultural engineering virtual teams: *"We studied the same concepts, which makes us think in the same way. However, every member provides a different idea based on the many different backgrounds we have"*.

Intercultural Communication. Students recognized and appreciated that communication across cultures and geography was essential for virtual teams. Students used technology to promote agency (e.g., *"Ensuring that everyone has a chance to voice their opinion"*), seek clarifications (e.g., *"Asking a lot of questions to understand the tasks better and seek clarification to reduce the number of miscommunications."*), resolve issues (e.g. *"The team has produced very efficient meetings as everyone are willing to help and discuss the problem together"*).

Trust and Commitment. To create a sense of community, teams made efforts to build a sense of belonging (e.g., "*We are a diverse team, spanning different continents. We are proud of our chemistry as a team."*). This served to build trust and encourage deeper social interactions. (e.g., *"Each of us is active on the group chats, answering any questions our teammates may have as soon as possible"*), and commitment (e.g., *"When someone has a busier schedule, the team helps by taking on a larger role for that time period, which is usually worked out later"*).

Conclusions

This program has demonstrated a successful strategy to help students develop intercultural competencies. The international teams provided a basis for our interventions around intercultural sensitivities and virtual teamwork. This was a relatively small-scale study, conducted against the backdrop of the COVID-19 pandemic, but nonetheless demonstrates proof of concept and revealed four themes that were salient to student participants.

References

Downey, G., Lucena, J., Moskal, B., Parkhurst, R., Bigley, T., Hays, C., Jesiek, B., Kelly, L., Miller, J., Ruff, S., Lehr, J. and Nichols-Belo, A. (2006). The Globally Competent Engineer: Working Effectively with People Who Define Problems Differently. Journal of Engineering Education, 95(2), 107-122.

Han, S., & Beyerlein, M. (2016). Framing the effects of multinational cultural diversity on virtual team processes. *Small group research*, *47*(4), 351-383.

Slotta, J. D., Quintana, R. M., & Moher, T. (2018). Collective inquiry in communities of learners. In *International Handbook of the Learning Sciences,* 308-317, Routledge.

Exploring Turtle Blocks in an online collaborative environment

Ravi Sinha, Gnowledge Lab, Homi Bhabha Centre for Science Education, Tata Institute of Fundamental Research, Mumbai, India, ravis@hbcse.tifr.res.in
Amit Dhakulkar, Research Unit Self-Directed Learning, Faculty of Education, North-West University, Potchefstroom, South Africa, damitr@gmail.com

Abstract: In this study, we explore the design and implementation of an online engagement on Turtle Blocks (microworld) and how it supports the introduction of core programming concepts in alignment with the constructionist approach. The engagement provided opportunities for discussions and collaborations amongst peers via synchronous online session and asynchronous instant messenger (IM). The preliminary analysis helps in understanding the nature of collaboration, tinkering by the learners and varied styles used for problem-solving.

Introduction

Visual programming languages such as Turtle Blocks (https://turtle.sugarlabs.org/) have several advantages for novice learners over text-based languages. Turtle Blocks is an implementation of constructionism and can be categorised as a *microworld*. The microworlds have several affordances for novice learners which include, having a very low threshold, “objects to think with” and tinkerability and as such allow different ways of thinking and knowing (Resnick and Rosenbaum, 2013, Noss & Hoyles, 2017). In this study, we report the design and implementation of an engagement aimed at introducing novice learners to a microworld in the form of a visual programming language Turtle Blocks. The objective of the study was to explore how the design of an online collaborative engagement influenced the teaching-learning processes in the context of a microworld. The main research question had two parts: first to understand the learning processes of the participants, and second, was to identify which features of our designed engagement were beneficial in the constructionist framework of learning. The insights gained from the analysis so far provide us with some broad findings and ideas for the next iteration of the engagement. We discuss some episodes from the engagement in the context of problem-solving, computational thinking and collaborative interactions.

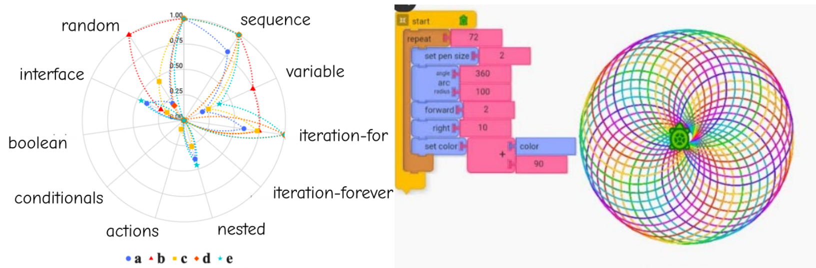
Designing microworld engagement and context

The six synchronous online sessions, increasing in complexity of one hour each, spaced across three weeks introduced some of the foundational programming concepts in the form of various *code blocks* and their combinations in Turtle Blocks using the context of projects. The sessions were anchored by one researcher with other mentors (four) coming in for particular sessions and facilitating discussions on the IM platform. The code blocks covered in these sessions were *movement blocks* (forward, left, back, right), *colour blocks* (colour, shade), *loops and conditionals* (repeat, forever, if, if-else), *variables* (box), *logical* (OR, AND), *numbers* (number, random), *functions* (action). These blocks demonstrated some computational concepts such as sequences, iterations, loops (simple, nested), data (stepping variables and their manipulation), events, conditionals, operators, and parallelism (Brennan and Resnik, 2012). The challenges were voluntary and cohort-driven, and learners were encouraged to present their own design challenges to their peers. The IM group was used to share artefacts and for communications and discussions between the learners and the mentors. A curated wiki page containing exemplar projects from the Turtle Art community was created for the participants (https://metastudio.org/t/turtleart-challenges-wiki/4036). Several unique variations, indicating tinkering and different ways of problem-solving were seen in the submitted projects.

An open invitation for the engagement was sent out to potential participants. Five participants (out of a total of nine) submitted the project files and reflection sheets, while others participated partially on the IM platform and attended few live sessions. The distribution of learners was as follows: Grade 4-5 (1), Grade 6-8 (2), Grade 9-12 (1), Adult (1) and were from urban and semi-urban Indian settings. The data was collected primarily from the following sources: (i) *Reflection Sheet,* a reflective questionnaire completed by the participants, (ii) the project files, (iii) discussions on the IM platform, (iv) researcher notes and observations. A mixed-methods approach was used for the analysis. The reflective entries were coded according to the following themes: episodes of discovery, challenge, overall experience, problem-solving. This analysis is not yet complete.

Findings

The project files were investigated for the presence of the categories of coding blocks that maps to the computational concepts described earlier. Figure 1 (a) shows the preference for certain types of simple blocks, while some blocks were not used at all. Also, a few projects utilised new blocks which were not discussed in the sessions. Most of the projects just involved the usage of sequencing, iteration and randomization. One of the recurring patterns in the reflection sheet of the participants indicated the fondness for the *random block* in combination with *set-color* which produces non-repeatable, non-reproducible color patterns (Figure 1 (b)). The participants found the concept of randomness quite intriguing, as it was a novel concept for them which was echoed in the sentiments as expressed in their Reflection Sheet.



(a) (b)

Figure 1. (a) Relative usage of blocks, colours indicate different learners (b) Multi-color wheel with *random*

The open-ended nature of the challenges encouraged creative thinking as there was no single right answer. One of the powerful features of computationally powered graphics is the versatility of the designs that are possible during the learning process. The tinkering at times led to serendipitous discoveries by the learners as evidenced by entries in reflective journals. Even with the use of limited blocks for programming, variations could be found in approaches to problem-solving and styles in different learners.

Discussion and reflections

The study highlights the importance of gradual scaffolding of the computational concepts along with opportunities for non-linear explorations and proposes that these should be elicited within the context of exploring personally meaningful projects challenges. Our preliminary observations addressing the first part of the research question indicate that certain programming concepts are not very intuitive for every learner and require more facilitation on part of the designer. The artefact analysis of the submitted projects indicates that most participants found it easy to apply simple blocks in combination with repeat and random to create various designs. Using a microworld like Turtle Blocks with its design features it is possible to support novice learners irrespective of their age in online settings with minimal didactical teaching. Some aspects of the engagement design that worked well in this iteration included presenting only the challenges and the idea of introducing all the exemplar projects in the beginning. This approach encouraged the participants to attempt challenges and experiment with designs which matched their readiness levels and interest. We can perhaps say that the different levels of learners in the cohort and their differing familiarity with Turtle Blocks created a *zone of proximal development* that supported collaborative learning and the designs themselves became *objects to think with* and hence having a diverse group and open-ended challenges proved to be useful. Further analysis of the existing data can lead to more insights. How some of the infrastructural challenges identified during the study influence the learning process is an open question that needs further study. Being such a small sample and limited time frame of three weeks, substantial claims cannot be made about the teaching-learning process and remains an obvious limitation of the study.

References

Brennan, K., & Resnick, M. (2012). *New frameworks for studying and assessing the development of computational thinking*. 25.

Noss, R., & Hoyles, C. (2017). Constructionism and Microworlds. In E. Duval, M. Sharples, & R. Sutherland (Eds.), *Technology Enhanced Learning* (pp. 29–35). Springer International Publishing.

Resnick, M., Rosenbaum, E., (2013). Designing for tinkerability, in: Honey, M., Kanter, David.E. (Eds.), *Design, Make, Play: Growing the next Generation of STEM Innovators*. pp. 163–181.

Acknowledgments

The authors would like to thank Prof Nagarjuna G and Prof Jako Olivier for their comments and feedback. The authors would also like to thank all the participants including mentors of Gnowledge Lab, HBCSE for their engagement which made this work possible.

ArguNotes: Collaborative Problem Solving and Argumentation Tool

Janet Rafner, Arthur Hjorth, Carrie Weidner, Shaeema Zaman Ahmed, Christian Poulsen, Clemens Klokmose, Jacob Sherson
janetrafner@mgmt.au.dk, arthurhjorth@phys.au.dk, cweidner@phys.au.dk,
shaeema@phys.au.dk, cbp@mgmt.au.dk, clemens@cavi.au.dk, sherson@mgmt.au.dk
Aarhus University, Aarhus, Denmark

Abstract: As educators push for students to learn science by *doing* science, there is a need for computational scaffolding to assist students' evaluation of scientific evidence and argument building. In this paper, we present a pilot study of ArguNotes, a CSCL tool allowing educators to integrate third-party software into a flexible and collaborative workspace. We explore how ArguNotes enables teams to build data-driven arguments in inquiry-based learning activities.

Introduction and theoretical considerations

The last three decades have seen an ever-increasing call for educational activities in which learners *do* science instead of *learning about* science. A growing number of initiatives support the general public's participation in research, but there is an urgent need for tools to facilitate these processes in both formal and informal learning contexts. When moving *beyond mere data-gathering, scientific argumentation* is key to actively engaging students in epistemological and methodological considerations. The CSCL community has sought to support this inquiry-oriented turn by providing computational scaffolding (Quintana et al, 2004) to learners engaged in argumentation in a variety of ways: AL (Wambsganss et al, 2020) gives learners real-time feedback on written argumentation; Critical Thinking (Sun, Yuan, Rosson, Wu, & Carroll, 2017) supports dyads discussing a particular proposition; and ThoughtSwap (Dickey-Kurdziolek, Schaefer, Tatar, & Renga, 2010) assists students in discussing the fitness of an answer to a question. In this pilot study, we introduce ArguNotes, a CSCL tool that both supports collaborative problem solving by integration of scientific data generated from third-party tools and assists learners in evaluating scientific evidence and building scientific arguments.

Introducing ArguNotes and pilot study

ArguNotes is an online tool, built on Webstrates (Klokmose, Eagan, Baader, Mackay, & Beaudouin-Lafon, 2015**)**. The goal of ArguNotes is to scaffold students' inquiry-based learning experience through two modes: *mode 1:* problem solving, sharing and building upon other's solutions to collectively explore the solution space and arrive at the best possible solutions (Fig. 1a, above dashed line) and *mode 2:* collective argumentation in which insights from mode 1 can be refined and aggregated into a deep phenomenological investigation of the underlying topical challenge (Fig. 1a, below dashed line). To do so, ArguNotes contains three specific design features: 1) data integration with third party software, 2) 'mini-papers' that consist of data (numerical results or arguments) and 3) 'citations', a means for students to refer to each other's mini-papers as they collectively build arguments. In Figure 1a we show a screenshot of ArguNote. All of a team's mini-papers are visualized as small rectangles showing the title of the evidence and a link to all mini-papers that it cites (or that cite it). Clicking a mini-paper opens it to full screen so students can see its data and argument. The students only directly input data when they are creating a mini-paper. Students cannot edit the content of other groups' mini-papers; they can only cite them. When a paper is cited, it appears in a reference section at bottom of the mini-paper. In this pilot study, we focus on investigating whether such a tool can successfully form the backbone of an inquiry-based learning activity (see Fig. 1b, inset).

Methods and Context

The study was done with 17 students in a graduate-level quantum mechanics course at a Danish university. We integrated the ArguNotes backend with Quantum Composer (QC), a quantum physics education and research tool (Ahmed et al, 2020). Students used QC to explore a difficult physics problem with an enormous number of possible solutions; any solution can be evaluated in terms of its proximity to an ideal solution by a score ranging from 0 (a poor solution) to 1 (a perfect solution). Onboarding took place via a 30-minute in-person presentation introducing students to ArguNotes and the challenge. Over the next hour, the students worked on the challenge in four teams. Students could talk to each other during the activity, but sharing of the details of a given solution (encoded as a JSON file) was only possible using the real-time-updating ArguNotes interface. To encourage reflection on the process of scientific argumentation, each team was explicitly evaluated on the complexity of their final joint argumentation and the extent to which they had used citations to acknowledge and build on each other's contributions.

 © ISLS

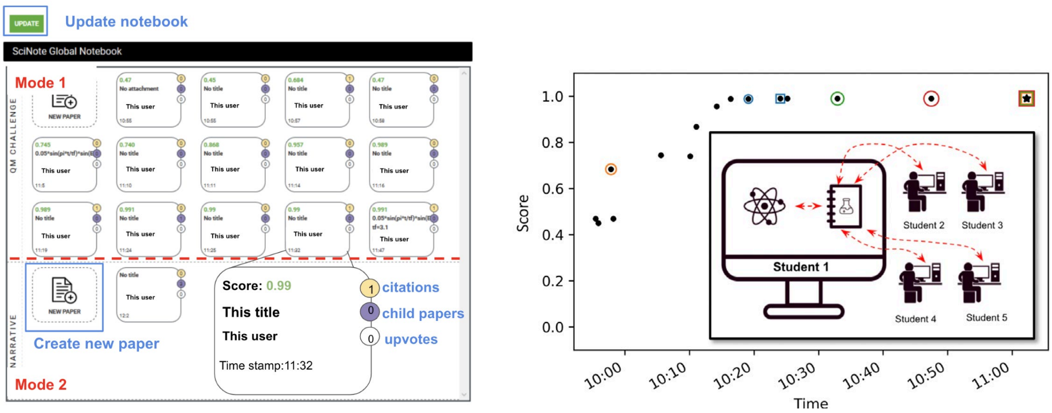


Figure 1. (a, left) *ArguNotes interface*: screenshot from Team 1 during the study, including an inset to explain each feature (b, right) *Iteratively improving on each other's results:* Team 1 score vs. time, with rings (boxes) around cited (citing) papers, culminating in a final analysis paper (star). (inset) Collaborative quantum problem solving using the ArguNotes interface.

All teams' scores increased roughly monotonically over time, indicating that teams built upon their previous solutions via citations and intra-team oral communication. For example, Team 1 submitted four papers at the beginning of the exercise (Fig. 1b), and all further work built on the best of these (score = 0.72), as the score never dropped below this level for the rest of the activity; other teams' results were similar. Finally, teams with more citation activity achieved higher scores and demonstrated higher levels of physical understanding in their analyses. The number of citations (final score) for each team were (4, 0.991), (1, 0.72), (9, 0.88), (1, 0.785). Therefore, while this is a pilot study with a small number of participants, these are indications that increased engagement with ArguNotes led to better knowledge exchange and ultimately better results.

Conclusion and outlook

Here, we demonstrate that ArguNotes can facilitate collective, complex problem-solving activities with third party tools, previously only possible with tedious manual file sharing. Furthermore, we see indications that ArguNotes can help bridge the gap between such activities and collective scientific argumentation. One student notes:

> It seems to work great with these 'investigative' types of problems/challenges but I have rarely seen such investigative problems during my normal studies. Perhaps it could be an idea to discuss with other teachers/professors at the university, if they want to make such investigative problems and create some (fun) challenges.

Future work will explore how ArguNotes can be used in other scientific contexts, including the social sciences, in other educational contexts, in large-scale studies, and in data integration using different third-party interfaces.

References

Ahmed, S.Z. et al. (2020). Quantum Composer: A programmable quantum visualization and simulation tool for education and research. *arXiv:2006.07263*.

Dickey-Kurdziolek, M., Schaefer, M., Tatar, D., & Renga, I. P. (2010). Lessons from ThoughtSwap-ing: Increasing participants' coordinative agency in facilitated discussions. *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, 81–90.

Klokmose, C. N., Eagan, J. R., Baader, S., Mackay, W., & Beaudouin-Lafon, M. (2015, November). Webstrates: shareable dynamic media. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology,* 280-290.

Quintana, C. et al. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, *13*(3), 337–386.

Sun, N., Yuan, C. W., Rosson, M. B., Wu, Y., & Carroll, J. M. (2017). Critical Thinking in Collaboration: Talk Less, Perceive More. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2944–2950.

Wambsganss, T. et al. (2020). AL: An Adaptive Learning Support System for Argumentation Skills. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14.

The Value of Using Roles while Collaboratively Writing Synthesis Texts in University

Karen Putzeys, Bram De Wever
karen.putzeys@ugent.be, bram.dewever@ugent.be
Tecolab, Department of Educational Studies, Ghent University

Abstract: Writing synthesis texts fosters students' writing skills as well as their conceptual learning. Writing collaboratively can improve synthesis text quality. This poster presentation focuses on an intervention study aimed at improving the writing quality in collaborating groups by means of guidance through roles. This proposal focuses on the background and design of the study, and we will present preliminary findings at the conference.

Keywords: collaborative learning, roles, synthesis writing, higher education

Theoretical background

In the current knowledge-based society, university students are required to initiate and direct their own learning. They should be able to produce knowledge based on reliable sources, and externalize that knowledge in, for example, a synthesis text. Synthesis writing is characterized by recurring and alternating reading and writing (Vandermeulen et al., 2020). It is an effective form of writing to learn, and subsequently an often provided task at university. Therefore, students should be taught synthesis strategies, through which their writing skills as well as their conceptual learning can be improved. This way, learning to write and writing to learn are connected in synthesis tasks (van Ockenburg et al., 2019).

When writing, students need to keep the audience in mind to achieve the text's communicative goal. In this study, the goal of the text is to inform the reading audience. Students with less experience in writing tend to overlook the impact of their writing on the reader, leading to difficult to understand texts (Bereiter & Scardamalia, 1987). Collaborative writing has the potential to overcome this problem, since it increases students' audience awareness (Storch, 2018).

However, writing a synthesis text collaboratively is not an easy task, since it involves two components. First of all, students need to learn how to write a synthesis text. Second, they need to know how to collaborate and how to build on each other. In view of the first component, students need strategy instruction: they need to learn how to select, organize and connect the information (van Ockenburg et al., 2019). Furthermore, the writing task can be pre-structured by means of a list of key elements a particular synthesis text should contain (Weinberger et al., 2005). In view of the second component, students need support to collaborate efficiently, which can be provided through an explicit role structure (Wang et al., 2017). Roles are prescribed functions that guide individual behavior and group collaboration. It is a way to distribute tasks and responsibilities in collaborative groups, supporting interdependence and at the same time individual accountability (De Wever & Strijbos, in press; Slavin, 1995). Students need to be additionally scaffolded regarding how to carry out the role (Rummel & Spada, 2005), for example by means of role descriptions. The role descriptions explain the function of the role, and can provide some examples of sentence openers or question stems guiding students in what to possibly say within a specific role (Weinberger et al., 2005).

Problem statement

As illustrated above, synthesis writing is an important task at university. However, students with little experience in writing may not be sufficiently aware of the audience, resulting in difficult to understand texts. Writing collaboratively seems to be a solution, if students receive strategy instruction regarding how to write a synthesis text. In addition, supporting students by assigning them roles throughout the collaborative writing process may reinforce the process and subsequently the quality of the synthesis text. A reason for this is that roles can foster more balanced participation and this way indirectly impact the quality of the writing product (Olson et al., 2017). Examples include motivating others and asking for contributions (Wise et al., 2012). Moreover, roles can hold responsibilities directly related to the text which may impact the quality of the writing product directly. Examples are checking for the presence of key elements in the text or improving clarity. This leads to the following research question: What is the added value of roles in addition to strategy instruction on the quality of university students' synthesis texts?

Methods

A randomized quasi-experimental design was set up with one experimental condition and one control condition. During the intervention, 41 groups of three university students in the third year of their bachelor's degree taking a course on academic writing wrote two synthesis texts. The first was an abstract of a provided research article, the second was a conclusion of another provided research article. These synthesis texts are comparable in terms of type, communicative goal and audience. For the second text, groups of students were randomly assigned to either a control condition (receiving only strategy instruction), or an experimental condition (receiving strategy instruction and roles guiding them through the collaboration process).

The strategy instruction comprised instruction on selecting, organizing and connecting information, in line with the design principles for synthesis writing interventions by van Ockenburg et al. (2019) and a list comprising key elements of an abstract and a conclusion respectively. Students in the experimental condition were in addition asked to distribute the following roles in their group as they saw fit. The *initiator* gives direction in terms of task division, planning of the synthesis text and time-management. The *moderator* ensures consensus, asks everyone's opinion, compromises and summarizes throughout the process. The *proof reader* reads and revises the text from the target audience's perspective, and checks if all key elements are present in the text. These roles are partially based on research by Wise et al. (2012). The students received a description of the functions and examples of sentence openers as additional scaffolds for the three roles.

Text quality will be measured by means of pairwise comparisons based on the used design principles, i.e. presence of key elements (selecting), logical order (organization), and degree of integration (connection). Differences in the quality of the synthesis texts between the two conditions will be analyzed. Detailed collaboration processes of a total of eight groups in the experimental and control condition will be captured for further investigation.

Points of discussion

Key findings will be presented during the conference. Next to the research question, i.e. the added value of roles on synthesis text quality, points of discussion will comprise the selection of the roles, the measurement of synthesis text quality, the limitations of the current study and the next steps following this study.

References

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Routledge. https://doi.org/10.4324/9780203812310

De Wever, B., & Strijbos, J. W. (n.d.). Roles for structuring groups for collaboration. In U. Cress, C. Rosé, A. Wise, & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning*. Springer.

Olson, J. S., Wang, D., Olson, G. M., & Zhang, J. (2017). How people write together now: Beginning the investigation with advanced undergraduates in a project course. *ACM Transactions on Computer-Human Interaction*, *24*(1). https://doi.org/10.1145/3038919

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. https://doi.org/10.1207/s15327809jls1402

Slavin, R. E. (1995). *Cooperative learning: Theory, research and practice* (2nd ed.). Allyn & Bacon.

