Computers in Human Behavior 98 (2019) 267–276

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

The effects of transformative and non-transformative discourse on individual performance in collaborative-inquiry learning

Wanli Xing[a,*], Vitaliy Popov[b], Gaoxia Zhu[c], Paul Horwitz[d], Cynthia McIntyre[d]

[a] Department of Educational Psychology & Leadership, Texas Tech University, Lubbock, TX, 79409, USA
[b] Jacobs Institute for Innovation in Education, University of San Diego, San Diego, CA, 92110, USA
[c] Ontario Institute for Studies in Education, University of Toronto, Toronto, ON M5S 1V6, Canada
[d] The Concord Consortium, Concord, MA, 01742, USA

ARTICLE INFO

Keywords:
Collaborative-inquiry learning
CSCL
STEM
Learning analytics
Discourse analysis

ABSTRACT

The effectiveness