Storch, N. (2018). Collaborative writing. In *The TESOL Encyclopedia of English Language Teaching*. https://doi.org/10.1002/9781118784235.eelt0395

van Ockenburg, L., van Weijen, D., & Rijlaarsdam, G. (2019). Learning to write synthesis texts: A review of intervention studies. *Journal of Writing Research*, *10*(3), 402–4028. https://doi.org/https://doi.org/10.17239/jowr-2019.10.03.01

Vandermeulen, N., Van Den Broek, B., & Van Steendam, E. (2020). In search of an effective source use pattern for writing argumentative and informative synthesis texts. *Reading and Writing*, *33*(2), 239–266. https://doi.org/10.1007/s11145-019-09958-3

Wang, D., Tan, H., & Lu, T. (2017). Why users do not want to write together when they are writing together. *Proceedings of the ACM on Human-Computer Interaction*, *1*, 1–18. https://doi.org/10.1145/3134742

Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science*, *33*(1), 1–30. https://doi.org/10.1007/s11251-004-2322-4

Wise, A. F., Saghafian, M., & Padmanabhan, P. (2012). 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. https://doi.org/10.1007/s11423-011-9212-7

Identifying Productive Conflict during Upper Elementary Students' Collaborative Programming

Zarifa Zakaria, Jennifer Tsan, Jessica Vandenberg, Collin F. Lynch
zzakari@ncsu.edu, jtsan@ncsu.edu, jvanden2@ncsu.edu, cflynch@ncsu.edu
North Carolina State University
Kristy Elizabeth Boyer, University of Florida, keboyer@ufl.edu
Eric N. Wiebe, North Carolina State University, wiebe@ncsu.edu

Abstract: Although pair programming has been found to be an effective tool for classroom instruction, conflict is a common occurrence during such collaborative environments. It is important to distinguish between conflicts that may engender an unfavorable learning environment from one that can be productive. In this poster, we report a study exploring 78 conflict episodes that occurred during pair programming by 9-11 years old students. Findings showed Exploratory talk being the most productive talk, occurring at a significantly higher level during conflict than pre- or post-conflict. Conversation leading up to conflict shows significantly higher levels of suggestions.

Introduction

Collaboration encompasses cognitive processes such as knowledge sharing and critical thinking, as well as emotional processes that may involve conflicts (Isohätälä et al., 2018). Conflict, being one of the challenges of collaboration, occurs when partners have difficulties in understanding each other's thinking, or negotiating multiple perspectives (Kirschner et al. 2008). Research on conflicts in computer-supported collaborative learning (CSCL) typically involves older students (e.g., Basil-Shachar et al., 2015); contexts outside of computer science (CS) (e.g., Wise et al., 2015); or mainly focuses on argumentation as a teaching approach (e.g., Noroozi, 2020). It is important to directly study conflicts to distinguish the characteristics of productive versus unproductive conflicts with upper elementary students.

In order to distinguish between productive and less productive conversation, we integrated a theoretical framework defined by Mercer and Littleton (2007) which describes students' dialogue in ways that suggest these distinctions. Mercer defines three types of talk that can occur in the course of a conversation: Cumulative, Disputational and Exploratory. In Cumulative talk, speakers build positively but uncritically on what their interlocutor has said. Disputational talk, by contrast, is characterized by disagreement and individualized decision-making. Finally, in Exploratory talk, participants engage critically but constructively with each other's ideas, which leads to improved reasoning and conceptual understanding (Mercer and Littleton, 2007). Guided by this theoretical framework, we studied the following research questions: 1) What are the characteristics of conversations during a conflict? 2) How does the characteristics of conversations differ between pre, during and post conflict?

Context and method

A total of 14 5th grade students (8-11 years; 5 females, 9 males) forming 10 dyads participated in pair programming using a block-based programming environment, NetsBlox. We used a mixed-method approach (Creswell & Clark, 2017), where we coded and transformed qualitative data into quantifiable form to conduct quantitative Webcam video and screen capture data were collected.

For our analysis, first, we identified episodes of conflicts by coding the start and end of conflicts during each dyadic collaboration (Tsan etal., 2021). Then, we applied discourse analysis to the students' conversation in conflict episodes that consisted of multiple turns of talk. Two authors coded the videos using a coding scheme developed across multiple studies (Zakaria et al., 2021). The coding scheme had categories that reflect Exploratory talk: *Alternative ideas*, *higher-order questions* and *justification.* For Cumulative talk: *simple questions*, *suggestions, comments, self-explanations, coordination, agreement.* We also had categories of *disagreement* and *others*. D*isagreements* were further re-coded into two categories. *Justified disagreements* (Exploratory talk) and *Unjustified disagreements* (considered Disputational talk). To explore the conversation occurring immediately before and after conflict episodes, we coded 6 turns of talk before and after conflict, called pre-conflict and post-conflict, respectively.

Results and discussion

We examined a total of 78 conflicts (*mean*=5.9, *mode*=3, *min*=2, *max*=17 per dyad) across 10 dyads collaborating for a total of 504.95 minutes (*mean*=50.5 mins per video). For the first research question, we calculated the percentage of times the categories were used. During conflicts, conversations were 5.4% *justified disagreements,* 11.6% *justification,* 2.9% *alternative idea,* and 2.4% *higher-order questions;* 18.6% *self-explanation*, 16.6%

suggestion, 4.6% *comment,* 3.7% *agreement;*10.3% *Unjustified disagreements*; 14.5% *other* and 3.1% *coordination*.

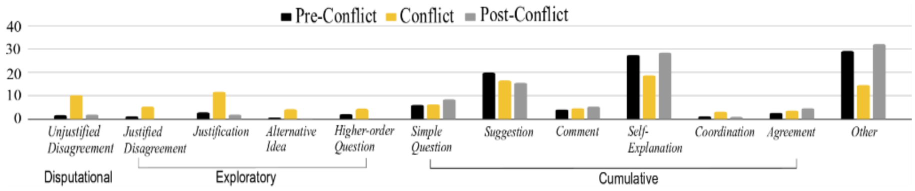


Figure 1. Average percentage of conversation categories used in episodes of pre, during, and post-conflict.

For our second research question, we examined the relative proportion in the form of percentage of each category used during episodes of conflict (*n*=78), pre-conflict (*n*=78), and post-conflict (*n*=63) using ANOVA tests. Results show all categories of Disputational (*Unjustified disagreement*) and Exploratory talk (*Justified disagreement, justification, alternative idea, higher-order thinking*) and only two Cumulative categories (*self-explanation* and *suggestion*) being significantly different between pre, during, and post-conflict. Pairwise comparison of the ANOVA showed that the mean percentage of Disputational and all Exploratory talk categories were higher during conflict than pre- and post-conflict episodes (Figure 1). Additionally, pre-conflict had higher *suggestions.*

Our findings show all the Exploratory categories are significantly higher during conflict episodes. This seems to indicate that conflict and Exploratory talk are not mutually exclusive. In addition, conversation that immediately precedes a conflict involves more *suggestions* which might indicate that conflicts in most cases arise from a suggestion. Also, *self-explanation* and *others* are significantly lower during conflicts. Considering that *self-explanations,* by definition, are mostly used by drivers, one explanation could be that during conflicts, drivers stopped coding and instead justified or addressed navigators' ideas. Thus, during conflicts, it may seem that pairs are not working on coding but talking; however, we believe this is a necessary element of productive conflict.

Conclusion

Conflicts are common in many collaborative activities. However, conflicts can be made more productive by listening to and challenging partners' arguments and building upon each other's ideas (Lee et al., 2015). Our work demonstrates that, as inferred from Mercer's theory (Mercer and Littleton, 2007), conflict can be productive through Exploratory talk, which occurs most often during conflicts rather than before and after conflicts. By bringing together Mercer's framework with parallel research on conflict, we get a richer understanding of what constitutes a *productive conflict*. That is, conflicts involve students challenging and responding with justification and talk that moves the problem solutions forward. These findings can help guide researchers and teachers on ways to identify how productive and unproductive conflicts discursively emerge and unfold in pair programming.

References

Basil-Shachar, J., Hod, Y., Ben-Zvi, D., & I-CORE, L. I. N. K. S. The Emergence of Norms in a Technology-Enhanced Learning Community.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.

Isohätälä, J., Näykki, P., Järvelä, S., & Baker, M. J. (2018). Striking a balance: Socio-emotional processes during argumentation in collaborative learning interaction. *Learning, Culture and Social Interaction*, *16*, 1-19.

Kirschner, P. A., Beers, P. J., Boshuizen, H. P., & Gijselaers, W. H. (2008). Coercing shared knowledge in collaborative learning environments. *Computers in human behavior*, *24*(2), 403-420.

Lee, D., Huh, Y., & Reigeluth, C. M. (2015). Collaboration, intragroup conflict, and social skills in project-based learning. *Instructional science*, *43*(5), 561-590.

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Routledge.

Noroozi, O. (2020). Argumentation-based computer supported collaborative learning (abcscl): the role of instructional supports. *European Journal of Open Education and E-learning Studies*, *5*(2).

Tsan, J., Vandenberg, J., Zakaria, Z., Boulden, D. C., Lynch, C., Wiebe, E., & Boyer, K. E. (2021, March). Collaborative Dialogue and Types of Conflict: An Analysis of Pair Programming Interactions between Upper Elementary Students. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 1184-1190).

Wise, A. F., Antle, A. N., Warren, J., May, A., Fan, M., & Macaranas, A. (2015). What kind of world do you want to live in? Positive interdependence and collaborative processes in the tangible tabletop land-use planning game Youtopia. International Society of the Learning Sciences, Inc.[ISLS].

Zakaria, Z., Vandenberg, J., Tsan, J., Boulden, D. C., Lynch, C. F., Boyer, K. E., & Wiebe, E. N. (2021). Two-Computer Pair Programming: Exploring a Feedback Intervention to improve Collaborative Talk in Elementary Students. *Computer Science Education*, 1-28.

Technology’s Role in Supporting Collaborative Interactions: An Ecological Approach

Dhvani Toprani, Marcela Borge
dhvanitoprani@psu.edu, mborge@psu.edu
The Pennsylvania State University

Abstract: In this paper we analyze design of a technology-based learning environment facilitated using digital technologies to understand its trade-offs as children work on collaborative design projects. Using interaction analysis, we analyze one group’s interactions around digital technology from three afterschool lessons. Our findings track the change in collaborative interactions as the technology was owned individually and collectively by the group, followed by a discussion on how this approach of analysis can build transformative learning environments.

Policymakers around the world have invested billions of dollars in introducing innovative technologies in K-12 classrooms anticipating improvements in learning outcomes. 2019 set a record number of investments into education technology companies around the world, which was greater than the combined investment between 1998 – 2017 (Adkins, 2020). What follows this investment is a need to leverage technology for creating transformative learning environments, instead of merely replicating existing teaching approaches (Kirkwood & Price, 2014). Researchers have explored how different technologies help or hinder aspects of learning for individuals or groups, however, few researchers have examined the impact of technology to understand its trade-offs from an ecological perspective, where multiple learning outcomes can be examined at different social levels (Toprani et al., 2021; Borge & Mercier, 2018). Towards that goal we analyze design of a technology-based learning environment facilitated using digital technologies to understand its trade-offs as children work on collaborative design projects. We focus on answering: How does the use of digital technology impact children’s collaborative interactions within design learning environment? We define collaborative interactions at the small-group level where learners are constructing knowledge by engaging in collective sense-making (Stahl, 2006).

The study was conducted at an afterschool club in a Northern US charter school that develops technologically enhanced informal learning environments for promoting design thinking among children between ages 8 – 10 years. We analyzed one group’s interactions over lesson 5 (L5) and lesson 6 (L6) as they worked together to redesign a local mall pertaining to the larger problem of malls in America going out of business. Members of the group; Baylee, Lance, and Joey; in L5 began by understanding the problem and gathering information about it, followed by organizing the gathered information in L6. Students worked with digital and physical technologies i.e. big size post-it notes, and laptops with google slides and information website about the design challenge.

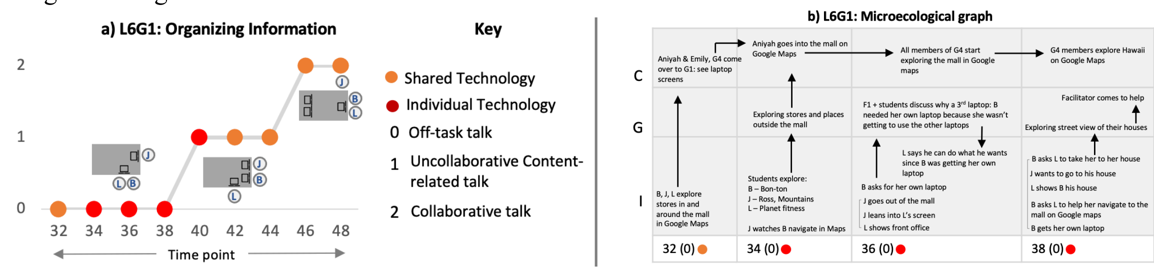


Figure 1. a) Ecological interaction graph representing the nature of talk and technology position in L6 b) Microecological graph of group 1’s Google Map exploration in lesson 6

Using Interaction Analysis (Jordan & Henderson, 1995) and quantification of qualitative data, we examined the pattern of collaborative interactions in comparison to the position of technology. We first created detailed content logs of the two lessons to identify instances (two-minute segment of video data) of collaborative interaction. Each instance was coded and plotted on a graph to understand how the position of the technology impacts collaborative interactions. We analyzed technology’s position to determine if it is individually owned where a student is working on a laptop, without sharing access to manipulate the laptop or view of the screen with others; or collectively owned where two or more students are working on one laptop, with each student having access to manipulate the laptop and view the screen. These are represented on the x-axis of the graph, shown in Fig. 1 (a) as red dots and orange dots respectively. We categorized interactions as 1) off-task i.e. students interacting with each other about topics unrelated to the design project 2) uncollaborative content-related i.e. students exchanging design project related ideas but not collaborating with each other, and 3) collaborative i.e. two or more students working together to generate, negotiate, and build on each other's ideas. These are represented on the y-axis of the graph as 0, 1, 2 respectively.

In L5 and L6 group 1 was working with different digital technologies. We found that digital technologies supported collaborative interactions when they were used as shared tools by each student in the group. When owned individually, they hindered collaboration but fostered other relevant 21st century skills. In Fig. 1 (a) from timepoint 34-44 students didn't engage in collaborative interactions as they owned technology individually to explore the mall on Google maps (34-38). At timepoint 40 the facilitator comes to the group and helps them reflect on how they were interacting with the technology. This led them to rearrange the technology in the shared space (42–48). Following this rearrangement students started negotiating ideas and searched ideas using one laptop.

As students begin to own technology individually, collaborative interactions decreased. The case study presented below, using microecological graph (Borge & Mercier, 2018), describes the interactions that took place from timepoints 32-38 in L6 when technology was owned individually, as shown in Fig. 1 (b). Although students didn't engage in collaborative interactions when technology was individually owned, there were other relevant skills that were developed through these interactions.

When students owned technology individually, they engaged in exploratory talk, conducting self-driven explorations of novel online applications. These interactions often emerged out of a design related need but gradually evolved to focus more on exploring the tool. These interactions didn't directly enhance the designs that the students were working on but helped them to develop other skills like technology fluency and learning by observation. In this case study students are exploring different features in Google maps, at first to understand the structure of the mall, and later to learn more about the application. The microecological graph captures the timepoints on the x-axis and the level of interaction on y-axis as individual (I), group (G), and community (C). The explorations triggered interactions at multiple ecological levels where students were exploring Google maps on their own, to troubleshoot technical difficulties with team members, and learn by observation as a community. At timepoint 32 the laptops were in the shared space. Joey and Lance were imitating each other on their own screens and Baylee used Joey's laptop to help him move around in Google maps. From timepoint 34, Joey learned how to navigate in Google maps and began exploring individually by walking to stores outside the mall and to nearby mountains. Lance on his screen tries to open the online virtual tour of another store (34). Seeing Joey and Lance explore individually, Baylee asks the facilitator for another laptop (36). After exploring around the mall area, at timepoint 38 students learn how to switch between satellite and street views in Google maps and find their houses. Between timepoint 32–36, level C, Aniyah and Emily from group 4 come to observe what group 1 was doing. Group 1 didn't interact with group 4 intentionally, but Emily and Aniyah watched group 1 navigate in Google maps and eventually explored places on their own in their small group (38, level C).

As we are trying to expand the access to technology by giving every student a laptop in school, it is equally important to design competent learning environments to foster relevant skills like collaboration and digital fluency so that children can build innovative designs to solve complex problems with technology. Taking the approach to analyze technology's affordances from an ecological perspective will bring about the shift in viewing technology as a tool for simplifying existing learning processes to transforming what children learn pedagogically.

References

Adkins, S. (2020). The 2019 Global Learning Technology Investment Patterns: Another Record Shattering Year. Metaari.

Borge, M. & Mercier, E. (2018). Towards a Cognitive Ecological Framework in CSCL. In Kay, J. and Luckin, R. (Eds.) Rethinking Learning in the Digital Age: Making the Learning Sciences Count, 13th International Conference of the Learning Sciences (ICLS) 2018. London, UK: International Society of the Learning Sciences.

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. Journal of the Learning Sciences, 4(1), 39–103.

Kirkwood, A., & Price, L. (2014). Technology-enhanced learning and teaching in higher education: what is 'enhanced' and how do we know? A critical literature review. *Learning, media and technology*, *39*(1), 6-36.

Stahl, G. (2006). Group cognition. London, UK: The MIT Press.

Toprani, D., AlQahtani, M., & Borge, M (2021) Examining Technology Use and Evaluation in Computer-Supported Collaborative Learning: A Systematic Review. In M. Spector, B. B Lockee, & M. D Childress (Eds.), *Learning, Design, and Technology. An International Compendium of Theory, Research, Practice, and Policy.* Springer.

Embodiment and Social Interactions in a Class Virtual Reality

Gennie Mansi, Blair MacIntyre, Jessica Roberts
gennie.mansi@gatech.edu, blair@cc.gatech.edu, jessica.roberts@cc.gatech.edu
Georgia Institute of Technology

Abstract: There is a growing enthusiasm to use VR to improve remote student learning experiences. However, incongruities between students' *virtual embodiment* – as avatars in the virtual environment – and *physical embodiment* – from their biological bodies – can significantly impact them. We observed a university class poster session held entirely in an immersive Mozilla Hubs environment. We found that incongruities in embodiment created both challenges and significant opportunities for students to collaborate and learn in a shared space.

Introduction and background

Virtual reality (VR) fills the gap between the physical and digital worlds by enabling interactions with objects in the virtual world (Liu et al., 2019). In recent years, lowered costs and improved system capabilities have encouraged applications that provide students with immersive experiences that permit them to interact with content in meaningful contexts (Abadia et al., 2018) and encourage active collaboration through shared experiences (Gao et al., 2019). Some research has shown that VR changes physical interactions within the environment, impacting cognition (Bailey et al., 2016) and student learning (Abadia et al., 2018). Additionally, embodiment in VR can provide a high level of social presence and cultivate a sense of community (Liu et al., 2019; Holt et al. 2020). However, little research explores how incongruities between *physical embodiment* – the sensations that come with having and controlling a physical body – and *virtual* embodiment – the sensations that come with controlling a body in VR – can affect learning in formal higher education settings.

We studied this gap in a virtual poster session in a postsecondary course. Through observations and student interviews, we investigated the following research questions: (RQ1) How do incongruities between physical and virtual embodiments impact social interactions in a classroom? (RQ2) How do these incongruities open opportunities that are otherwise missing from remote classroom environments?

Methods

This study centers on a VR poster session conducted in a 3D User Interfaces class taught by the second author in Fall 2020. The 75-minute session was held in Mozilla Hubs – a free, open source, virtual collaboration environment that can be accessed without installation or app store (see https://hubs.mozilla.com/docs) – as a part of the regular course syllabus. The course's 33 students were divided into three Hubs rooms. The lead researcher recorded observations of the 11 study participants (2 female graduate students, 4 male graduate students, and 5 male undergraduate students) and performed follow-up interviews with 8 of them.

During the poster session the first author participated as an avatar and took field notes focused on interactions in which congruities and incongruities in embodiment were observed to impact students' social interactions. A group interview was conducted after the session with seven students. One student was interviewed individually within a week using the same interview protocol. Interviews focused on in-person and remote classroom experiences and on their social interactions. Episodes of collective social interaction were sorted into two emergent themes at the intersection of embodiment, collaborative problem solving, and social interactions.

Findings

The first emergent theme was opportunities for collaborative learning in VR. A large part of students' shared experience was exploring their physical and virtual embodiment in the Hubs environment. This opened opportunities for playfulness, creativity, and genuine peer-to-peer interactions as students discovered novel aspects of the environment. For example, at one point a student asked a large group of peers whether each of them was physically sitting or standing. The students then realized that avatars were at different heights depending on their physical posture, prompting excited conversation and several bobbing avatars:

Student A: *Oh! I didn't realize that's why I'm so high up.*

Student B: *That's why I've been looking up to everyone this whole time.*

Student C: *Yeah. You can even set it up to crouch.*

Student D: *Haha. You can even dance a little.*

At other times, embodiment facilitated student communication and peer teaching episodes. For example, students used gestures to instruct others how to navigate the virtual space and control their avatars. These episodes were prevalent during class and were referenced as positive experiences by multiple students during the interviews. Such peer mediation mimics the social learning supports that happen naturally in face-to-face classrooms but are often missing in remote spaces.

While virtual embodiment through avatars provided positive collaborations when replacing some aspects of co-located interactions at typical poster sessions, we found it could also have unexpected effects. For example, virtual walking is relatively inefficient and can cause motion sickness, so most students teleported through the space. However, they often accidentally teleported into other presentation listeners. Additionally, Hubs mimics sound in physical embodiment so people farther away sound fainter, but students, unfamiliar with this feature, felt they needed to get distractingly close to a presenter to hear. In combination, students perceived the "normal" behavior of walking by a poster presentation as more disruptive in the virtual space and avoided doing so.

Discussion

While the interactions described may seem trivial, they highlight two key strengths of VR. First, it provides the opportunity for students to interact naturally within a shared space. In the interviews, students reported that in classes held through video conferences they felt alone in what one student referred to as a "sea of information" with no unstructured interactions such as casual conversations. Students reported that VR allowed for interactions that decreased feelings of isolation, indicating these interactions built a sense of community. Second, the VR environment helped students participate in peer-to-peer learning, which made class more engaging and comfortable.

In remote collaboration settings, students felt that interactions were awkward when they could not have spontaneous interactions. While VR solves some challenges such as moving around a shared space, incongruities in embodiment can create unusual or frustrating social situations, such as with teleporting. Although such incongruities may make VR seem undesirable, it was exactly these challenges that presented opportunities for students to engage in open discourse and collaborative exploration. This in turn engaged students throughout the session and created a sense of community. VR technology increasingly grows more advanced, and as interactional elements are improved, opportunities to learn and explore novel environments together will expand.

Limitations

This is a small observational study to serve as a foundation for future empirical work. We interviewed a small number of students who were novice VR users, but interactions between non-novices may differ. Additionally, we explored an immersive VR experience, so students without headsets may not share the same effects of embodiment.

Conclusion and future work

VR can powerfully engage students in remote settings, allowing them to share digital artifacts, talk, and move in a shared space. However, incongruities between physical and virtual embodiment can alter social interactions between students, but also provide new opportunities for students to explore and learn from one another. Educators who want to use VR in their courses need to design their courses with embodiment in mind.

Future research should identify how to balance structure and spontaneity in remote education settings. For example, the integration of problem solving and platform exploration sessions in a course could provide a supportive class structure that helps students collaboratively explore navigation features while acclimating to how their peers act, move, and speak in VR. We remain optimistic that free, open-source technologies such as Hubs will continue improving accessibility to VR technology, paving the way to impactful remote student experiences.

References

Abadia, R., Calvert, J. & Tauseef, S.M. (2018). Salient features of an effective immersive non-collaborative virtual reality learning environment. In *Proc ICETC '18*. New York, NY, USA, 268–278. ACM.

Bailey, J.O., Bailenson, J.N., and Casasanto, D. (2016). When does virtual embodiment change our minds? *Presence: Teleoper. Virtual Environ.* 25(3): 222–233.

Gao, C., Bai,Y., and Goda, B. (2019). Are We Ready for a VR Classroom? A Review of Current Designs and a Vision of Future Virtual Reality Classrooms. In *Proc of the 20th Annual SIG Conf on Information Technology Education* (*SIGITE '19*). New York, NY, USA, 39. ACM.

Holt, E. A., Heim, A. B., Tessens, E., & Walker, R. (2020). Thanks for inviting me to the party: Virtual poster sessions as a way to connect in a time of disconnection. *Ecology and Evolution*, 1–8.

Liu, Y., Fan, X., Zhou, X., Liu, M., Wang, J. and Liu, T. (2019). Application of Virtual Reality Technology in Distance Higher Education. In *Proc of the 2019 4th Int'l Conf on Distance Education and Learning* (*ICDEL 2019*). New York, NY, USA, 35–39. ACM.

Teaching about COVID-19: Using a Virtual Epidemic to Contextualize and Problematize Infectious Disease Epidemiology in a High School Class

Amanda Strawhacker (co-chair), Tufts University, amanda.strawhacker@tufts.edu
Tyler Hansen (co-chair), Utah State University, tyler.hansen@usu.edu
Colby Tofel-Grehl (co-chair), Utah State University, colby.tg@usu.edu
Yasmin Kafai (co-chair), University of Pennsylvania, kafai@upenn.edu

Abstract: *SPIKEY-20* is a virtual epidemic within the Whyville.net virtual world that infected over 400 online players during a month-long outbreak. We report on a design study of one high school teacher who used *SPIKEY-20* in a NGSS-aligned curriculum on epidemiology with his AP Biology class during the COVID-19 pandemic. Class observations and teacher interviews illustrated how students' online experiences with the virtual epidemic helped the teacher to contextualize and problematize their understanding of community spread, prevention, and economics with their lived experiences of COVID-19.

Keywords: COVID-19, SPIKEY-20, High school biology, virtual worlds, epidemiology

Overview

As the COVID-19 pandemic impacted schools around the world, much attention has focused on how teachers and students handle the transitions to physically distanced, hybrid, or online learning. However, a paucity of research exists explaining student understanding of infectious disease. The absence of core ideas underpinning topics of epidemiology and disease prevention in the Next Generation Science Standards (NGSS), K-12 science learning standards, highlights the lack of attention on these now critical areas of science education. As students, families, and communities have been dramatically and personally impacted by the COVID-19 pandemic, infectious disease epidemiology is an increasingly relevant subject for science classes (Straif-Bourgeois, Ratard & Kretzschmar, 2014). One critical issue is how to engage students in meaningful learning about relevant aspects of disease transmission and prevention that can promote understanding and impact their behaviors and actions. A participatory epidemic simulation could address this issue by providing a context to model and examine the dynamic interactions of disease vectors while also adding real-time experience to inquiries (Kafai & Dede, 2014).

The SPIKEY-20 virtual epidemic in the Whyville.net online community was launched as a timely opportunity to immerse students in this type of experiential learning about infectious disease outbreaks and their prevention in an safe, free online environment that is accessible from home or school. In this study, the research team developed a NGSS-aligned science curriculum about infectious disease spread and prevention, using the SPIKEY-20 outbreak in Whyville as a testbed for student explorations of epidemiology topics, including community spread, testing and public health prevention practices, and population infection modeling (see Figure 1).

(a) (b) (c)

Figure 1. Screenshots from Whyville.net, including (a), a player suffering SPIKEY-20 and a disease rate chart, (b), the Whyville City Hall Lobby, where players can get tested for SPIKEY-20 and purchase PPE, and (c), simulators of infectious disease spread, for students to explore impact of infection duration and population size.

A teacher implemented the curriculum in a high school AP Biology classroom with 18 students in weekly 1-hour lessons during the roughly 4-week outbreak. Lessons consisted of individual computer-based play in Whyville, and a mix of individual, small-group, and large-group off-screen activities designed to support reflection about Whyville observations and experiences. Building on a decade of research on virtual epidemics within Whvyille (e.g. Fields et al., 2017; Neulight et al., 2007), SPIKEY-20 seeks to extend this design research

 © ISLS

approach, using iterative trial and revision informed by teacher and student feedback to shape the epidemic and curriculum (Edelson, 2002). Our analysis of recordings and transcripts of classroom observations (two off-screen lessons) and teacher interviews (following each lesson) focused on how students connected the real COVID-19 and the virtual SPIKEY-20 epidemic.

Students made connections between the virtual SPIKEY-20 and real COVID-19, that relate to critical aspects of infectious disease epidemiology (Straif-Bourgeois, Ratard & Kretzschmar, 2014): an understanding of (a) biological concepts such as germs and infection, (b) processes such as incubation and immunity within larger ecological contexts and, most importantly, (c) community factors that contribute to or hinder an epidemic outbreak. The curriculum connected the Whyville unit directly to AP biology concepts of cell anatomy and genetics, which led to discussions about how viruses spread by attacking cell organelles to replicate viral DNA.

While students demonstrated a strong grasp of virus anatomy and physiology, they had challenges understanding how these relate to public health behaviors. Students' experiences with the virtual SPIKEY-20 epidemic supported their conceptual understanding of factors that impact spread of COVID-19 in the following ways: (1) *Understanding data visualizations:* Students engaged with data visualizations of the virtual epidemic (see Figure 1a) to track player behaviors and identify disease vectors, interpret population infection rates, and compare to published visualizations about COVID-19; (2) *Examining data inconsistencies*: In tracing SPIKEY-20 infections in their own class, they noticed a counter-intuitive finding about spread (students who visited more places in Whyville were not more likely to become infected, due to infection being algorithmically initiated at login rather than through organic player interactions), leading to discussions about sampling error and experimentation limits in COVID-19 population statistics; (3) *Simulating vectors of epidemic outbreaks*: Students used online simulators to model infection and epidemic spread based on factors like duration and rate of infection, and compared outcomes in relation to public messaging around herd immunity in COVID-19 (see Figure 1c); and (4) *Discussing economic barriers to personal protective equipment:* Students noticed differential access to individual SPIKEY-20 preventive measures in Whyville such as cheap masks and expensive full-body suits (see Figure 1b), and discussed how economic affordance impacts community health.

We found that students' participation in SPIKEY-20 provided an experientially motivating and academically salient context for learning about factors of infection, incubation, and community spread critical to understanding real-world epidemics such as COVID-19. In contrast to traditional epidemiology curricular interventions (e.g. Panou et al., 2013), results also revealed an emotional investment in SPIKEY-20, with students drawing on personal anecdotes, current events, and first-hand experiences contracting COVID-19 to interpret SPIKEY-20 trends. This personal connection inspired students to compare public health attitudes and preventive behaviors observed in Whyville and in their real-life communities, leading some to intensify their SPIKEY experience by logging extra time in Whyville to earn more PPE. We interpret these findings to mean that virtual epidemics can offer students an immersive outlet to explore emotional and epidemiological aspects of the real-life COVID-19 pandemic, yielding heightened awareness of public health prevention measures in everyday life.

References

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *The Journal of the Learning sciences*, *11*(1), 105-121.

Kafai, Y. B. & Dede, C. (2014). Learning in Virtual Worlds. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences,Second Edition* (pp. 522-544). New York, NY: Cambridge University Press.

Fields, D. A., Kafai, Y. B., Giang, M. T., Fefferman, N., & Wong, J. (2017). Plagues and people: Engineering player participation and prevention in a virtual epidemic. In *FDG'17, International Conference on the Foundations of Digital Games*, Hyannis, MA, USA, August 14-17, 2017. New York, NY: ACM.

Neulight, N., Kafai, Y., Kao, L., Foley, B. & Galas, C. (2007). Children's participation in a virtual epidemic in the science classroom: Making connections to natural infectious diseases. *Journal of Science Education and Technology, 16*(1), 47-58.

Panou, Å., Souglis, A., Zachos, D., Sotiropoulos, A., & Hatziharistos, D. (2013). Evaluation of an educational programme on the influenza in elementary schools. Journal of Physical Education and Sport, 13(4), 570.

Straif-Bourgeois S, Ratard R, Kretzschmar M. (2014). Infectious disease epidemiology. In W. Ahrens and I. Pigeot (Eds.), *Handbook of Epidemiology* (pp. 2041–2119). New York, NY: Springer.

Acknowledgements

This work was supported by a grant (NSF RAPID 2031748) from the National Science Foundation to Yasmin Kafai. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation or University of Pennsylvania.

Promoting College Students' Systems Thinking during Pandemic

Yawen Yu, The University of Hong Kong, yawenyu@connect.hku.hk
Xiaofeng Li, Clarion University, xli@clarion.edu
Shengwen Dong, Tianjin Medical University, dongshengwen@tmu.edu.cn
Gaowei Chen, The University of Hong Kong, gwchen@hku.hk

Abstract: This study explores the possibilities of using asynchronous discussion forums to engage students in inquiry learning to promote their systems thinking during the pandemic. Twenty-eight students from a medical university in China participated in this study. Compared with the traditional online teaching unit, students had better learning outcomes and developed systems thinking skills in the inquiry-based teaching unit. Qualitative findings unpacked how students engaged in inquiry-learning and why such an approach promotes systems thinking.

Supporting systems thinking with inquiry-based pedagogical approaches

Systems thinking is considered an important reasoning skill for medical students (Michael, 2007). Lira and Gardner (2017) defined systems thinking as a reasoning process that involves predicting and explaining phenomena by reasoning within and between levels of biological organization and across system components. However, systems thinking is hard to grasp for medical students and for all students across all educational levels. Inquiry-based pedagogical approaches, which occurred in regular classrooms with months of extended designs have been shown as an effective way to support students' systems thinking (Hmelo et al., 2000; Hmelo-Silver et al., 2017; Liu & Hmelo-Silver, 2009). However, the sudden outbreak of COVID-19 requires teachers to make a rapid transition to online teaching. How to better support students' systems thinking in an online learning environment remains an issue. In this study, we explore the possibilities of using asynchronous discussion forums (e.g., Chen et al., 2012) to engage students in inquiry learning to promote their systems thinking in a medical school in China. We ask following questions: 1. To what extent do students improve their understanding of key concepts through engaging in inquiry-based learning? 2. To what extent do students develop systems thinking in asynchronous discussion forums that incorporated inquiry-based pedagogical approach? 3. How do students engage with inquiry-based learning in asynchronous discussion forum to promote their systems thinking?

Method

This study took place in a 16-week midwifery course in a medical school in China. The course is mandatory for junior nursing students. As one of the four co-teachers, Dr. D participated in this study and provided instructions that covered the following topics: nursing care for women with complications during pregnancy (Unit 1) in weeks 2–3, and the other covered home care for newborns (Unit 2) in weeks 5–6. In Unit 1, students watched video lectures and attended synchronous video conferences. In Unit 2, students read the assigned readings, asked questions in discussion forums, and participated in student-led discussions. Twenty-eight students participated in this study. The following data were collected during the study. An online assessment made up of twenty multiple-choice questions designed to measure student understanding of the key concepts in both units. In addition, the instructor posted five open-ended questions in the forum by the end of each unit. The questions asked students to explain why certain phenomena would happen. Overall, students posted 106 answers in Unit 1 and 121 in Unit 2, respectively. In addition, students posted 39 questions in the forum for Unit 2. Students made 33 replies to the questions posted by their peers. As an important scaffold, the teacher made 39 responses to student questions

Student online assessment scores were analyzed using a *t*-test to answer RQ1. To answer RQ2, content analysis was conducted to analyze student answers for the five open-ended questions. The coding scheme is a four-level trajectory starting from student belief in everything happening in the system was detrimental to the human body, moving up to using some phenomena to explain other phenomena without unpacking any mechanisms, advancing towards identifying certain mechanisms for the phenomenon. Ultimately, students located multiple mechanisms to the problem and connected them with appropriate relationships (Hmelo-Silver et al., 2007). The analysis unit used was each individual post in the discussion forum. To answer RQ3, content analysis was conducted to analyze the nature of questions (e.g., explanation-seeking questions or facts-seeking questions; van Aalst, 2009) students posted in the forum. An open coding (Charmaz, 2006) was conducted to unpack how students answered their peers' questions.

Results

A two-sample *t*-test was performed to examine if the learning outcomes in Unit 2 was significantly better than for Unit 1. Table 1 shows that students performed better in Unit 2 than Unit 1 ($t = 10.974$, $p < .001$). The result suggested students had a better understanding of key concepts in Unit 2 when they conducted inquiry-based learning.

Table 1: Students' assessment scores across two units

	Unit 1	Unit 2
M	75.04	94.20
SD	7.49	4.49
n	25	25

Content analysis showed that students' answers appeared to be mostly at the second (77.3% in Unit 1 and 41.3% in Unit 2) and third levels (12.3% in Unit 1 and 38.4% in Unit 2) in which students began to identify mechanisms to explain the phenomena. However, fewer posts were observed at the first (6.6% in Unit 1 and 1.7% in Unit 2) and fourth level (3.8% in Unit 1 and 18.2% in Unit 2). Moreover, students demonstrated fewer instances of low-level systems thinking (levels 1 and 2) in Unit 2 than Unit 1 but more instances of high-level systems thinking (levels 3 and 4) across the two units. The results suggested that in Unit 2, students were able to unpack the underlying mechanisms and make connections across different levels to explain the phenomenon, which was an important indicator of improved systems thinking (Eberbach et al., 2012).

Data analysis results showed that students generated 22 (out of 39) fact-seeking questions and 17 explanation-seeking questions to distinguish two similar syndromes or challenge the validity of certain operations. In addition, students made replies to their peers' questions. They connected prior knowledge or applied intuitive opinions to the question, answered the question with external authoritative resources, critically assessed answers provided by their peers and reflected upon their own ideas, and asked further questions. Among these, there are 9 instances in which the authors revisited their questions and made further replies to their peers' replies. In these cases, students demonstrated reflections and even challenged their peers' answers by asking further questions.

Moreover, there were 13 instances in which the depth of the discussion exceeded one which indicates there were replies to initial replies. This suggested student engaged with sustainable inquiry. Among these 13 instances, eight were answers to explanation-seeking questions. The result indicates that explanation-seeking questions are more likely to lead sustainable inquiry and the authors of explanation-seeking questions have a greater willingness to reflect how their peers' answers confirmed or challenge their existing answers. In summary, explanation-seeking questions elicited more responses and further questions.

Discussion

The findings of this study demonstrated that students developed systems thinking when asynchronous discussion forums incorporated an inquiry-based pedagogical approach. In these online discussion forums, students are given venues to ask explanation-seeking questions. These questions provided students a chance to focus on phenomena and use mechanisms to explain such phenomena. This encourages students to make meaningful connections between phenomena and mechanisms, which is an indicator of improved systems thinking. The results extend existing understanding that phenomena-oriented questions promoted systems thinking, but demonstrating the effectiveness of encouraging students to ask questions in supporting their systems thinking (e.g., Hmelo et al., 2000; Liu & Hmelo-Silver, 2009). In addition, students made replies to these questions. Such diverse information confirms or challenges students' existing understanding, which encourages them to reflect on how to better incorporate new information to their existent understanding.

References

Charmaz, K. (2014). *Constructing grounded theory.* Sage.

Chen, G., Chiu, M. M., & Wang, Z. (2012). Social metacognition and the creation of correct, new ideas. *Computers in Human Behavior, 28*(3), 868-880. doi: 10.1016/j.chb.2011.12.006

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. doi:10.1207/s15327809jls0903_2.

Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *The Journal of the Learning Sciences, 16*(3), 307-331.

Hmelo-Silver, C. E., Jordan, R., Eberbach, C., & Sinha, S. (2017). Systems learning with a conceptual representation: a quasi-experimental study. *Instructional Science, 45*(1), 53-72.

Lira, M. E., & Gardner, S. M. (2017). Structure-function relations in physiology education: Where's the mechanism? *Advances in Physiology Education, 41*(2), 270-278. doi:10.1152/advan.00175.2016

Michael, J. (2007). What makes physiology hard for students to learn? Results of a faculty survey. *Advances in Physiology Education, 31*(1), 34-40. doi:10.1152/advan.00057.2006.

Acknowledgments

This work is supported by the Public Policy Research Funding Scheme from the Policy Innovation and Co-ordination Office of the Government of the HKSAR (Grant No. 2017.A8.073.18C).

Grasping Evidence with EDDiE: A CSCL Tool to Support Collaborative Reasoning about Disagreements in Multiple Documents

Toshio Mochizuki, Senshu University, Japan, tmochi@mochi-lab.net
Clark A. Chinn, Randi M. Zimmerman
clark.chinn@gse.rutgers.edu, randi.zimmerman@gse.rutgers.edu
Rutgers University, USA
Myat Min Swe, Htay Min Khaung, Seiji Sekine,
myatmins@spiceworks.co.jp, htay.kaung@spiceworks.co.jp, sekine@spiceworks.co.jp
Spiceworks Myanmar, Myanmar & Japan,

Abstract: We have developed and investigated a web-based CSCL system to support students as they analyze multiple controversial documents. The application, "EDDiE" (Electronic Documents Disagreements Evaluation), allows students to interactively create a visual representation of information from multiple conflicting documents in order to resolve disagreements among the documents. The design is based on the Grasp of Evidence (GoE) Framework (Duncan et al., 2018). Using EDDiE, students collaboratively analyze and compare the quantity and quality of evidence (including testimony by knowledgeable sources) for different claims. In a preliminary study, EDDiE's epistemic scaffolds promoted productive epistemic discourse.

Introduction

Citizens in the 21st century find it challenging to make sense of the rampantly conflicting information they encounter (Barzilai & Chinn, 2020; Chinn et al., 2020). People must learn to resolve the widespread conflicts and disagreements in digital media. Thus, recent scholarship has made increasing efforts to help people deal with conflicting information and disagreements across multiple sources (e.g., Barzilai et al., 2020a; 2020b). Effective integration across documents requires reasoners to use competent strategies both to identify and explain the disagreements that exist (Thomm et al., 2016) and to resolve these disagreements.

This poster paper describes the design of a CSCL system, referred to as EDDiE (Electronic Documents Disagreements Evaluation), a multi-user, interactive web application. Users collaboratively read a set of multiple documents and create a visual graphic organizer to synthesize information as well as analyze and resolve disagreements. In a preliminary efficacy trial, EDDiE's epistemic scaffolds promoted productive epistemic discourse directed at disagreement resolution; details of the results are not included in this paper due to its brevity.

Epistemic scaffolds designed and implemented in the web-based CSCL

The epistemic scaffolds of EDDiE (Figure 1) are grounded theoretically in the Grasp of Evidence (GoE) framework (Duncan et al., 2018). EDDiE encourages students to resolve disagreements by engaging systematically in five evidential practices identified by the GoE framework. The graphical elements invite students to participate in productive discussions regarding each document's evaluation, the quality and strength of evidence (including empirical evidence described in the documents, the evidence of testimony by experts, etc.), the relationships between evidence and positions, what the disagreements are, and how to resolve them. In short, EDDiE aims to promote collaborative discussions about how evidence can be used to resolve disagreements.

The GoE framework posits five dimensions of evidence evaluation. Below, we describe each, explain how EDDiE scaffolds each, and explain how these scaffolds can promote productive epistemic discourse.

- **Evidence analysis** (understanding the components of empirical studies and how they fit together). Students record important elements of evidence (e.g., sample size, critical comparisons, results) in the tableau. *Epistemic discourse: They discuss the study details as they compare and evaluate studies across documents, analyzing differences between studies as potential reasons for disagreements.*
- **Evidence evaluation** (determining if evidence is of high methodological quality). Students denote the quality of evidence via *color* of evidence circles in the tableau. *Epistemic discourse: Students evaluate methodological processes (e.g., appropriate sample size, proper controls, etc.); they may conclude that some lower-quality studies should be weighted less.*
- **Evidence interpretation** (determining how strong evidence is in supporting or weighing against explanations). The thickness of arrows between evidence and claims reflect evidence strength in supporting or opposing various claims. Dotted arrows mark disagreements. *Epistemic discourse: Students discuss what the evidence shows, how relevant it is, how diagnostic it is, how directly it supports claims, etc.*

These deliberations can also illuminate which positions are best supported by the available evidence.

- **Evidence integration** (determining the extent to which larger bodies of evidence support or weigh against theoretical claims). *The size* and *shape* of evidence circles mark evidence quantity. Larger ovals denote more evidence; smaller circles denote less evidence. *Epistemic discourse: Students discuss issues such as how much evidence there is for each position, whether there are multiple lines of evidence for different positions, the degree to which the evidence is consistent, etc.*
- **Lay use of evidence** (determining the credibility of scientific claims in everyday communication, such the trustworthiness of sources, consensus among experts, and cross validation by knowledgeable others). The *knowledgeable supporters* circles and arrows (size and color of circles, boldness of arrows) reflect lay evaluation of quality and consensus of knowledgeable others. *Epistemic discourse: Students discuss the extent to which the sources are competent, biased or unbiased, in consensus, and so on.*

In short, we expected the scaffolds to support students in addressing and potentially resolving disagreements through systematic engagement with evidence along these five dimensions.

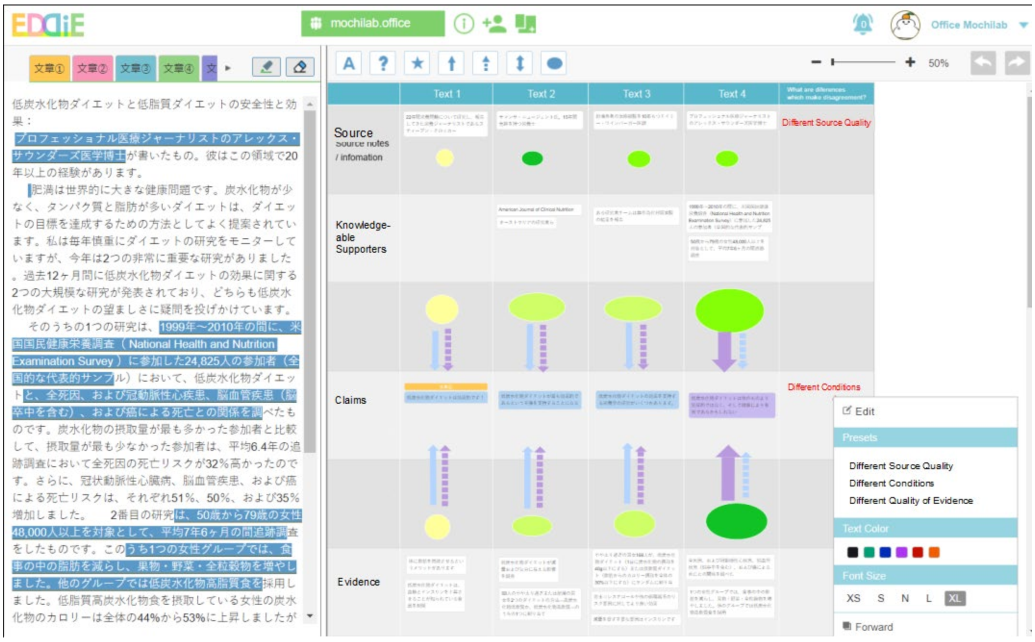


Figure 1. EDDiE's interface and functions

References

Barzilai, S., & Chinn, C. A. (2020). A review of educational responses to the "post-truth" condition: Four lenses on "post-truth" problems. *Educational Psychologist*, *55,* 107–119.

Barzilai, S., Mor-Hagani, S., Zohar, A. R., Shlomi-Elooz, T., & Ben-Yishai, R. (2020a). Making sources visible: Promoting multiple document literacy with digital epistemic scaffolds. *Computers & Education*, 157, 103980.

Barzilai, S., Thomm, E., & Shlomi-Elooz, T. (2020b). Dealing with disagreement: The roles of topic familiarity and disagreement explanation in reasoning about conflicting expert claims and sources. *Learning and Instruction*, 69, 101367.

Chinn, C. A., Barzilai, S., & Duncan, R. G. (2020). Education for a "post-truth" world: New directions for research and practice. *Educational Researcher, 52* (1), 51–60.

Duncan, R. G., Chinn, C. A., & Barzilai, S. (2018). Grasp of evidence: Problematizing and expanding the next generation science standards' conceptualization of evidence. *Journal of Research in Science Teaching*, 55(7), 907–937.

Thomm, E., & Bromme, R. (2016). How source information shapes lay interpretations of science conflicts: Interplay between sourcing, conflict explanation, source evaluation, and claim evaluation. *Reading and Writing*, 29(8), 1629–1652.

Acknowledgments

This work is supported by Telecommunication Advancement Foundation and the Japan Society for the Promotion of Science (JSPS KAKENHI Grant numbers JP16K12796, JP19H01720, JP20H01729, & JP20K20829).

International Society of the Learning Sciences

OptimizerSpace: A CSCL Tool for Search and Optimization

Arthur Hjorth, Ida Bang Hansen, Jacob Sherson
arthur@mgmt.au.dk, idabanghansen@gmail.com, sherson@mgmt.au.dk
Center for Hybrid Intelligence, Business and Social Sciences, Aarhus University, Denmark

Abstract: Optimization is important, both as a mathematical method, and as a general, descriptive and analytical perspective on human activity. However, taking a formal, mathematical approach to teaching optimization will leave out the majority of learners who do not have the necessary calculus training. We present a pilot study of OptimizerSpace, a CSCL-tool designed to elicit students' intuitive ideas and strategies for optimization, and to enable teachers to easily facilitate classroom discussions about them.

Keywords: cscl, optimization, formal education, pilot study

Introduction: Optimization and challenges

A large number of human activities can be thought of as searching for optimal solutions in a big landscape; foraging for food, or generally solving many-dimensional problems in which the result of a particular set of actions yield a more or less stable result (Wright, 1932). Consequently, optimization could be productively used as a framing device across curriculum in school settings. Optimization traditionally relies on calculus, but in our national curriculum, calculus is only taught in high school electives. We therefore need to think of alternative approaches in which we draw on students' intuitive ideas (Smith et al, 1994) and use non-formal representations (Sherin, 2000). As one such approach, we present a pilot study of OptimizerSpace, a computer-supported collaborative learning tool, that we designed to help teachers *elicit* students' intuitive ideas about optimization and *discuss* these ideas collectively in the classroom. In this poster, we show that being able to easily access students' different optimization approaches was useful for collaborative reflection and learning in the classroom.

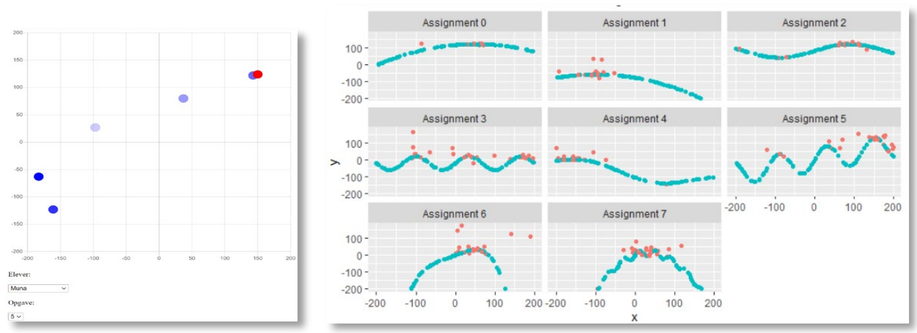


Figure 1. Left: OptimizerSpace Teacher View. Right: The eight functions that students worked with.

OptimizerSpace and pilot Study

The purpose of OptimizerSpace is to provide a shared reflection space for students to think about optimization. OptimizerSpace provides a simple data API that accepts a JSON object containing an activity ID, the name of the student(s) who submitted the datapoint, and an X/Y-value pair, or X/Y/Z-value triplet. OptimizerSpace then provides a teacher view [Figure 1, left] that lets the teacher quickly scroll through students' submissions. Students can work through assignments that relate to optimization, and the teacher can show and discuss how each student approached the optimization task. In this pilot, we tested OptimizerSpace in conjunction with learning activities built in Netlogo (Wilensky, 1999). We collaborated on designing the unit, spanning six hours divided into four lessons, with a math teacher with 10+ years of teaching experience. 30 students from a science-focused high school class, aged 16-18, took part in the pilot study.

OptimizerSpace: Eliciting, identifying and discussing intuitive strategies

To briefly contextualize, we want to mention that during lesson 1, we introduced students to optimization with a simple pen and paper activity. In Lesson 2, students were given a NetLogo model with 8 mathematical functions with one parameter in a Cartesian space [Figure 1, Right]. For each function, students were given five attempts in which they could click the Cartesian space and get a Y-value for the corresponding X-value, and in their sixth attempt they had to click where they thought the optimum of the function was located. Each student click was sent to OptimizerSpace. Students had not initially been introduced to formal optimization strategies like exploration or exploitation. However, we saw that students intuitively took a variety of approaches to optimizing functions that resemble these strategies. Below, we present an episode during a classroom discussion of students' solution to function 5 - a sine wave with four local optima and a global slope [Figure 1, right, 5].

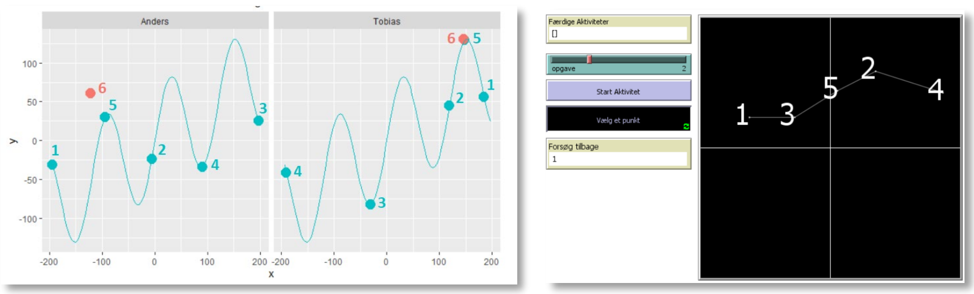


Figure 2. (Left) Anders' and Tobias' approaches to optimizing function 5. (Right) The NetLogo interface where students were optimizing the eight functions. Five "guesses" have been placed here.

As students had finished their assignments, the teacher used OptimizerSpace to scroll through their solutions to the function, and quickly found one that was successful and one that was not. OptimizerSpace shows the order of "guesses" and because of this, it is possible for students and the teacher to reconstruct what happened. In Figure 2 left, we visualize Anders' and Tobias' approaches to optimizing the function: They both used their first four guesses to *explore* the full space, and even relatively close to each other's guesses. In their second to last guess, they took different strategies: Anders *explored* the largest space (between his 1 and 2), whereas Tobias *exploited* that his 1 and 2 were the highest he had found and put his fifth guess between them. Consequently, this led Anders to incorrectly believe that the global maximum was near the first peak of the function, whereas Tobias gathered enough information to exploit it again, and get very close to the global maximum in his final guess. Anders insisted that he had just been unlucky, but the teacher used the contrasting cases of Anders and Tobias to facilitate a collaborative learning experience: By being able to show a concrete example of a students' optimization process, it was possible to discuss when to change from exploring to exploiting, and this in turn fed into a larger discussion throughout the unit on how optimization is a process of *making increasingly educated guesses*.

Conclusion and future work

OptimizerSpace allows a teacher to easily view and show students' solutions to optimization problems, and it helps students walk the classroom through their thinking and strategizing during classroom discussions. By facilitating the sharing of students' optimization approaches, OptimizerSpace creates a productive and collaborative reflection process and helps identify concrete examples of what is difficult about optimization. The simple interface (both visual and API) of OptimizerSpace makes it easy to work with, and the simple visualizations can engage students' naïve, yet productive intuitions about optimization. While this is very early work, we believe that our pilot showed that OptimizerSpace can enable collaborative classroom learning through sharing and visualizing students' strategies. In future work, we will use OptimizerSpace to explore learner strategies in more sophisticated optimization problems in both formal educational settings, and in museums and citizen science settings.

References

Sherin, B. L. (2000). Meta-representation: An introduction. *The Journal of Mathematical Behavior*, *19*(4), 385-398.

Smith III, J. P., Disessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The journal of the learning sciences*, *3*(2), 115-163.

Wilensky, U. "NetLogo: Center for Connected Learning and Computer-Based Modeling." *Northwestern University, Evanston, IL*, 1999.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature human behaviour*, *2*(12), 915-924.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proceedings of the Sixth Internatio nal Congress on Genetics, ed Jones DF (Brooklyn Botanic Garden, Menasha, WI), Vol 1, pp 355–366

The Interplay of Knowledge Construction and Regulation of Learning in CSCL Settings

Kateryna Zabolotna, Jonna Malmberg, Sanna Järvelä, Hanna Järvenoja
kateryna.zabolotna@oulu.fi, jonna.malmberg@oulu.fi, sanna.jarvela@oulu.fi, hanna.jarvenoja@oulu.fi
University of Oulu

Abstract: This study explores how different types of regulated learning (CoRL and SSRL) interplay with phases of knowledge construction in a CSCL task. Secondary school pupils (N=36) were videotaped while working in groups on a poster about a topic “Center of gravity”. Based on microlevel interaction video analysis, we present an example of how a group engages in CoRL and SSRL while co-constructing knowledge in CSCL and point out connections between regulated learning and knowledge construction.

Theoretical background

Collaborative learning (CL) refers to the process of working together, when students set shared goals, create plans how to achieve these goals together, negotiate throughout the whole process, and bear shared responsibility for the learning outcomes (Hmelo-Silver & Barrows, 2008). In other words, learners in groups find ways how to produce or co-construct knowledge about a certain phenomenon or a problem by reaching common understanding. They can engage in these complicated CL sequences by performing regulatory actions, i.e. take strategic control over thoughts, actions, and motivations, to achieve goals as individual learners and group members by regulating themselves, each other and the group as a whole (Järvelä & Hadwin, 2013). However, as regulation and knowledge construction occur through similar cognitive and metacognitive processes, in social settings it is still challenging to recognize how they are different from each other (Järvelä et al., 2013) and how they intertwine with each other. There is also a need to understand how groups apply regulatory strategies to construct shared knowledge and what regulatory mechanisms lead to shared understanding of task and ensure its completion (Lee et al., 2017).

Present study

In this study we attempt to characterize how collaborating groups engage in knowledge co-construction (KC) and co-regulated (CoRL) and socially shared regulated learning (SSRL). Thus, the research question is: *How do types of regulation and phases of knowledge construction occur and intertwine during computer-supported collaborative learning?* 34 secondary-school students were divided into 9 groups. Their task was to collaboratively create a joint poster on the science topic “centre of gravity”. The drawing task was carried out on an interactive table with the application of the SimSketch software at the university research facility. The task instructions and knowledge about the concepts related to the center of gravity were embedded in the software. The students were videotaped with a 360° camera and they wore a wireless microphone each.

Data analysis

For this paper we selected one group that was most actively engaged in KC activities. Microlevel interaction analysis for phases of KC was implemented by using Gunawardena et al.’s (1997) Interaction Analysis Model (IAM). The model defines *5 phases of KC*, i.e. sharing/comparing of information; discovery and exploration of dissonance or disagreement among ideas, understandings, statements; negotiation of meaning/co-construction of knowledge; testing, evaluating and modifying proposed synthesis or co-construction; and agreement statements/application of newly constructed meaning or solution.

Further, we coded the selected video for indicators of regulatory processes (task understanding, planning, monitoring and evaluating, strategy use) and regulation types (CoRL and SSRL) (Malmberg et al., 2017**).** Task execution was added as an additional category covering the executive processes (drawing, writing). The interrater reliability coding indicated high agreement between coders (for KC $\kappa = .82$; for RL $\kappa = .87$). At the final stage we descriptively analysed the connections between CoRL, SSRL and KC and graphically represent them below.

Results

The results indicate how students working in groups engage in CoRL, SSRL and KC and how these processes interrelate in CSCL. The selected group spent 45% of the total duration of the session on KC activities. The most frequently occurring KC phase was negotiating and co-constructing ($f = 11$), whereas the least frequently students engaged in sharing opinions/comparing information (f = 1). All 5 phases of knowledge construction appeared in the selected group. As for the regulatory activities, 22% of the whole time the students spent on CoRL (planning, monitoring and evaluating, strategy use) and 7% on SSRL (planning, monitoring and evaluating). SSRL occurred mostly in the beginning of the learning sessions, while CoRL was occurring at different moments, coinciding with a number of KC phases.

In Figure 1 we provide a visual example of how regulated learning and knowledge construction interrelate in our selected case.

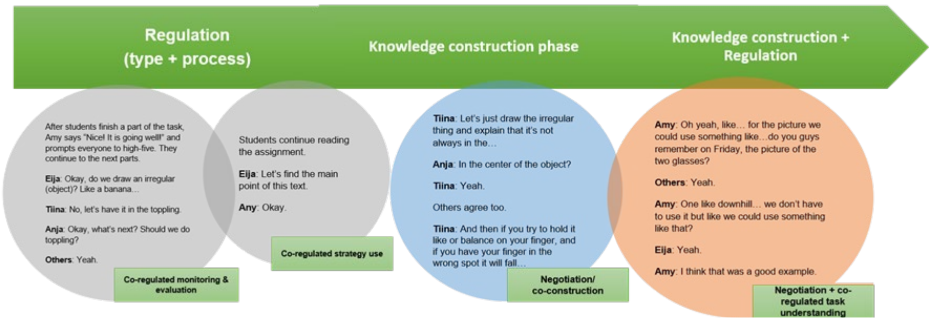


Figure 1. Occurrence of regulated learning and knowledge construction in a learning episode.

In this situation, co-regulated planning was followed by co-regulated strategy use (identifying the main points to use in the poster) which led to negotiation and co-construction of mutual understanding (KC Phase 3). At the same time, while negotiating mutual meaning, students were co-monitoring their task understanding (CoRL). In another situation, co-regulation of task understanding (students stating that they do not understand a phenomenon) led to and continued during negotiation phase (students were prompted to interpret and discuss task instructions, thus co-construct a mutual understanding of the phenomenon). Additionally, negotiation and co-construction was the only knowledge construction phase that happened simultaneously with CoRL or SSRL. The case example shows that the students' awareness of the need to understand a task (task understanding) may prompt and support active negotiation of a mutual understanding of the task (co-construct knowledge). However, more detailed analysis will allow us to take a deeper insight into these processes and may provide implications for how to differentiate between RL and KC and illustrate how and when they intertwine with each other.

References

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.

Hadwin, A. F., Järvelä, S., and Miller, M. (2017). Self-regulation, co-regulation and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of Self-Regulation of Learning and Performance*, (2nd ed.). NY: Routledge.

Hmelo-Silver, C. E., & Barrows, H. S. (2008). Facilitating collaborative knowledge building. *Cognition and Instruction*. https://doi.org/10.1080/07370000701798495

Järvelä, S., & Hadwin, A. F. (2013). New Frontiers: Regulating Learning in CSCL. *Educational Psychologist*, *48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006

Lee, L., Lajoie, S. P., Poitras, E. G., Nkangu, M., & Doleck, T. (2017). Co-regulation and knowledge construction in an online synchronous problem-based learning setting. *Education and Information Technologies*, 22(4), 1623–1650. https://doi.org/10.1007/s10639-016-9509-6.

Malmberg, J., Järvelä, S., & Järvenoja, H. (2017). Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemporary Educational Psychology*, *49*, 160–174. https://doi.org/10.1016/j.cedpsych.2017.01.009

Analyzing Peer Interaction as Asynchronous Online Professional Development Scales Up to Include More Teachers

Thomas Richman, Susan A. Yoon, Amin Marei, Katherine Miller
trichman@upenn.edu, yoonsa@upenn.edu, amarei@upenn.edu, kmmiller@upenn.edu
University of Pennsylvania

Abstract: This study examines how participant group size affected peer interactions in an asynchronous online professional development as it scaled from a small (8) to a larger (91) participant pool. Analysis of the discussion forums indicated that transactivity was higher in the larger iteration during discussions around classroom implementation. Analysis of participant interviews indicated that access to diverse perspectives on implementation was particularly important to participants, which may have supported collaboration in the larger iteration.

Introduction

With a growing number of teachers accessing professional development (PD) in asynchronous online contexts (e.g., Parsons et al., 2019), online PD designers must be able to provide participants with collaborative learning opportunities that play a critical role in knowledge-building processes and sustained improvement of teaching practices (Desimone & Garet, 2015). One design characteristic that warrants further attention is the size of the participant pool in an asynchronous online learning platform. While smaller group sizes (i.e., fewer than 10 participants) can support effective learner interaction (e.g., Akcaoglu & Lee, 2016), larger participant pools may alleviate concerns around the immediacy of responses and may make it easier for participants to find peers that share relevant classroom experiences (e.g., Frumin et al., 2018). In this study, we examined how the nature of peer interaction in discussion forums differed between two iterations of the same online PD for high school Biology teachers as it scaled from 8 to 91 participants. We ask the following questions: (1) how does peer discourse manifest differently in online discussion forums as asynchronous PD scales up to engage more teachers; and (2) what factors contribute to those differences?

Methods

Both iterations of our online PD were conducted on the edX platform and were designed to take participants approximately 40 hours over the course of 6 weeks. The first iteration (Iteration 1) was offered in July and August 2018 and consisted of eight teachers. The second iteration (Iteration 2) was offered in July and August 2019 and consisted of 91 teachers posting at least once in the discussion forum. Few alterations were made between the two iterations of the PD. More information about the PD and its participants can be found in Yoon et al. (2020).

In both iterations, participants were asked to participate in discussion forums with open-ended prompts that scaffolded discourse. These prompts were categorized as "implementation" (i.e., considering topics in reference to past or future classroom implementation), "content" (i.e., considering topics or ideas reflectively), or "collaboration" (e.g., "read and reply to a few of your peers' posts"). Both iterations contained 55 discussion board prompts. For Iteration 1, the prompts resulted in 694 coded utterances. For Iteration 2, the prompts resulted in 6138 coded utterances. Using a transactivity coding scheme to measure instances of peers operating on and interacting with each other's reasoning, discussion posts following each prompt were qualitatively coded from 1 (lowest levels of transactivity) to 5 (highest levels of transactivity). We conducted equal variances t-tests on each of the three prompt categories to determine if the nature of transactive discourse varied between Iteration 1 and 2. More information about the prompts and the transactivity coding procedure can be found in Yoon et al. (2020).

Interviews were conducted with eight Iteration 1 participants and ten Iteration 2 participants. These interviews used the same semi-structured interview protocol to probe participants about their overall PD experience and their interactions with peers. Interviews were qualitatively analyzed for comments that elucidate the ways that the larger number of participants in Iteration 2 may have impacted participants' overall ease of discussion and social presence, access to diverse and relevant perspectives on classroom implementation, and peer responsiveness relative to Iteration 1. A more comprehensive analysis and discussion of these interviews will be detailed in future publications.

Findings

Transactivity analysis

The average transactivity scores (with standard deviation) for Iterations 1 and 2 can be found in Table 1.

Table 1: Average transactivity score of each iteration by prompt-type (standard deviation in parentheses)

Prompt Type	Iteration 1	Iteration 2
Implementation	2.51 (0.86)	2.75 (1.01)
Content	2.66 (0.96)	2.66 (1.01)
Collaboration	2.91 (1.06)	2.99 (1.07)

The average transactivity scores for implementation prompts were significantly greater in Iteration 2 than in Iteration 1 ($t(2649) = 4.003$, $p < .0001$). However, we found no significant differences in transactivity for prompts categorized as content ($t(2162) = .081$, $p = .936$) or collaboration ($t(2025) = .876$, $p = .381$) between the two iterations. This indicates that there was significantly more transactive discourse between teachers following implementation prompts in Iteration 2 when compared with Iteration 1. No difference in transactivity was found between the iterations following content prompts or collaboration prompts. These findings were corroborated with findings in the interviews that are briefly discussed in the next section.

Analysis of interview transcripts

The majority of teachers interviewed across both iterations of the PD (14 of the 18 interviewees) discussed the importance of accessing contextually-relevant guidance for implementing the PD curriculum in their own unique classrooms. In seeking this guidance, 11 of the 18 interviewees mentioned that they turned to the discussion forum for support and insights on implementation from peers. As one teacher describes, "it was really neat to be able to hear how some of the teachers are implementing things in their classroom… that was, I thought, very impactful."

Despite the value that teachers placed on these discussions, the small size of the Iteration 1 cohort seemed to act as a barrier. According to five of the eight teachers in Iteration 1, the PD's small participant pool and asynchronous nature meant teachers were rarely working in the same forums simultaneously, making ease of discussion and peer responsiveness a challenge. Additionally, according to two participants, the limited number of perspectives in the PD made discussions of implementation difficult. One teacher explained. "the [student] age-level matters, the demographic matters, the learning levels matter. If I could have known [teachers working] in a similar context to me then maybe a relationship could have been built there."

For Iteration 2 teachers, there appeared to be greater success in accessing relevant perspectives on implementation in the discussion forum. In this iteration, seven of the ten interviewees described benefitting from peer discussion to better understand implementation concerns, with three of those seven describing these discussions as exceedingly valuable. As one teacher described, "It was very, very informative for me to see all the different perspectives [in the discussion forum] and how we could take the same lesson and play with the crosscutting concepts." Another teacher described how the sheer diversity of perspectives on implementation in Iteration 2 was valuable. According to her, "even if I was reading somebody who taught a younger grade level for instance, I could foresee some ways that I might scale that up to fit in my high school classes."

References

Akcaoglu, M., & Lee, E. (2016). Increasing Social Presence in Online Learning through Small Group Discussions. *The International Review of Research in Open and Distributed Learning, 17*(3).

Desimone, L., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. *Psychology, Society and Education, 7*(3), 252–263.

Frumin, K., Dede, C., Fischer, C., Foster, B., Lawrenz, F., Eisenkraft, A., Fishman, B., Levy, A.J., & McCoy, A. (2018). Adapting to large-scale changes in Advanced Placement Biology, Chemistry, and Physics: the impact of online teacher communities. *International Journal of Science Education, 40*(4), 397-420.

Parsons, S. A., Hutchison, A. C., Hall, L. A., Parsons, A. W., Ives, S. T., & Leggett, A. B. (2019). U.S. teachers' perceptions of online professional development. *Teaching and Teacher Education, 82*, 33–42.

Yoon, S. A., Miller, K., Richman, T., Wendel, D., Schoenfeld, I., Anderson, E., Shim, J., & Marei, A. (2020). A social capital design for delivering online asynchronous professional development in a MOOC course for science teachers. *Information and Learning Sciences, 121*(7/8), 677-693.

Acknowledgements

This research was funded by a grant from the U.S. National Science Foundation, #1721003. We would also like to thank our collaborators on the design and development of the PD course who include Daniel Wendel, Ilana Schoenfeld, Emma Anderson, Jooeun Shim, and Eric Klopfer.

Synergies Between Humans and Machines to Support the Orchestration of CSCL Scripts at Different Scales

Ishari Amarasinghe, Davinia Hernández-Leo
ishari.amarasinghe upf.edu, davinia.hernandez-leo@upf.edu
ICT Department, Universitat Pompeu Fabra, Barcelona, Spain

Abstract: This study presents the orchestration challenges associated with scripted collaborative learning situations at different scales and how different Learning Analytics (LA) interventions may facilitate to address those issues. The proposed LA interventions were characterised as machine-in-control, human-in-control and hybrid approaches given different agents in charge of orchestration actions. A framing of the proposed LA interventions is presented considering also the different scales within which those interventions were deployed, in an attempt to seek the balance between different types of interventions.

Keywords: Computer-Supported Collaborative Learning, Orchestration, Learning Analytics.

Introdcution

The notion of orchestration captures the complexity associated with the real-time management of educational scenarios seeking effective learning (Dillenbourg, 2013). Data collected from online learning platforms can be analysed using different Learning Analytics (LA) techniques to support and improve orchestration. On the one hand, machine-oriented LA interventions such as adaptive group formation strategies that tailor group formation according to students' profiles or intelligent (conversational) agent techniques that support peer interactions are expected to assist the orchestration of collaboration automatically. On the other hand, LA tools in the form of teacher-facing LA dashboards may support teachers' orchestration actions. In a middle space, humans and machines can inform each others' actions hence taking the advantage of complementary strengths of both ends (Holstein, Aleven, & Rummel, 2020). The challenges associated with scripted Computer-Supported Collaborative Learning (CSCL) deployed at different scales and the degree of human/machine control for effective orchestration are yet to be explored. To this end, in this poster, we present orchestration challenges identified with respect to Pyramid pattern based CSCL scripts (Manathunga & Hernández-Leo, 2018) at different scales and a design space framing of LA interventions as human-in-control to machine-in-control in nature, given the feasibility and regulation needs of the learning contexts under investigation.

Framing human and machine support to orchestrate collaboration

Deployment of CSCL activities in a MOOC using PyramidApp (Manathunga & Hernández-Leo, 2018) revealed that sustained student participation in multiple phases of the script was a primary challenge. The uncertainty associated with learners' continuous participation along the consecutive Pyramid script phases undermined the pedagogical benefits of the Pyramid pattern. Moreover, the choice of script design parameters, e.g., activity duration, require adaptive modification according to participation levels (Amarasinghe, Hernández-Leo, Manathunga, & Jonsson, 2018). In the classroom context, the findings of the sessions conducted with teachers, in addition to the knowledge acquired through the literature review, revealed teachers' desire for tools that augment their actionability, which informed the design decisions of a teacher-facing LA dashboard (Amarasinghe, Hernández-Leo, Michos, & Vujovic, 2020).

In the distance context at a large scale, due to the nature of activity distribution in time and lack of continuous instructor involvement, we designed an automatic LA-based orchestration intervention agent that implements different intervention strategies adapting to the activity participation differences of students. The proposed interventions were automatic in nature and can be characterised as machine-in-control. In the classroom learning context, LA interventions in the form of teacher-facing LA dashboards were implemented to support teachers in regulating collaboration. The dashboard implemented two different types of support. In mirroring support, the interpretation of information and use of dashboard controls were decided by teachers without additional guidance, whereas in guiding support teachers were guided to take action via an alert mechanism that flagged critical moments in collaboration. The mirroring support thus scaffolds human-in-control sense-making and orchestration actions, whereas in guiding support, automatic machine-generated alerts suggest orchestration actions and offload teachers' decision-making responsibilities to some extent, all the while amplifying their actionability and respecting their agency (Soller, Martínez-Monés, Jermann, & Muehlenbrock, 2005). This can be characterised as a hybrid human-machine approach. Another LA intervention, which formulates adaptive

 © ISLS

groups using inputs from prediction algorithms and incorporates them into the Pyramid activity flow, has also been proposed and evaluated in both classroom and MOOC settings (Amarasinghe, Hernández-Leo, & Jonsson, 2019). It was important for this group formation policy to be implemented in both small-scale and large-scale situations for minimising the number of non-participating groups which would deter collaboration and break the continuous flow of learning. This intervention can be positioned under machine-in-control, as it automatically generates group formation policies based on predictions. Figure 1 shows an overview of the positioning of human-in-control and machine-in-control LA interventions in a design space that consider the orchestration challenges associated with CSCL activities deployed at different scales.

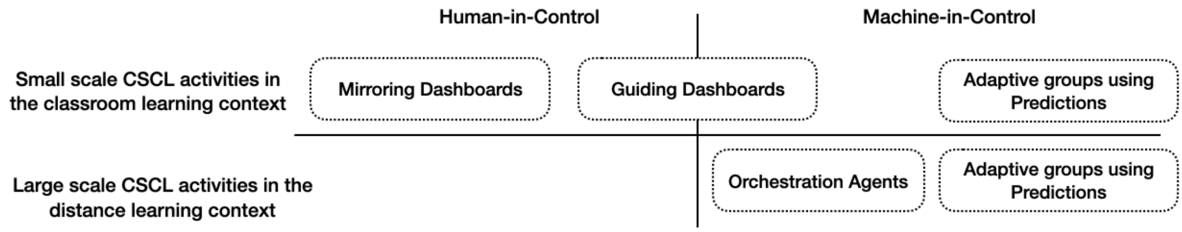


Figure 1. Positioning of different LA interventions to support orchestration at different scales.

Conclusions and future work

As presented above, the orchestration challenges related to CSCL activities deployed at different scales are different and it is possible for different agents to be in control of orchestration. In the classroom-learning context, teachers can be supported with dashboards ('human-in-control'), whereas in MOOCs, intelligent agents may take over collaboration regulation ('machine-in-control'). However, in the middle space, these two extremes (human-in-control and machine-in-control) can benefit the complementary strengths resulting in a hybrid approach (as illustrated using the guiding dashboards) that spans across a broad design space, that requires further exploration (Holstein et al., 2020). In the future, we are interested in exploring further the added values of the proposed human-in-control and machine-in-control approaches in terms of several evaluation metrics (such as improved teaching and learning, human agency, orchestration load, ethical aspects) to understand how to balance the human and machine support for orchestration at different scales.

References

Amarasinghe, I., Hernández-Leo, D., Manathunga, K., & Jonsson, A. (2018). Sustaining continuous collaborative learning flows in MOOCs: Orchestration agent approach. *Journal of Universal Computer Science, 24*(8), 1034–1051.

Amarasinghe, I., Hernández-Leo, D., Michos, K., & Vujovic, M. (2020). An Actionable Orchestration Dashboard to Enhance Collaboration in the Classroom. *IEEE Transactions on Learning Technologies, 13*(4), 662–675.

Amarasinghe, I., Hernández-Leo, D., & Jonsson, A. (2019). Data Informed design parameters for adaptive collaborative scripting in across spaces learning situations. *User Modeling and User-Adapted Interaction, 29*(4), 869–892.

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education, 69*(1), 485–492.

Holstein, K., Aleven, V., & Rummel, N. (2020). A Conceptual Framework for Human–AI Hybrid Adaptivity in Education. *Proceedings of the International Conference on Artificial Intelligence in Education (AIED 2020)* (pp. 240–254).

Manathunga, K., & Hernández-Leo, D. (2018). Authoring and enactment of mobile pyramid-based collaborative learning activities. *British Journal of Educational Technology, 49*(2), 262–275

Soller, A., Martínez-Monés, A., Jermann, P., & Muehlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal on Artificial Intelligence in Education, 15*(4), 261–290

Acknowledgments

This work has been partially funded by FEDER, the National Research Agency of the Spanish Ministry of Science, Innovations and Universities MDM-2015-0502, TIN2017-85179-C3-3-R. Davinia Hernández-Leo acknowledges the support by ICREA under the ICREA Academia programme.

International Society of the Learning Sciences

Transition, Lamination, and Personification: Affordances for Teacher Learning in a Mixed Reality Simulation

Jamie Gravell, Anne Garrison Wilhelm, Ann Marie Wernick
jamiedaragravell@gmail.com, awilhelm@smu.edu, awernick@smu.edu
Southern Methodist University

Abstract: Collaborative approaches to teacher professional development have the potential to support teacher learning. We studied the affordances of a Mixed-Reality Simulation of a mathematics discussion with a "notice and wonder" protocol in supporting teacher development in a collaborative setting. Qualitative analysis of 8 in-service teachers' participation in the rehearsals revealed how teacher sensemaking was supported by conversations during transitions between sessions as well as the lamination of different semiotic resources and personification of avatars within those conversations.

Introduction

Recent scholarship encourages a turn to the learning of adults, particularly teachers, as a necessary pre- and co-requisite for facilitating learning with children (e.g., Kawasaki & Sandoval, 2019). This paper identifies the affordances of a Mixed-Reality Simulation (MRS) for adult learners building pedagogical expertise during episodes of pedagogical reasoning (EPR). We can uncover the building of expertise in cooperative contexts by identifying the different semiotic and physical resources that participants bring to interactions, and how the lamination of those resources constructs the learning in interaction (Goodwin, 2013). In the case of this study, participants are bringing together semiotic resources from their own classrooms, their historical experience as a learner, previous assignments and discussions in the course, and their current experience in the MRS. These resources are 'laminated' or built upon one another to create a set of layers that are organized in relation to one another to construct new communication and meaning making. Uncovering those layers is useful for understanding how new knowledge is built in collaboration.

The MRS that we utilized involves a virtual classroom interface with student avatars controlled by a trained "interactor" (Dieker et al., 2014). This type of simulation environment offers a unique opportunity for a group of teachers to engage in teaching rehearsals without having to assume the role of students and allows for multiple tries at engaging in the same practice or set of practices (Cohen et al., 2020). In this paper, we build on these different bodies of work to study teachers as learners in collaborative settings, and ask: What are the affordances of a MRS along with a 'notice and wonder' protocol used to provide opportunities for teacher learning? Rehearsing teachers and observing teachers build pedagogical content knowledge together during the EPR provided by the instructional hand-off between the pair of teachers leading instruction and the whole-class debrief following the rehearsal.

Methods

The setting for this project was a masters-level mathematics education course for in-service teachers. Our analysis focuses on 8 teachers who were enrolled in the course and attended this session of the simulation lab where they were able to interact with five avatars (Figure 1) on a TV screen at the front of the classroom (see Figure 2). Teachers served as both observers and members of two-person teams instructing the avatars while the rest of the group watched from the side. Participants include novice to veteran teachers across contexts: primary grades, Montessori, bilingual math and science teachers, and one high school AP Statistics teacher.

Figure 1. *Mursion Avatars*

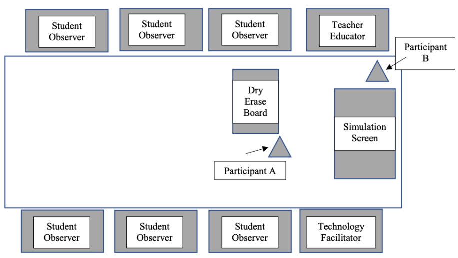


Figure 2. *Bird's eye view of simulation lab*

The course focused specifically on one instructional activity, Number Talks (Parrish, 2011). During expected pauses, the teachers utilized a "notice and wonder" protocol to identify what stood out to them in the MRS they just watched (notice) and what they might have more questions or concerns about (wonder) (van Es, 2011) to think together about next steps. Data for this study includes both video footage of the learning lab and photos of teachers' writing from the dry erase board used during the simulation. In accordance with facilitating a close investigation into how teachers co-construct knowledge with one another and their teacher educator (TE), we focus on the conversations between the rehearsing teacher team, observers and TE during transition points in the enactment of a number talk. Video data was content logged deductively through identification of salient episodes of content discourse defined as statements or questions that furthered classroom discourse or metacognitive reflection.

Findings

Based on our work-in-progress analysis of 4 teams of 2 teachers, we identify three major themes in the data related to the affordances of the MRS environment: (1) transitions and debriefs among participants during the simulation facilitate EPR; (2) participants are laminating semiotic resources across time and space onto the shared experience of the MRS; and (3) the personification of avatars by teacher participants supports that lamination across time and space. By engaging in the MRS combined with a "notice and wonder" protocol during transitions, teachers surface EPR focused on "collective interpretation linked to future work" (Horn et al., 2017) laminating their prior personal classroom experience *and* a shared experience with the simulation to build rich pedagogical content knowledge in the moment. This lamination, or layering of resources that results in a new entity, was supported by the shared MRS experience and authentic to the everyday work of teachers. The validity of the experience is evidenced by the personification of the MRS avatars by the participant teachers—contrary to their expectations, the teachers were readily able to suspend reality and interact with the avatars as if they were real children in many of the ways that matter for pedagogical reasoning.

Conclusion

The mixed-reality simulation and the "notice and wonder" protocol not only provide utility for learning the instructional activity of number talks and probing for student understanding, but also provide space for EPR within a shared experience which allows for a supported inquiry stance integral to professional learning (Opfer & Pedder, 2011; Russell et al., 2020). We are interested in three avenues for additional analysis of this data: (1) conducting a more detailed interaction analysis to trace language uptake and topic distillation across participants; (2) comparing this lab session to another example that was completed completely online through video conferencing software to see what variation the location of collaborative sensemaking has on the process; and (3) comparing the teacher sensemaking in the mixed-reality environment to that of the same students when they use the "notice and wonder "protocol asynchronously on video recordings of each other's live classrooms.

References

Cohen, J., Wong, V., Krishnamachari, A., & Berlin, R. (2020). Teacher coaching in a simulated environment. *Educational Evaluation and Policy Analysis*, *42*(2), 208–231.

Dieker, L. A., Rodriguez, J. A., Lignugaris/Kraft, B., Hynes, M. C., & Hughes, C. E. (2014). The potential of simulated environments in teacher education: Current and future possibilities. *Teacher Education and Special Education*, *37*(1), 21–33.

Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics*, *46*(1), 8–23. https://doi.org/10.1016/j.pragma.2012.09.003

Horn, I. S., Garner, B., Kane, B. D., & Brasel, J. (2017). A Taxonomy of Instructional Learning Opportunities in Teachers' Workgroup Conversations. *Journal of Teacher Education*, *68*(1), 41–54. https://doi.org/10.1177/0022487116676315

Opfer, V. D., & Pedder, D. (2011). Conceptualizing Teacher Professional Learning. *Review of Educational Research*, *81*(3), 376–407. https://doi.org/10.3102/0034654311413609

Parrish, S. D. (2011). Number talks build numerical reasoning. *Teaching Children's Mathematics*, *18*(3), 198–206.

Prior, P., & Shipka, J. (2003). Chronotopic lamination: Tracing the contours of literate activity. *Writing Selves, Writing Societies: Research from Activity Perspectives*, 180–238.

Russell, J. L., Correnti, R., Stein, M. K., Bill, V., Hannan, M., Schwartz, N., Booker, L. N., Pratt, N. R., & Matthis, C. (2020). Learning From Adaptation to Support Instructional Improvement at Scale: Understanding Coach Adaptation in the TN Mathematics Coaching Project. *American Educational Research Journal*, *57*(1), 148–187. https://doi.org/10.3102/0002831219854050

Please Introduce Yourself: Exploring Student Identity in Academic Online Spaces

Preeti Raman, Dina Soliman, Jim Hewitt
preeti.raman@mail.utoronto.ca, dinaa.soliman@utoronto.ca, jim.hewitt@utoronto.ca
University of Toronto

Abstract: The purpose of this research is to understand the ways in which new students identify themselves in online social spaces. More specifically, we performed a linguistic analysis of graduate students' introductory notes in an asynchronous discussion forum. Students described themselves in terms of their epistemic beliefs, qualifications, professional experiences, motivations, career plans and cultural background. In building their identity in the community, students brought out their individuality, justified their presence and established social capital.

Introduction and objectives

As Universities around the world move their courses online, students are increasingly being asked to 'get to know one another' in online social spaces. Such asynchronous discussion-based spaces provide a platform for students to find resources, support one another and build community (McInnerney & Roberts, 2004). Additionally, they become a place for students to identify themselves to their peers and build social capital. Self-identity carved out in these spaces can be subjective (how we think of ourselves), representative (what is depicted), or self-presentation (how we present ourselves to others) (Marwick, 2013). One way of understanding such self-presentation is through the information people choose to share with others in their first introductions in a community. Discussion style introductory notes are schema-less, thus providing for a free slate for expression.

The academic online context makes these student self-introductions different than previously studied social media narrations or identities shared in classroom settings. Our study seeks to examine students' self-presentation of their identity through a systematic qualitative analysis of student notes in the 'Introductions' section of an online academic social space. This is an important first step towards building scholarship around the construct of academic online identity and how the process of identity development subsequently influences community building and learning.

Perspectives and theoretical framework

This study is framed within social constructivism and situated within growing conversations around placing students in the center of their learning within a learning community (Vygotsky, 1978). Sharing identities is an important element in building a sense of community and being accepted by one's peers. This acceptance encourages students to share their ideas and learning needs without being embarrassed and to receive support from their peers (Brown & Campione, 1994; Scardamalia, 2003). Moreover, students engage better with both their peers and their learning when there is identity congruence (Hughes, 2007). Thus, we seek to find ways to affirm students' self-identities in a way that allows for the development of everyone's full potential.

Methods and data sources

This is a case study of a class of incoming graduate students who participated in a department-wide online social space at a large public Canadian university when their programs were moved fully online during the Covid-19 pandemic. Students engaged in various discussion threads that provided useful information for new students alongside social activities. All 190 new students were encouraged to participate in this web space by introducing themselves in the discussion forum. Data were collected from the posted student notes in this forum and then anonymized prior to analysis. Central codes for the analysis were informed by literature on academic social identity (deductive); however, our analysis was also open to a variety of emerging themes that occurred during the content analysis (inductive). 63 student introduction notes were coded by two researchers independently with an inter-rater agreement of 89.932%.

Results and discussion

Across the introductory notes of 63 students, 330 separate codes were analyzed. Greetings and repeated sentences were ignored. The nine themes that emerged from the content analysis are summarized in Table 1 with examples from student writings. All notes contained multiple and themes repeated themselves multiple times in the same note. Frequency counts from the thematic analysis are summarized in Table 2.

Table **1**: Themes identified in introductory notes

Theme	**Code**	**Examples**
Qualifications	Q	*"I have a BSc in Mathematics and Statistics, BEd in Mathematics, and MEd in..."*
Professional Experience	X	*"I currently teach high school English and have been teaching for the past 10 years."*
Motivation (academic)	M	*"I'm interested in understanding how teachers navigate the moral terrain of their…"*
Interests (non-academic)	I	*"I like to run, cook, bake, play music, and just generally be outside"*
Social Presence	S	*"I look forward to meeting you all and discuss our endless possibilities."*
Cultural	C	*"I'm from Shanghai and I lived in Ottawa before coming to Toronto for university."*
Future/Career	F	*"My plan is to be an entrepreneur in Education"*
Epistemics	E	*"...let learning and teaching praxis be informed by students..."*
Other	O	*"I'm also a parent of two young boys…"*

The purpose of this study was to examine the ways in which graduate students self-identify in online contexts. The research question asked what elements of self-identity are observed in student introduction notes. As can be seen from Table 2, 94% of students tried to establish a social connection. Notes also mostly started (29%) and ended (67%) with a social theme. Students mostly spoke about their past professional experiences (92%) and their qualifications (84%), perhaps to justify their presence and foster academic belonging. Only three students spoke about their epistemic beliefs.

Table 2: Results of content analysis on all introductory notes

	Q	X	M	I	S	C	F	E	O
Notes that include this theme	53 (84%)	58 (92%)	52 (82%)	45 (71%)	59 (94%)	34 (54%)	7 (11%)	3 (5%)	19 (30%)
Notes that begin with this theme	21 (33%)	9 (14%)	1 (2%)	0 (0%)	18 (29%)	12 (19%)	0 (0%)	1 (2%)	2 (2%)
Notes that that end with this theme	2 (3%)	0 (0%)	1 (2%)	14 (22%)	42 (67%)	1 (2%)	1 (2%)	0 (0%)	2 (3%)

Surprisingly, we did not find any notes where students directly stated their personality traits such as extroversion or openness. We are furthering this work with student interviews to better understand how student identities grow as their learning progresses.

Scholarly significance

The purpose of this research is to identify how students introduced themselves in an online community. Our results suggest that students' efforts typically involved: 1) An attempt to legitimize one's involvement in the community by describing one's academic and/or professional credentials; 2) A statement of one's personal learning goals and objectives, which serves to legitimizes one's involvement in a learning community and open up possibilities of forging social connections with others in the community that share similar interests; 3) An attempt to humanize oneself by sharing personal interests; and 4) An announcement of a willingness to engage with community members. In sum, students' initial efforts to establish an identity within a learning community involve discursive moves to establish a foundational identity that sets the stage for future involvement.

References

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* (p. 229–270). The MIT Press.

Hughes, G. (2007). Diversity, identity and belonging in e-learning communities: Some theories and paradoxes. *Teaching in Higher Education*, 12(5-6), 709-720.

Marwick, A. E. (2013). *Online identity. A companion to new media dynamics*, 355-364.

McInnerney, J. M., & Roberts, T. S. (2004). Online learning: Social interaction and the creation of a sense of community. *Journal of Educational Technology & Society*, 7(3), 73-81.

Scardamalia, M. (2003). Knowledge building environments: Extending the limits of the possible in education and knowledge work. In A. DiStefano, K. E. Rudestam, & R. Silverman (Eds.), *Encyclopaedia of distributed learning* (pp. 269–272). Thousand Oaks, CA: Sage.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Distributed Interactions During "Hands-On" Labs with Paraeducators

Seth Van Doren, BSCS Science Learning, svandoren@bscs.org
Lisa Hardy, Concord Consortium, lhardy@concord.org
Colin Dixon, BSCS Science Learning, cdixon@bscs.org
Sherry Hsi, BSCS Science Learning, shsi@bscs.org

Abstract: We highlight potential areas of interest surrounding the dynamics of student-paraeducator collaborations, a type of collaboration that has largely been overlooked and could well inform future computer-supported collaborative learning (CSCL) research. We observed that engagement with materials was necessarily distributed and that this made implicit material practices in science activity explicit. We also observed that leadership in the group was emergent and distributed amongst both students and paraeducators.

Introduction

Current and projected shortages of fully credentialed special education teachers nationwide have forced many school districts to turn to paraeducators to provide the personalized support students need. Paraeducators often serve not only as teacher aides, but language interpreters and helping hands in material activity (Daniels & McBride 2001). These circumstances have made student-paraeducator groups a common arrangement, and these collaborations will become more widespread as employment of paraeducators is projected to grow over the next decade (Bureau of Labor Statistics). Group composition has implications on the nature of a group. Broadly, the current focus is designated on understanding student-student, student-teacher, and peer-peer collaboration with little to no research on student-paraeducator groups. CSCL spaces themselves can promote equity by redistributing participation and promoting engagement from underrepresented learners (Ramey & Stevens 2019) and when designing equitable learning spaces, it will be important to better understand student-paraeducator collaboration. This study unpacks student-paraeducator interaction and highlights potential areas of interest within student-paraeducator collaboration, particularly distributed interaction with materials and emergent leadership.

Methods

This study was conducted in a 9th grade biology classroom in a San Francisco Bay Area public high school with a predominantly Black, Latinx, and Asian student body. The focal group being highlighted here is composed of three students on individualized education plans and three paraeducators who worked together during a three week design-based CSCL biology curriculum. During the curriculum, learners used Internet of Things (IoT) enabled hardware and software to design and perform authentic investigations related to photosynthesis and cellular respiration in plants. All activities throughout the curriculum are meant to be collaborative, computer and hardware availability necessitated that learners work in small groups.

We collected video and/or screencast data from ten class periods during the curriculum. Activity within the focal group of three students and three paraeducators was captured via video for 531 minutes and via screencast for 311 minutes. Screencast data was summarized and logged and video data was thematically coded using MAXQDA software. Analysis of emergent leadership was informed by examining the groups problem solving discourse (Li et al. 2007, Sun et al. 2017).

Results: Distributing interactions with materials and leadership

Interactions with materials were necessarily distributed amongst members of the group. Two of the three students in the group were physically disabled in a way that inhibited them from directly interacting with hardware and other physical materials needed to complete investigations, meaning that if one of these students wanted to see a change to tangible components of their experimental setup they needed to communicate and cooperate with a paraeducator. None of the students were able to regularly record answers to the questions in their journals by hand, so students and paraeducators needed to discuss what would be recorded and submitted for credit. Similarly, the paraeducators understood that the investigations were not theirs alone, and they couldn't make large decisions about the experiment or journal answers without assent from students. No one person was able to make all decisions about materials, or was responsible for answers in a journal. This necessary cooperation forced some implicit material practices of science to become explicit.

Analysis of classroom video data revealed that leadership within the group was emergent and distributed amongst three key members, one student and two paraeducators. During the first class period of the curriculum, a discussion-focused day where no IoT equipment was used, paraeducators primarily acted as aides and worked with their individual students. Once lab activities began, associations between paraeducators and specific students dissolved and the group worked together. We found that a student, Bernice, assumed leadership when the group was planning investigations and running an investigation. A paraeducator, Ms. Ana, assumed leadership when the group needed to orient toward their lab handouts and answer questions in order to receive credit for the day. Another paraeducator, Ms. Camilla assumed leadership over materials, ensuring that they were properly set up and delegated tasks toward Harrison, a student who was able to manually interact with materials. Bernice and Ms. Camilla co-lead discussions about biology phenomena, and each represented the group at different times when they were asked questions by the teacher or members of the research team.

One event that exemplifies the group's distributed dynamic occurred while the group was running an investigation to learn how light conditions influence carbon dioxide gas concentration in a sealed container full of spinach. Bernice directed Ms. Camilla on how high to hold a lamp above their lab setup. Bernice watched the stream of data they were being produced, and asked Ms. Camilla to raise or lower the lamp if the light data started to move away from their desired value. Concurrently, Ms. Ana would ask both Bernice and Ms. Camila what was occurring biologically, and what to write down in the lab journal.

Conclusion and implications

The study highlights two potential areas of interest for better understanding student-paraeducator collaborations in CSCL settings. Distributed interactions with materials and necessary cooperation provided ample opportunities for the group to distribute leadership duties and make implicit science practices explicit. Through their interactions with each other, not only did students build knowledge and engage in more authentic science practices, but the paraeducators did as well. Further research on this group composition will allow us to design more equitable learning spaces and better understand the implicit material practices of learners in CSCL settings. Furthermore, we believe that forming a better understanding of this and other collaborations between students and non-credentialed adults may be beneficial for better designing learning spaces within remote and pod learning contexts.

References

Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook, Teacher Assistants, at https://www.bls.gov/ooh/education-training-and-library/teacher-assistants.htm (visited *October 16, 2020*).

Daniels, V. I., & McBride, A. (2001). Paraeducators as critical team members: Redefining roles and responsibilities. *Nassp Bulletin*, *85*(623), 66-74.

Li, Y., Anderson, R. C., Nguyen-Jahiel, K., Dong, T., Archodidou, A., Kim, I. H., ... & Miller, B. (2007). Emergent leadership in children's discussion groups. *Cognition and Instruction*, 25(1), 1-2.

Ramey, K. & Stevens, R. (2019). Girls as Experts, Helpers, Organizers, and Leaders: Designing for Equitable Access and Participation in CSCL Environments. In Lund, K., Niccolai, G. P., Lavoué, E., Hmelo-Silver, C., Gweon, G., & Baker, M. (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019*, *Volume 1 (368-375).* Lyon, France: International Society of the Learning Sciences.

Sun, J., Anderson, R. C., Perry, M., & Lin, T. J. (2017). Emergent leadership in children's cooperative problem solving groups. *Cognition and Instruction*, *35*(3), 212-235.

International Society of
the Learning Sciences

Towards Estimating Classroom Orchestration Load using Physiological and Self-Perception Measures

Ishari Amarasinghe, Milica Vujovic, Davinia Hernández-Leo
ishari.amarasinghe@upf.edu, milica.vujovic@upf.edu, davinia.hernandez-leo@upf.edu
ICT Department, Universitat Pompeu Fabra, Barcelona, Spain

Abstract: This poster presents the exploration of a method to estimate the notion of orchestration load using physiological measures in triangulation with self-perception measures in the classroom computer-supported collaborative learning (CSCL) context. Details of a pilot study conducted in which a teacher orchestrated CSCL activities under different supporting conditions are presented. Different facets of the orchestration load were disentangled in light of the study findings.

Keywords: Orchestration Load, Physiological measures, Classroom Collaboration.

Introduction

The notion of orchestration load has been described as the total effort teachers need to put in when using a certain technology for the real-time management of classroom activities (Prieto, Sharma, & Dillenbourg, 2015). Most of the existing studies refer to this notion as a high-level concept without disentangling its multi-faceted elements (Prieto, Sharma, Kidzinski, & Dillenbourg, 2018), due to its complex nature and a lack of robust measures available to assess them in real-time. In this study, we examined how novel technologies, i.e., Electrodermal Activity (EDA), could be used in triangulation with self-perceptions of the teachers (collected using post-activity questionnaires and stimulated-recall interviews) to deconstruct the notion of orchestration load.

Methods

A female teacher from a Spanish University conducted scripted CSCL activities using PyramidApp (Manathunga & Hernández-Leo, 2018). Teachers' orchestration actions under different support provisions namely, no dashboard, mirroring and guiding dashboard conditions were recorded. In the no dashboard condition, the teacher did not have access to a teacher-facing dashboard. In the mirroring condition teacher had access to a teacher-facing dashboard, however interpretation of information and use of dashboard controls were left to be decided by the teacher. In the guiding condition additional guidance, in terms of warnings were generated automatically to upfront critical moments. The teacher conducted nine sessions in the three conditions, having used the PyramidApp also extensively in the past without the support of the LA dashboard. To avoid a novelty effect, data was gathered in the three latest CSCL sessions reflecting the three conditions. Each activity lasted around nine minutes. During the sessions, the teacher was equipped with a wearable EDA sensor to measure and compare affective state under different supporting conditions. By visual inspection of the signal, frequency of peaks were taken into consideration and triangulated with the teacher's self-perception measures. The additional information regarding baseline data collection and calibration of the sensor is described in detail in our previous work (Amarasinghe, Vujovic, & Hernández-Leo, 2020).

Results

Figure 1 shows the graphs that were plotted using the EDA data collected. As shown in Figure 1(a), in the no dashboard condition, the presence of peaks in graphs implies changes in the affective state of the teacher. The teacher's affective state is changing as a reaction to the activity. A visual inspection of signal change indicates that there are some differences between the three conditions. For instance, in the no dashboard condition the signal shows an increase in the number of peaks and skin conductivity towards the end of the activity. In the mirroring condition [see Figure 1(b)], the signal implies that the physiological state was not constant during the whole activity. According to the peaks, teachers' physiological state changes over time, where less arousal can be noticed towards the end. Also, this physiological response declines towards the end of the activity. In the guiding condition [see Figure 1(c)] the signal was more constant and showed that there was physiological response (according to the peaks), but that state remained more-less constant during the whole activity.

In the no dashboard condition, the teacher was frustrated and felt discomfort: "Very difficult to obtain the whole picture. I was stressed regarding the planned time as some students were taking more time and frustrated for not having the means to control the script progressions." We infer that the EDA signal shows arousal which could be related to frustration that increased towards the end of the activity. In the mirroring condition, the teacher expressed that thinking and making decisions to take orchestration actions became demanding in real-time: “I am more relaxed when I use the dashboard and I can monitor the progression of the activity, but thinking and decision making was somewhat demanding.” However, towards the end of the activity, the physiological response declines which means less arousal, and the teacher mentioned that she felt more in control of the activity and became calmer over time. In the guiding condition, the teacher felt comfortable and was in control due to the automatic guidance: “I really felt I was in control, alerts were very helpful, I could relax and read on student’s submissions, discussions, etc.” We infer that this state remained more-less constant during the whole activity.

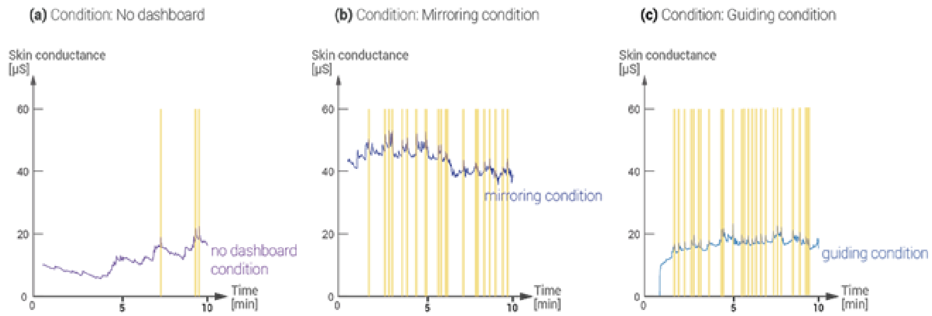


Figure 1. EDA signal (peaks) in three conditions.

Conclusions and future work

This paper presents preliminary work showing that EDA physiological measurements can be tentatively explained in terms of different facets of the orchestration load, i.e., goal formation, situation evaluation and action-taking, through its triangulation with the subjective reflections by the teacher. The obtained data shows that there are differences across the three conditions. For instance, the teacher was less comfortable in the mirroring dashboard condition as the teacher has to formulate goals, evaluate the learning situation and take actions in real-time without additional support-which may add to the orchestration load. However, in the guiding condition, the additional guidance provided using warnings may have supported the teacher in goal formation, situation evaluation and action taking hence resulting in a low orchestration load which created a much more comfortable situation. These results are interesting but should be taken with caution given the limited data set analyzed (only one teacher). In the future, we plan to extend the analysis of EDA and self-perception measurements (also with a bigger sample of teachers) with additional information such as a pre-survey about teachers’ contexts (e.g., activities completed before the data collection) and video recordings of orchestration actions.

References

Amarasinghe, I., Vujovic, M., & Hernández-Leo, D. (2020). Towards teacher orchestration load-aware teacher-facing dashboards. *Proceedings of Cross-MMLA in practice: Collecting, annotating and analyzing multimodal data across spaces co-located with 10th International Learning Analytics and Knowledge Conference* (Vol. 49 2610, pp.7–10). CEUR-WS.org. Available: http://ceur-ws.org/Vol-2610/paper2.pdf

Manathunga, K., & Hernández-Leo, D. (2018). Authoring and enactment of mobile pyramid-based collaborative learning activities. *British Journal of Educational Technology, 49*(2), 262–275.

Prieto, L. P., Sharma, K., & Dillenbourg, P. (2015). Studying teacher orchestration load in technology enhanced classrooms. *Design for teaching and learning in a networked world* (pp. 268–281). International Publishing: Springer.

Prieto, L. P., Sharma, K., Kidzinski, Ł., & Dillenbourg, P. (2018). Orchestration load indicators and patterns: In-the-wild studies using mobile eye-tracking. *IEEE Transactions on Learning Technologies*, *11*(2), 216–229.

Acknowledgments

This work has been partially funded by FEDER, the National Research Agency of the Spanish Ministry of Science, Innovations and Universities MDM-2015-0502, TIN2017-85179-C3-3-R. Davinia Hernández-Leo acknowledges the support by ICREA under the ICREA Academia programme.

PK–12 Computing Teacher Interactions in an Online Professional Learning Experience

Paulina Haduong, Carter Zenke, Jess Eng, Karen Brennan
haduong@g.harvard.edu, carterzenke@gse.harvard.edu, jessicaeng@college.harvard.edu, karen_brennan@gse.harvard.edu
Harvard University

Abstract: Online professional learning experiences can support computing teachers in developing content knowledge and expanding professional networks. Within the context of an 11-day online professional learning experience for computing teachers working with Scratch, this study explored interactions between participants and facilitators with varying levels of Scratch familiarity. Findings describe how facilitators and participants with lower and higher levels of Scratch familiarity interacted with one another, which can inform the design of professional learning experiences that support teacher collaboration.

Introduction

Broadening access to PK–12 computing education requires supports for teachers, and yet computing teachers report limited opportunities for professional learning (Yadav et al., 2016). While some scholars have argued that professional learning experiences should vary for teachers with different levels of prior experience (Qian et al., 2018), teachers can benefit from opportunities to interact with others of varying expertise (Penuel, Sun, Frank, & Gallagher, 2012). In this study, we explore the following research question: *How did teachers with varying prior Scratch familiarity interact when participating in an online professional learning experience?*

We take a situated learning approach (Sentance & Humphreys, 2018) and draw on Lave and Wenger's (1991) conception of *learning as participation* to understand computing teachers' experiences in Getting Unstuck, a 2020 11-day online professional learning experience that aimed to support PK–12 computing teachers' developing familiarity with Scratch, a widely-used introductory programming language and online community. Participation in this context includes not only creating projects, but also receiving and sending comments on projects. Each day of Getting Unstuck, participants received an email with an invitation to create a Scratch project in response to a prompt, share their projects with others in a studio public to the Scratch community, and reflect on their daily work. Five facilitators (elementary school teachers and university researchers) modeled interactions by leaving comments on participants' projects.

Method

1,009 participants signed up through an introductory survey to receive the daily emails with the project prompts. 298 participants created and submitted at least one project across the 11 days. 168 of the 298 participants (56%) completed 10 or more projects, and we collected participants' 10,576 comments and associated metadata across 2,251 projects submitted to all 11 studios. We then grouped the sample into three categories: lower familiarity (LF), higher familiarity (HF), and facilitators. In the introductory survey, participants self-reported their familiarity with Scratch across five categories: "I've never used Scratch before," "I've seen others use Scratch," "I've made a few projects," "I've helped others make projects," and "I'm well-acquainted with the ins and outs of Scratch." Participants were also asked to self-report prior participation in a 2018 pilot of Getting Unstuck.

Participants who placed themselves in the first three categories *and* did not participate in 2018 were grouped as lower familiarity (LF). Participants who placed themselves in the last two categories *and/or* participated in 2018 were grouped as higher familiarity (HF). Our sample of 298 participants included 69 (23.2%) LF users, 224 (75.2%) HF users, and 5 (1.7%) facilitators. We sought to understand interactions between participants with the same degree of familiarity, over time. Our analysis adopts Krackhardt & Stern's (1988) measures to evaluate the relative proportion of in-group and out-group ties within a social network, visualizing the evolving network with directed graphs and calculating a normalized index to quantify the extent to which participants interacted within or outside their group. We then track how these measures evolved throughout Getting Unstuck to identify and compare participation trajectories for LF and HF teachers.

Findings and discussion

Participants engaged in substantial and consistent interactions with others. Across the 11 days, 45 out of 69 LF teachers (65%) and 185 out of 224 HF teachers (83%) made at least one comment. Though the number of participants decreased over time, participants remained consistently engaged with others' projects, leaving an average of 4.1 to 6.0 daily comments per project (cumulative average of 4.7 over all days).

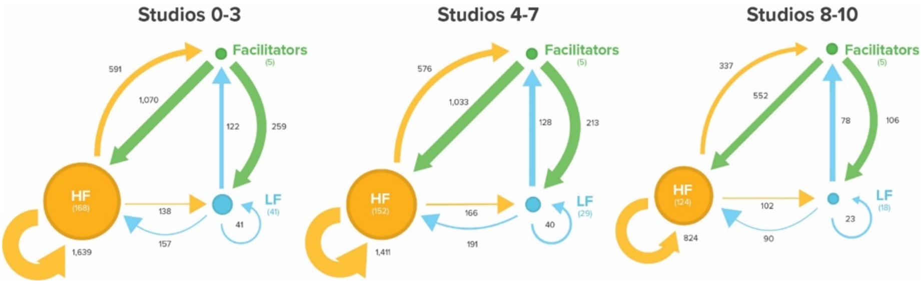


Figure 1. Directed comment interactions among HF teachers, LF teachers, and facilitators.

We found evidence of different trajectories of participation between HF and LF teachers (see Figure 1). We calculated the difference between HF and LF teachers' externally directed comments per person (i.e., across groups) and internally directed comments per person (i.e., within groups), dividing by the sum of these metrics to reach a normalized index of externality and internality, which has two poles: -1, indicating engagement only within the group, and 1, indicating engagement only outside of the group. LF teachers began Getting Unstuck with a moderate tendency to interact with HF teachers and facilitators (EI = 0.23), rather than interacting with one another. Over time, however, LF teachers strengthened interactions with those who had similar levels of prior Scratch familiarity (EI = 0.19 and 0.01 in studios 4–7 and 8–10, respectively). On the fifth day, for example, one LF user debugged another LF user's project: "I remixed it and I think I got it to work!" HF teachers also began Getting Unstuck with a moderate tendency to reach outside of their group (EI = 0.24) but exhibited an opposite trajectory of participation: as they progressed, they engaged more often with facilitators and LF teachers (EI = 0.40 and 0.48 in studios 4–7 and 8–10, respectively). In Studio 6, an HF user asked an LF user about the project they had made: "You need way less code than I did. Do I understand it correctly that you use a broadcast block to change the color of the clones and that all clones that need to change color do so when that message is broadcasted?" The LF user responded, "Indeed, every time a clone is clicked it sends a message to all the others and stores its value on a shared, temporary variable."

Teachers' interactions with one another in Getting Unstuck offer evidence for how participation in an online community can evolve, even within a short timeframe. These opportunities to learn from and with others may also support teachers in deepening their own familiarity with Scratch. Understanding how teachers engage with the work of others with varying computing backgrounds can inform the design and facilitation of professional learning experiences that create opportunities for teacher collaboration. Future work could examine the substance of participants' comments as well as the role of skilled facilitators, offering directions for the design of future professional learning activities.

References

Krackhardt, D., & Stern, R. N. (1988). Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly*, 123-140.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Penuel, W. R., Sun, M., Frank, K. A., & Gallagher, H. A. (2012). Using social network analysis to study how collegial interactions can augment teacher learning from external professional development. *American Journal of Education, 119*(1), 103-136.

Qian, Y., Hambrusch, S., Yadav, A., & Gretter, S. (2018). Who needs what: Recommendations for designing effective online professional development for computer science teachers. *Journal of Research on Technology in Education*, *50*(2), 164-181.

Sentance, S., & Humphreys, S. (2018). Understanding professional learning for Computing teachers from the perspective of situated learning. *Computer Science Education*, *28*(4), 345-370.

Yadav, A., Gretter, S., Hambrusch, S., & Sands, P. (2016). Expanding computer science education in schools: Understanding teacher experiences and challenges. *Computer Science Education*, *26*(4), 235–254.

Acknowledgments

This work has been supported by the National Science Foundation, Grant #1908110. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s).

Studying Shared Regulation in Immersive Learning Environments

Rinat Levy Cohen, Fordham University, rlevycohen@fordham.edu
Aditi Mallavarapu, Leilah Lyons, Stephen Uzzo
amalla5@uic.edu, llyons@nysci.org, suzzo@nysci.org
New York Hall of Science

Abstract: We examined the regulation of shared problem solving in a museum exhibit. We found that we had to augment our dialogue codes to properly embrace the dynamic nature of the observed learning regulation. These changes reflect aspects of shared regulation that occur when learning takes place (1) in an immersive open-ended learning environment, where (2) learners work together in large groups. We present preliminary results, arguing that designers and researchers may benefit from recognizing how planning and evaluation acts can be tactically embedded in immersive learning environments.

Keywords: SRL, SSRL, informal learning, immersive learning environments.

Introduction and background

This work is part of a larger project with the goal of designing a digital, data-driven dashboard to help museum educators facilitate an immersive multi-user exhibit, Connected Worlds (CW). The exhibit allows visitors to engage in collaborative problem solving, and it became apparent that they needed support with the exhibit content and with managing the problem solving process as a group. Designing a dashboard that would bolster socially shared regulation of learning (SSRL) (Järvelä & Hadwin, 2013) required us to understand how visitors work with each other in the exhibit. We began studying the visitors by developing a dialogue coding scheme rooted in the self-regulated learning (SRL) and SSRL literature, but found that the traditional SRL and SSRL conceptions didn't quite capture our group dynamics.

Generally, SRL models have a preparatory phase, which includes familiarizing with the task, followed by a performance phase, which includes strategy use and monitoring progress, and finally an appraisal phase, where learners evaluate and reflect on their performance and plan for future performance, when appropriate (Panadero, 2017). The SSRL framework (Järvelä & Hadwin, 2013) adapted SRL to social learning scenarios, viewing SRL phases as co-constructed by individuals and distinguishing between three levels of learning regulation (self, group members, and the group as a whole).

Phases of SSRL inform the design of scaffolds in asynchronous CSCL settings. For example, goal setting requires different types of information than monitoring, and depending on the target of the regulation, the presentation of the scaffolds may change (Järvelä & Hadwin, 2013). This paper demonstrates that when a learning environment is highly immersive and includes large numbers of learners, additional aspects of SSRL surface. We developed a modified coding scheme that builds on and extends SSRL to immersive CSCL environments (Levy-Cohen et al., 2021), revealing the prevalence of tactical planning and evaluation in group regulation. The following questions guided our study: *How do groups regulate their learning in an immersive digital learning environment? What SRL processes and sub-processes come to play?*

Methods

Setting, sample, and procedure

Connected Worlds is an ecological simulation at the New York Hall of Science. Visitors need to work collaboratively -- within their team and with the other team -- to help the interconnected biomes (Desert, Grasslands, Jungle, and Wetlands) thrive, by routing water to biomes and engaging in forestry management (see Mallavarapu et al., 2019). A volunteer group new to the simulation was recruited (N=26, 22-57 years of age, M=33), and randomly divided into two separate 30 minute sessions (N= 12, and N=14) and again into one of four smaller teams (one team per biome). Visitors wore digital lapel recorders to capture their conversations.

The data from the recorders were transcribed and segmented into speaking turns. Three transcripts (~12% of the total corpus) were randomly selected. Only the speaking turns of the person wearing that recorder were coded. Initial data analysis applied *a priori* codes taken from the SRL and SSRL literature and further elaborated through inductive coding by three researchers (Miles et al., 2014). Emerging codes and themes were recorded and then compared to create a revised codebook. For interrater reliability scores below 80%, we reviewed the disagreements and used those discussions to revise the codebook.

Initial findings and discussion

All the SRL and SSRL codes were represented in the transcripts (e.g., Grounding: 19%, Monitoring: 44%, Planning 5%, and Evaluating 5%; of 103 total speech turns). We were surprised by the low incidence of Planning and Evaluating codes, as the participants were clearly coordinating in an organized fashion. The challenge was that the traditional Planning and Evaluating definitions didn't quite fit in the context of an immersive learning environment. Traditionally, self-evaluation takes place when individuals assess their learning progress and compare it to a goal they set for themselves (Zimmerman, 2008). We could seldom detect clear goals in learners' talk, although participants were clearly making judgments about the current scenario, and making and adjusting plans. For example, viewing one participant's traditional codes over time (Figure 1) suggests the participant only evaluated sporadically. We thus decided to code Evaluations with clear, persistent goals (e.g., "Wait, they're getting enough [water] too, and we are, so they're good.") as *Strategic* Evaluations (SE) while an evaluation of the satisfaction of an emergent, implied goal as a *Tactical* Evaluation (TE) (e.g., "Okay, we need to direct it more that way."). This added lens reveals that TE is a major part of group coordination (55% of speech turns) and is often interleaved with SE (18% of speech turns). This suggests that the support tool needs to incorporate dynamic indicators that can be easily inspected, like gauges, in addition to data visualizations that require longer time to derive insights (e.g., line graphs). We came across similar challenges when coding for planning events. For example, while participants proposed plans with larger, strategic goals (e.g., "Let's set up an irrigation system") they also proposed more immediate, tactical plans (e.g., "Let's shunt the reservoir water over there"). The information needed to decide how much water to move in the moment is different from the information one needs to construct a strategic division of resources.

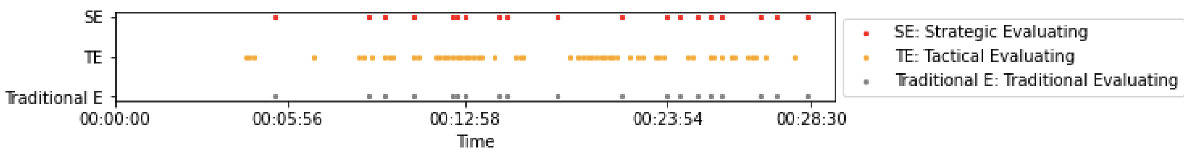


Figure 1. One participant's speech, comparing traditional Evaluating and revised Tactical/Strategic Evaluating.

Conclusion

The coding of our full transcript corpus is incomplete, but by inspecting a limited sample through the lens of our coding scheme we have been able to infer several preliminary findings. First, the degree to which groups collaborate seems to align with social regulatory processes, especially high frequencies of monitoring and grounding processes. We also found that using the traditional definition of Evaluating would grossly underestimate the amount of Evaluating work done by groups in immersive settings. The interleaving of Tactical and Strategic talk suggests that we should investigate what role Tactical Planning and Evaluating plays *vis a vis* Strategic Planning and Evaluating in immersive learning environments. For example, is TE decompositional (addressing smaller components of larger goals), or evolutionary (leading to shifting of larger goals over time)? A decompositional view suggests supports and strategies (like "divide and conquer") that break goals into tasks, whereas an evolutionary view would suggest that learners may need support comparing and contrasting.

References

Mallavarapu, A., Lyons, L., Uzzo, S., Thompson, W., Levy-Cohen, R., & Slattery, B. (2019). Connect-to-connected worlds: Piloting a mobile, data-driven reflection tool for an open-ended simulation at a museum. In *Proc. CHI '19*. ACM Press, New York, NY, 1–14.

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist, 48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006.

Levy-Cohen, R., Mallavarapu, A., Lyons, L., Thompson, W., & Uzzo, S. (2021). Studying Collective Problem Solving Regulation in an Immersive Open-Ended Museum Exhibit. In *AERA '21*.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook.* SAGE Publications: Los Angeles, CA.

Panadero, E., & Järvelä, S. (2015). Socially shared regulation of learning: A review. *European Psychologist, 20*, 190–203. https://doi.org/10.1027/1016-9040/a000226.

Acknowledgments

This material is based upon work supported by the National Science Foundation under grant nos. 1623094 & 1822864.

A Template for Facilitating Knowledge-Building Discourse in Online Teacher Professional Development

Bohdana Allman, Brigham Young University, bohdana.allman@byu.edu
Heather Leary, Brigham Young University, heather.leary@byu.edu

Abstract: Effective teacher professional development is participatory in nature, i.e., situated in practice, collaborative, dialogic, and inquiry-based. Current technologies make online participatory learning experiences increasingly possible. This design case study presents a pedagogy-based template, the Progressive Instructional Conversation (PIC), that guided the redesign of existing teacher professional development courses grounded in sociocultural practices into an online modality. The template elements and their role in facilitating progressive knowledge-building discourse online are presented and discussed.

Keywords: online, knowledge-building discourse, teacher professional development, design.

Background

This study was part of a more extensive design-based research project at a large private university in the western United States. The project's overall goal was to redesign six participatory teacher professional development (TPD) courses into fully online courses to improve access to resources and flexibility of instruction. The program supports teachers' learning through collaboration, dialogue, and inquiry-based approach situated in practice (participatory TPD). It incorporates principles of active learning and adult learning theory, modeling of effective practices, provides opportunities for reflection on one's practice, and offers coaching and expert support, representing best TPD practices (Borko et al., 2010; Darling-Hammond et al., 2017; Dede et al., 2009). Facilitating progressive knowledge-building discourse and promoting dialogic learning is at the core of enacting the program (Hofmann, 2019). It supports teachers as they develop a complex understanding of context-specific and situational issues and apply their knowledge and skills in their classrooms (Harasim, 2017; Scardamalia & Bereiter, 2014; Wells, 2002).

Maintaining the participatory character of instruction and supporting progressive knowledge-building discourse while taking full advantage of available online resources and affordances became the project's key design requirement. The Progressive Instructional Conversation (PIC) template was created to systematically guide the redesign and development processes and to support an effective facilitation of the completed courses. This investigation and related findings are presented as a design case study, which is especially useful in disclosing localized design practice details, related experiences and understanding, and innovative solutions to complex challenges (Boling & Smith, 2012). This design case aims to present the PIC template and explain its structure and functionality revealing its role in facilitating progressive knowledge-building discourse.

Design description

The PIC template represents the structure of a module. Typically, there are ten thematically-focused modules in a course guiding a practicum-based portfolio development. Content and activities within each module are organized in two directions: horizontal and vertical. The template's horizontal progression assists the learners as they move across the four module topics from theoretical focus to practical application. This theory-to-practice progression is an essential element of effective TPD and offers a logical structure for online instruction. It provides a way for teachers to reflectively transcend the theory-practice divide and become comfortable operating from a principled-based practice stance while flexibly responding to the messy yet the most important practical problems within their classrooms. Explicitly attending to this horizontal theory-to-practice progression around module topics creates a pattern of (1) unfolding of complex theoretical concepts in familiar and learner-friendly terms, (2) assisting learners as they interpret their prior practical experiences through a theoretical lens, and (3) supporting the learners as they anchor their classroom practices in a solid theoretical understanding.

The template's vertical progression supports the learning process advancing from individual and collaborative learning toward assisted instruction, reflection, and performance assessment. It is based on a well-known notion of the zone of proximal development (Vygotsky, 1978). All instructional elements, i.e., learning, assistance, assessment, are always present, but the emphasis changes with progression. Two key theory-related factors frame the pedagogical purpose of each phase: *the type of interactions* (Anderson, 2008) and *the phase* *within the inquiry cycle* (Garrison et al., 2001; Harasim, 2017; Wells, 2002). The template's vertical progression assists learners as they develop a mastery of conceptual complexity necessary to effectively and flexibly use their theoretical understanding in their everyday practice as they work in diverse situations with a variety of learners.

Learners begin each module by reviewing key concepts, developing background understanding, and reflectively connecting it with their prior knowledge and experiences. This prepares them for active participation in small group collaborative activities where they review each other's work, consider various perspectives, and seek and negotiate acceptable solutions for a common task that expands both their individual and group understanding. In the process, the group discourse becomes exploratory, an important characteristic of a knowledge-building discourse associated with sustained development of ideas. Meeting with others and the facilitator during the class conference is central for completing the intellectual convergence phase of collaborative work. As individuals articulate their conceptual knowledge, share gained understanding, and are exposed to different ideas and solutions, their understanding further develops. The facilitator assists learners as they make deeper connections across topics within the module, challenges their thinking, and highlights and shares notable examples of practical application. These meetings also provide an opportunity to begin an individual reflection on the module's most valuable ideas and their application in one's practice. Practicum assignment is the culminating experience for each module where learners reveal their ability to apply their theoretical understanding in their own practice and receive individualized feedback and coaching. Practicum assignments build on each and, together with other evidence, become a portfolio-based course assessment.

Conclusion

Developing a template that facilitates progressive knowledge-building discourse in online modality creates powerful opportunities for more accessible and flexible professional development for teachers and other professionals. There is a great urgency for quality collaborative models to bring together geographically distributed professionals, promote deep connections between theory and practice, and provide easy and flexible access to resources and interactions within communities of practice.

References

Anderson, T. (2008). Toward a theory of online learning. In T. Anderson & F. Elloumi (Eds.), *Theory and practice of online learning* (2nd ed., pp. 45–74). Athabasca University. http://cde.athabascau.ca/online_book/

Boling, E., & Smith, K. M. (2012). The design case: Rigorous design knowledge for design practice. *Interactions, 19*(5), 48–53. https://doi.org/10.1145/2334184.2334196

Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 7, pp. 548–556). Elsevier. https://doi.org/10.1016/B978-0-08-044894-7.00654-0

Darling-Hammond, L., Hyler, M., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. https://learningpolicyinstitute.org/product/teacher-prof-dev.

Dede, C., Ketelhut, D., Whitehouse, P., Breit, L., & McCloskey, E. M. (2009). A research agenda for online teacher professional development. *Journal of Teacher Education, 60*(1), 8–19. https://doi.org/10.1177/0022487108327554

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education, 15*(1), 7–23. https://doi.org/10.1080/08923640109527071

Harasim, L. (2017). *Learning theory and online technologies* (2nd ed.). Routledge.

Horfmann, R. (2019). Dialogue, teachers and professional development. In N. Mercer, R. Wegerif, & L. Major (Eds.), *The Routledge international handbook of research on dialogic education* (pp. 213-216). Routledge. https://doi.org/10.4324/9780429441677-18

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 97–115). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.025

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wells, G. (2002). Inquiry as an orientation for learning, teaching and teacher education. In G. Wells & G. Claxton (Eds.), *Learning for life in the 21st century: Sociocultural perspectives on the future of education* (pp. 195–210). Blackwell Publishing. https://doi.org/10.1002/9780470753545.ch15

International Society of the Learning Sciences

From "in a sleep" to "stayed every day": Engaging Students and Teachers with micro:bit Smart-Greenhouses

Yihong Cheng, Boston College, chengyz@bc.edu
David Jackson, Boston College, jacksondw@bc.edu

Abstract: We report findings from surveying and interviewing 10 middle-school science teachers after 14, 1-hour online trainings. Sessions prepared teachers for embedding a BBC micro:bit based automated table-top greenhouse curriculum, as developed by the research team for middle-school science classrooms. Teachers expressed strong interest and confidence in this curriculum as a tool to engage all students in exploring computation in science, suggesting a promising way of using cross-disciplinary hands-on projects in science teaching and learning.

Introduction

Physical computing, or programming digital devices that interact with the physical world, opens possibilities to make science learning more engaging for students (Przybylla & Romeike, 2018). The BBC micro:bit is gaining traction as a tool that introduces students to block-based coding and solving problems relevant to their lives, receiving positive feedback from students and teachers alike. In response to this trend, we are writing to share insights drawn from a series of 14, 1-hour online micro:bit training sessions the research team recently developed about programming automated table top smart-greenhouses.

Adopting the strategy of "decomposition of complex skills and tasks into minimal constituent components" (Reiser & Tabak, 2014, p. 47), we modularized learning activities such that teachers could focus on one sensor/actuator each time (see summary in Table 1). Two overarching research questions guided our work: 1) What are teachers' perspectives about using a micro:bit-based smart-greenhouse project to teach about computing?; 2) What scaffolding will benefit teachers as they prepare to engage with students in the project?

Table 1: Summary of Professional Development Sessions

Week	Topic(s)
1	Introduction to BBC Micro:bit & Grove shield hardware; introduction to MakeCode software; how to transfer files; use of micro:bit LEDs, Grove OLED screen, and Grove LED strip
2	Introduction to (programming) functions; using a Grove temperature-and-humidity sensor
3	Introduction to if-then(-else) loops; using a relay (switch) to activate or deactivate circulation or exhaust fans
4	Using the micro:bit's built-in light sensor to turn on or off an LED lamp; performing arithmetic in MakeCode
5	Calibrating and controlling a servo motor to open or close the greenhouse's windows
6	Communication between two or more micro:bits; using a Grove gesture sensor and/or passive infrared sensor to control various outputs
7	Demonstration of integration with Google Sheets; introduction to micro:bit Classroom (learning management software); planning future sessions; time for focus-group and survey

Material and design

The project focused on using micro:bit and related extension devices to create a tabletop smart-greenhouse. Using Microsoft® MakeCode, micro:bit's complementary block-based coding system, teachers learned to program a variety of sensors and actuators to make the greenhouse automatically measure environmental variables, then take actions to maintain these variables in ways that are ideal for plants' growth. While a full set of these devices can cost $150 per greenhouse, we've also developed a simulator based, completely free virtual version (Figure 1).

Methods

To answer our motivating questions about teachers' perspectives and future scaffolding, we made use of four forms of qualitative data analysis. First, we member-checked field memos for informative moments in the training sessions. Second, we conducted document analysis on Zoom® chat histories in search of conversations that revealed teachers' insights. Third, we transcribed audio recordings of two post-training focus group interviews to generate axial codes per grounded theory (Strauss & Corbin, 1990) via multiple iterations of forming, grouping, and organizing categories of open codes. The same interview protocol was used in both focus groups. Finally, we axially coded teachers' responses to open-ended questions in a post-training survey.

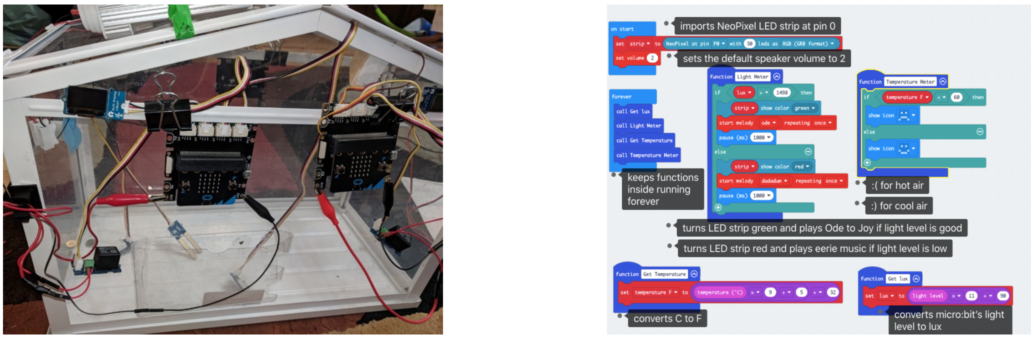


Figure 1. A micro:bit smart-greenhouse and annotated codes. The forever loop runs functions inside forever.

Results

With respect to our first research question, most teachers reported confidence in introducing their students to the micro:bit smart-greenhouse. Teachers who had worked with the research team before to integrate the Python smart-greenhouse were especially optimistic because of their observation of students' reactions. For example, one teacher recalled in a focus group interview that there were a few 8th grade students who had "been in a sleep" and were "not involved in anything all year" before turning into "entirely different human beings" when they began working on the smart-greenhouse.

With respect to our second research question, one aspect of the training sessions that worked well in supporting teachers to prepare for launching the micro:bit smart-greenhouse project was having enough time between sessions to reflect on what they learned and try it out. Meanwhile, teachers really liked the scaffolding for a complicated project into small components that turned "technology to be more productive for learning by doing" (Reiser & Tabak, 2014, p. 56), or in one teacher's words, "[the training] went forward in a really logical manner" that "built on itself well".

Discussion

This study supports the promise of integrating Micro:bit into middle school science classrooms to better engage students, in addition to raising a novel model of online professional development that scaffolds for teachers an original curriculum centered on building and programming automated smart-greenhouses. One limitation that currently restricts a full-fidelity, large-scale implementation of this project is funding. Our work is supported by a National Science Foundation grant, which enabled us to purchase necessary materials, spend time organizing materials into kits, and mail kits to participating teachers. For teachers with limited funding, we do realize that the cost of a fully functioning micro:bit greenhouse equipped with all the devices included in our current design can be unaffordable. One possible adaptation that can be made for the sake of cost reduction is to make use of the free simulator function at makecode.microbit.org to simulate what the micro:bit can do when it is programmed to respond to changes in light level and temperature.

References

Przybylla, M. & Romeike, R. 2014. Key competences with physical computing. In Brinda, T., Reynolds, N. & Romeike, R. (Eds.), *Proceedings of key competencies in informatics and ICT 2014*. 216-221.

Reiser, B. J. and Tabak, I. 2014. Scaffolding. In Sawyer, R. K. ed. The Cambridge Handbook of the Learning Sciences. Cambridge University Press, New York, NY, 44-62.

Strauss, A. and Corbin, J. 1990. Basics of qualitative research: grounded theory procedures and techniques. SAGE Publications.

Exploring the Role of Curriculum in Learning Analytics to Support Knowledge Building Practice

Chew Lee Teo, Aloysius Ong, Alwyn Vwen Yen Lee
chewlee.teo@nie.edu.sg, aloysius.ong@nie.edu.sg, alwyn.lee@nie.edu.sg
National Institute of Education, Singapore

Abstract: One of the challenges of knowledge building practice is for teachers to prioritize the process of using student questions to guide knowledge building in view of the demand of curriculum and assessment in a regular school structure. Knowledge building analytics was developed to empower learners' epistemic agency and support paths of idea improvement (Chen & Zhang 2016) but it was not explicitly designed to support teachers and students to develop ideas beyond the fulfillment of a prescribed curriculum. In this paper, we report the iterations of design of a set of analytics that compared keywords from Knowledge Forum (KF) notes against keywords mined from an expanded curriculum mapped across different grade years and topics. We termed the analytics, Curriculum-ideas-Analytics (CiA). Preliminary findings show that the balance between curriculum and students' ideas is not easy navigation. Still, there is much potential in this design to support teachers to consider students' diverse and seemingly naive ideas about big ideas embedded within the curriculum and better facilitate the identification of new lines of inquiry.

Introduction

Idea improvement is an iterative and critical process of knowledge building (Scardamalia and Bereiter, 2006). Teachers and students go through the process of elaboration, refinement and exchange of ideas. Teachers prioritize students' questions and ideas to build knowledge and teachers play this role by systematically analyzing students' ideas and identifying potential points of view from the discourse. This task is non-trivial as teachers need to sift through many notes on Knowledge Forum (KF) (Scardamalia and Bereiter, 2006) to gain insights into the discussion and to bring the identified ideas from online discussion into a face-to-face discussion (Knowledge Forum is an online discussion platform with features to support collective idea development). This process of mining, analyzing and evaluating online discussion is tedious and often lack objectivity. In this study, we report the design of CiA to help teachers compare students' textual contributions in KF against a connected and expansive cross-level curriculum. By tapping on the big ideas and unifying themes that are often expounded in curriculum documents, the CiA tool was observed to serve the purpose in generating useful lines of inquiry with sound professional development structure.

Design of Curriculum-Idea-Analytics as a scaffold for teachers

The iterative design process of CiA:

1. ***Creation of expanded curriculum map connected across grades and topic***: The Specific Instructional Objectives (SIOs) from related topics (e.g. electricity and magnetism) of different grades are mapped to a focal theme known as unifying ideas of 'system'. This mapping emphasizes how different parts and functions of an electrical and magnetic system may work together for common purposes. From this mapping, we derived four big ideas of electricity and magnetism in system. The four ideas are (a) energy conversion; (b) energy conservation; (c) electrical system; and (d) electrical energy. This mapping captured the trajectory depicted in curriculum knowledge across grades. For example, if a Grade Three student became interested about electromagnets (originally a Grade nine idea) before it is formally introduced in the curriculum. The teacher might label it as out-of-syllabus or they might let the children know that they would learn it in higher grade, or they could decide to embrace the electromagnetic question and make it the focus on the class' inquiry – and design inquiry activities on electricity, magnetism and electromagnetism. In our case, a blinking light experiment was adopted in a Grade five class when the interest emerged.
2. ***Text-mining mechanism*** was then used in CiA to benchmark student ideas to a cross-grade, expanded curriculum. This is done by visualizing the semantic space of student notes on Knowledge Forum with the semantic space of the expanded curriculum (Figure 1).

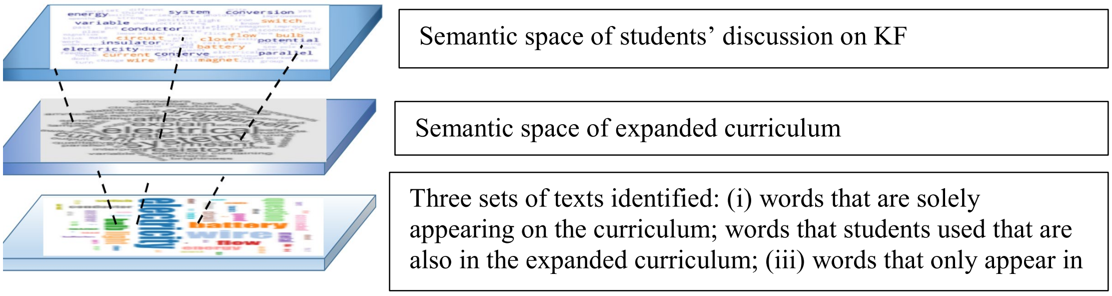


Figure 1. The three interacting semantic spaces that undergirds the design of CiA.

3. ***Exploring more connections***. CiA was further refined to allow the selection of key concepts and unifying ideas for different grade levels. Users can select to view curriculum ideas and key concepts from other grades Teachers can focus on analyzing specific semantic space of students to see its connection to the big ideas of curriculum through an extended network or word visualization. (see Figure 2).

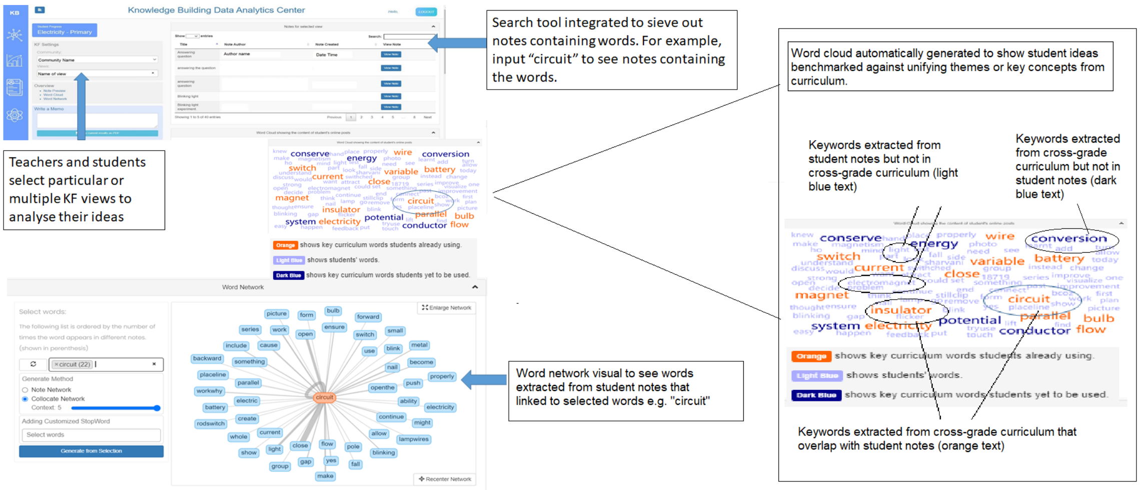


Figure 2. A word cloud (right) visualization showing the synthesized semantic space of students discourse and the expanded curriculum. Different colored text showing different semantic spaces and the user interface of CiA showing the word-network that teachers can view the connections between selected words.

Ongoing and future work

Students and teachers are often the main creators of the content in their community knowledge spaces (Hong & Scardamalia, 2014), and CiA provides a representation of their contributions from which they can quickly return to the context of use. The design for CiA has developed and expanded to include interdisciplinary topics and connects curriculum words with similar students' ideas. More case studies will be conducted to establish CiA's potential as intermediary analytics to support idea-centric pedagogy.

References

Chen, B., & Zhang, J. (2016). Analytics for knowledge creation: Towards epistemic agency and design-mode thinking. *Journal of Learning Analytics*, *3*(2), 139-163. https://doi.org/10.18608/jla.2016.32.7

Hong, H.-Y., & Scardamalia, M. (2014). Community knowledge assessment in a knowledge building environment. *Computers & Education*, *71*, 279–288. http://dx.doi.org/10.1016/j.compedu.2013.09.009

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.

Acknowledgments

This project was funded by the Edulab Research Grant, Office of Education Research, Nanyang Technological University, Singapore (AED 04/17 MC). The views expressed in this paper are the author's and do not necessarily represent the views of the host institution. We thank A/P Chen Mei Hwa for developing the CiA analytics tool.

Initial Analysis of Prompted Discourse Patterns in an Informal, Online, Global Collaborative Learning Environment

Danielle P. Espino, Pepperdine University, danielle.espino@pepperdine.edu
Seung B. Lee, Pepperdine University, seung.lee@pepperdine.edu
Megan Hokama, Pepperdine University, megan.hokama@pepperdine.edu
Eric R. Hamilton, Pepperdine University, eric.hamilton@pepperdine.edu

Abstract: This analysis examines the discourse patterns of adolescent participants from two countries while engaging in synchronous video conference calls to collaborate on STEM-oriented media projects. Epistemic network analysis (ENA) is used to examine the influence of prompting on student discourse. Results highlight how unscripted prompting helps to generate rich, diverse responses connected to content and curiosity that might not otherwise take place.

Introduction

As learners adjust to a collaborative space dynamic, high levels of discourse engagement and effective instructional methods are vital to creating a sound social space for learning to take place (Kreijns et al., 2013). The role of scaffolded and guided assistance promotes discussion, elaboration, explanation, and reflection in online settings (Morris et al., 2010). Prompts are identified to be composed of hints, guiding questions, and suggestions for improvement (Harney et al., 2015). This project involves adolescent learners who participate in afterschool clubs to collaborate online with learners in other countries on developing digital media projects with STEM (science, technology, engineering, math) focused content to enhance their understanding of those subjects. Participants interact synchronously through video conference calls known as online global meet-ups, where participants from at least two clubs join and a facilitator helps support the conversation. In an effort to foster engagement, the facilitator informally calls on participants with unscripted prompts related to the topic being discussed. Peers also contribute to prompting one another during the meet-ups. The exchange of information and feedback among students can lead to enhanced cognitive engagement, especially as they provide and receive explanations from their peers (Ge & Land, 2004). This study looks at the influence of prompting in promoting student engagement in an informal, online collaborative environment.

Methods

This analysis examines discourse data from two online meet-ups held in 2018 and 2019 involving students in Kenya and the United States. In the analysis, each turn of talk, or utterance, represented one line of data. Two raters separately coded all lines for eight constructs, identified as the most relevant from a grounded analysis of the data: Collaborative Disposition, Content Focus, Curiosity, Feedback, Information Sharing, Media Production, Participatory Teaching and Social Disposition (see Table 1). Each utterance was analyzed for whether it was prompted by another line (i.e., the response given by a participant to a specific, directed request for their input was classified as a prompted utterance). All other lines, including self-initiated elaborations, were categorized as unprompted utterances. The final coding was determined through a process of social moderation.

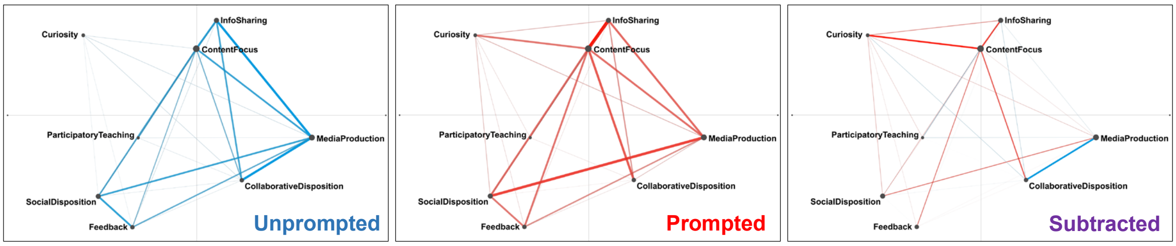
Table 1: Codebook of constructs included in the analysis

Code	Description
Collaborative Disposition	Promoting cooperation between two or more individuals to accomplish a project task
Content Focus	Dialogue focused on the meet-up's STEM-related educational content
Curiosity	Seeking clarification for better understanding of STEM-related content or project
Feedback	Communicating one's opinions/ideas or sharing suggestions on projects
Information Sharing	Sharing of personal experiences or contextual information relevant to the discussion
Media Production	Dialogue related to the production of media artifacts
Participatory Teaching	Helping others to learn STEM subject matter by providing fact information in explanation
Social Disposition	Demonstrating pro-social tendencies (e.g., appreciation, acknowledgement or validation)

The coded data was analyzed using epistemic network analysis (ENA), a technique in quantitative ethnography utilizing visualization and statistical methods to identify meaningful patterns in discourse. ENA models the connections among salient constructs in the data by quantifying the frequency of their co-occurrences within conversations (Shaffer, 2017). In this study, an individual participant was defined as the unit of analysis, and each meet-up constituted a conversation to which the connections were limited.

Results

This study analyzed a total of 699 utterances from meet-ups in November 2018 and May 2019. A total of 15 unique students from Kenya and the U.S. participated in the meet-ups. Of the 425 lines spoken by students, prompted utterances accounted for about 34% while unprompted utterances made up around 66%. Figure 1 displays the ENA network for the prompted and unprompted utterances spoken by student participants during the two meet-ups. The individual networks exhibit similarities in their prominent connections between several codes. In particular, it can be seen that Content Focus plays a central role in both networks, with strong associations to Information Sharing, Social Disposition, and Media Production, among others. As shown in the subtracted network model (c), the main difference between the two networks is the connection between Media Production and Collaborative Disposition, stronger in the unprompted utterances. On the other hand, the linkage between Content Focus and Curiosity was much more prominent in the prompted utterances, along with relatively thicker connections between Content Focus, Collaborative Disposition and Information Sharing.



(a) (b) (c)

Figure 1. ENA networks for (a) unprompted (b) prompted utterances, and the (c) subtracted network

Discussion

The results above show the impact of prompting on the discourse patterns exhibited by participants in a global, virtual collaborative learning environment. In comparing prompted and unprompted utterances from students, prompted responses have stronger connections to content and other constructs, notably with Curiosity, Information Sharing, Collaborative Disposition and Feedback. This reflects how much unscripted prompting promotes discourse related to content. Unprompted utterances uniquely drew a strong connection independent of content, between Collaborative Disposition and Media Production, speaking to the natural desire for students to discuss making media artifacts together in the project environment. These findings highlight how unscripted prompting helps to generate rich, diverse responses connected to content and curiosity that might not otherwise take place. At the same time, unprompted student utterances compliment strong content connections from prompts by focusing on collaborative and media aspects, such as expressing the desire to work together on media projects. While these initial results focused on prompted versus unprompted discourse, the use of ENA provides additional possibilities for analysis such as peer versus facilitator, type of prompt, and along other metadata (e.g., gender).

References

Ge, X., & Land, S. M. (2004). A conceptual framework for scaffolding ill-structured problem-solving processes using question prompts and peer interactions. *Educational Technology Research and Development, 52*(2), 5-22. https://doi.org/10.1007/BF02504836

Harney, O., Hogan, M., Broome, B., Hall, T., & Ryan, C. (2015). Investigating the effects of prompts on argumentation style, consensus and perceived efficacy in collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 10*(4), 367-394.

Kreijns, K., Kirschner, P. A., & Vermeulen, M. (2013). Social aspects of CSCL environments: A research framework. *Educational Psychologist, 48*(4), 229-242.

Morris, R., Hadwin, A. F., Gress, C. L. Z., Miller, M., Fior, M., Church, H., & Winne, P. H. (2010). Designing roles, scripts, and prompts to support CSCL in gStudy. *Computers in Human Behavior, 26*(5), 815-824.

Shaffer, D. W. (2017). *Quantitative ethnography*. Madison, WI: Cathcart Press.

Acknowledgements

The authors gratefully acknowledge funding support from the US National Science Foundation (#1612824) for the work this paper reports. Views appearing in this paper do not reflect those of the funding agency.

Scaffolding Epistemic Understanding of Discourse and Knowledge Building Using Knowledge-Forum Analytics and Reflective Assessment

Yuyao Tong, Carol K. K. Chan
yyttong@hku.hk, ckkchan@hku.hk
The University of Hong Kong

Abstract: This study investigates how students developed an epistemic understanding of the nature of discourse and engaged in productive inquiry using analytics and reflective assessment. The study reports on a case study of a class of Grade 11 students, working on collective inquiry in Artifacts Design in a knowledge-building environment supported by Knowledge Forum® (KF). A key design involved students engaging in reflective assessment by reflecting on their KF discussion using analytics information with the help of visualizations of ideas-building networks. Results indicated that the students developed their understanding of discourse and engaged in productive online discourse progressively. The analysis also suggested how students with different understanding levels of discourse performed differently in their KF work. Qualitative analyses revealed how students developed their epistemic understanding and productive inquiry through analytics-supported reflections.

Introduction

As educational goals change and technology advances, the development of the new educational skills and learning capabilities needed for the 21st century knowledge era is drawing widespread attention (OECD, 2018). Knowledge Building (KB), as an educational and CSCL model, aims to acculturate school-aged students to work as scientists and engage in scientific, disciplinary, and dialogic practice for knowledge advancement (Scardamalia & Bereiter, 2014). Students worked collectively, engaging in idea-centered progressive discourse to post ideas, generate questions and explanations, and reflect on and revise theories, supported by Knowledge Forum (KF), a collaborative discussion workspace developed to support progressive inquiry. It is widely believed that learning analytics affords students opportunities to engage in reflective assessment processes, particularly in a knowledge-building community. There is a body of empirical evidence indicating the significant impact of analytics-supported reflective assessment approaches on students' metacognition, productive discourse moves, and collective learning (Lee et al., 2016). Engaging in an analytics-supported reflective assessment approach involves reflecting on the state of knowledge-building efforts, using analytics to support that reflection from multiple perspectives. This study designed an analytics-supported reflective assessment environment augmented with the visualization of KF ideas-building networks, to examine and scaffold the development of students' epistemic understanding of the nature of discourse and progressive inquiry. Specifically, two research questions were addressed: (1) What characterized students' epistemic understanding of discourse, and what change over time? (2) How did students with high and low level epistemic understanding engage in the KF inquiry differently?

Methods

Pedagogical design and data sources

Thirteen Grade 11 students studying visual arts participated. The key design focused on students' collective meta-reflection of KF work using analytics of "ideas building" (1) *Cultivating a knowledge-building classroom culture* (Weeks 1-2). (2) *KF inquiry and collective meta-reflection* (Weeks 3-4). (3) *Deepening inquiry with collective reflection and coordinating of ideas for rise-above* (Weeks 5-6). Students deepened their inquiry on KF following reflective assessment. Various data were collected including (1) Students' pre- and post-test epistemic understanding of discourse. (2) KF discourse. (3) Students' reflective assessment journals.

Data analysis and results

RQ1. What characterized students' epistemic understanding of discourse?

Students' responses on epistemic understanding of discourse were analyzed using a three-point scale ranging from simple to more sophisticated understanding of discourse towards a knowledge-building approach. A second rater coded 30% of the data, K = .83 (Cohen's Kappa), indicating good inter-rater reliability. Paired sample t-test indicated that there is a significant change from pre-test to post-test, t (12) = - 9.815, p < .001.

RQ2. How did students with high level epistemic understanding engage in the KF inquiry differently?

The second research question examined how students contributed to and engaged in KB discourse. We used KBDeX, an analytic tool that uses social network analysis techniques employed in KB research to examine students' collective involvement in KB discourse (Oshima et al., 2012). KF discourse was exported to KBDeX and produced three analysis networks - students, discourse, and keywords. This paper examined how students with different levels of epistemic understanding performed differently in their KF inquiry using the keywords network. Students were divided into high- and low-level groups based on their post-test epistemic understanding of discourse. As Figure 1 shows, analysis of the keywords network and comparison of the two groups suggested that students with higher epistemic understanding engaged more productively in the progress discourse by integrating key ideas for advancing collective knowledge (words highlighted in red represent the keywords used by the students in the KF discussion).

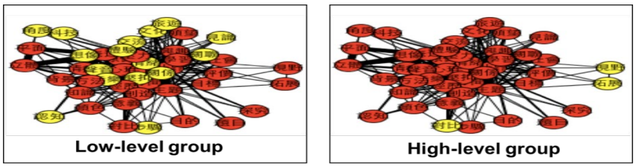


Figure 1. Visualization of keywords network between high- and low-level groups

Further to the social network analyses, content analysis was conducted to examine how students engaged in productive KB. Students' writings in KF were parsed into inquiry threads based on the conceptual problems (Zhang et al., 2009). Individual notes were used within each inquiry thread as the unit of analysis and coded using a theory- and data-driven coding scheme (Chuy et al., 2011; Hakkarainen, 2003). A second rater coded 30% of data, and the inter-rater reliability was .94 for questioning, .83 for theorizing, and .89 for community. Paired sample t-tests were conducted to examine whether students engaged in more productive discourse over time. Significant changes were obtained from Phase 1 to Phase 3, in terms of sustained inquiry, t (12) = - 2.250, p<.05; supporting an explanation, t (12) = - 2.551, p<.05; and connection, t (12) = - 2.245, p<.05. The results suggest students engaged in asking explanation-seeking questions and proposing ideas in Phase 1, sustained their inquiry by asking further questions and supporting their ideas with elaboration and examples in Phase 2, and reflected on the state of knowledge-building efforts in Phase 3. These analyses suggested the role of the designed analytics-supported reflective assessment environment in supporting students' productive discourse engagement.

Conclusion and implications

This study sheds light on the use of analytics and reflective assessment approaches to scaffold students' development of an epistemic understanding of the nature of discourse and knowledge building. Analysis of KF work using KBDeX indicated how students with different levels of epistemic understanding of discourse performed differently in their KF discussion. Content analysis of KF discussions suggested how the analytics-supported reflective assessment journal supported students' engagement in productive discourse progressively over time. In sum, this study is particularly important, as few studies have focused on how students' epistemic understanding of the nature of discourse can be scaffolded using analytics-supported reflections to help them engage in productive discourse.

References

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*, 57-87.

OECD (2018). *Science, technology and innovation outlook 2018*. OECD Publishing, Paris.

Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse. *Educational Technology Research and Development, 60*(5), 903-921.

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In R. K. Sawyer (Eds.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 397-417). New York, NY: Cambridge University Press.

Supporting Learning Interaction in a Distributed Learning Environment with Tangible User Interfaces

Yanhong Li, Beat Rossmy, Heinrich Hussmann
yanhong.li@ifi.lmu.de, beat.rossmy@ifi.lmu.de, hussmann@ifi.lmu.de
LMU Munich

Abstract: Tangible User Interface (TUI) brings an embodied learning experience and provides a good solution for collaborative learning in the distributed learning environment, such as the Interconnected Smart Classrooms (ISC) in this study. In the ISC, multiple classes are interconnected with a single teacher in one classroom and teaching assistants in the distributed classrooms. Students sat and studied in groups. However, ISC had the problems for learning interaction (1) between students, teaching assistants, and the teacher; (2) within group and between groups; (3) within classroom and between classrooms. Our study will: (1) analyze the requirements of learning interaction in the ISC; (2) design and implement four TUI prototypes to support above interactions; (3) discuss how to design TUIs for collaborative learning in a distributed learning environment.

Introduction

While our overall goal is to design TUIs to support learning interactions in the Interconnected Smart Classrooms (ISC), this paper provides the following contributions: (1) an analysis of the learning interaction requirements in the ISC; (2) four initial prototype concepts which aim to support these interactions; (3) a discussion about how to design TUI support learning in the distributed learning environment.

During a practical university class, 19 master students supervised by a team of 3 HCI researchers developed 4 tangible prototypes for ISC. The developments followed an iterative design process to generate insights based on the research through design approach. Final prototypes have physical functions, which include the required casing, sensors, actuators and electronics.

stayFOCUSed (see Figure 1) is a TUI that uses light projection on the ceiling and light-feedback on the device to support ISC learning activities. Group work is supported by light-feedback on the device that indicates the remaining time via a progress bar in traffic light colors (see Figure 1-1). To uncover the voting the light beam is focused via rotation of the projector lens. Subsequently, students can discuss the outcome of the poll. Colored disks are used to communicate group work status (green = finished, red = help) to other groups and the TAs (see Figure 1-1). Empty disks can even be used to write and share information freely (see Figure 1-4).

Figure 1. *stayFOCUSed* experience prototype (1. Prototype structure; 2. Progress bar in traffic light; 3. Rotate the projector to show answers; 4. Hand-write in the disk)

Group Hexagon (see Figure 2) is a modular TUI that supports different ISC learning activities. Via a secondary smart-device, the teacher can change the working mode of *Group Hexagon* (see Figure 2-5). Each group has one group-hexagon and six individual hexagons (see Figure 2-5). The individual hexagons are used in the detached mode by the students to pick answer options (see Figure 2-1) or if connected to the group-hexagon to show solutions of working tasks (see Figure 2-5). For the interaction with *Group Hexagon*, touch gestures are used for selection tasks and miscellaneous interaction (see Figure 2-1).

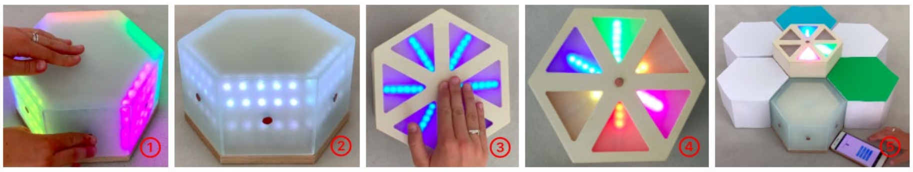


Figure 2. *Group Hexagon* experience prototype (1. Choose the answer with individual-hexagon; 2. Remaining time shown in the individual-hexagon; 3. Help seeking with group-hexagon; 4. Answer distribution shown in group-hexagon; 5. Teacher chooses the working mode)

Tower (see Figure 3) is designed to show the interactions both within and across groups. By placing magnets on the outer grid on the device surface, students can participate in voting. Different colored magnets are used to indicate students' certainty regarding their answers (see Figure 3-2). The rows of the grid demonstrate the response options and the columns represent the individual group member's work space. The top of the *Tower* is used for seeking help and signalizing the working status (see Figure 3-3). For communication and interaction with other groups such as (1) call for help from peers or provide them help, (2) rate your own or other groups' work, or (3) participate on discussions an App on students' personal mobile devices is used (see Figure 4-1).

Figure 3. *Tower* experience prototype (1. Discuss with other groups through App; 2. Place magnets on the *Tower* to choose an answer, green is "I'm confident", white is "I'm not sure"; 3. Rotate top bulb for help; 4. Touch top bulb to show finished)

Glowing Wand (see Figure 4) is a personal handheld TUI which is used by students to participate in ISC learning activities. It is modeled after a magic wand and thus motion gestures are used to control *Glowing Wand*. Different gestures indicate to change *Glowing Wand*'s color, whereas the inclination regulates its brightness (see Figure 4-1). The combination of color and brightness communicate the current working state of the student or can be used to quickly get an overview of the participants opinions in voting situations. Simple gestures that are considered to be broadly understandable and associated consistently are mapped to the traffic light color schemes. This system fits well into ISC learning activities, but can be used in self-defined cases or group processes such as voting due to its open design and tool character.

Figure 4. *Glowing Wand* experience prototype (1. Gesture designs; 2. Switch for a rainbow feedback; 3. Negative tick gesture to red light; 4. Circle gesture to yellow light; 5. Tick gesture to green light)

Conclusion

The study aims to support learning interactions with TUIs in the ISC learning environment: (1) communication in diverse learning contexts: group (within and across) and classroom (the same and different); (2) group process which contain different interactions (student-student in the same group, student-student in different groups, student-student in different classrooms, student-TA, student-teacher) at the same time. Four TUI experience prototypes were developed to: (1) support the learning interaction within group, inter-group, with teacher, with TAs, and across the classroom; (2) support learning activities in the ISC. The discussions, such as *how to design TUIs for learning* and *TUI for ISC: Closed or open* provide an insightful perspective for future study. As a unique distributed learning environment, ISC is a unique learning environment which contains the interactions among different users at the same time. The study is a good example to show how to provide a TUI solution for collaboration learning in the technology supported learning environment.

Acknowledgments

This study is funded by the Elite Network of Bavaria (Project number: K-GS-2012-209). TUI prototypes in the study were developed by master students who took our Practical Tangible Light class (SoSe 2020) at LMU Munich. *Tower*: Havy Ha, Mariam Ali Hussain, Jennifer Herner, Vanessa Balzer; *Group Hexagon*: Sybil Bast, Anna-Carina Gehlisch, Laura Haller, Ina Klautke; *stayFOCUSed*: Aleksa Ristic, Sarah Muser, Agnes Reda, Tianyang Lu; *Glowing Wand*: Anke Pellhammer, Jonathan Haudenschild, Sandra Wackerl.

Author Index

© ISLS

International Society of
the Learning Sciences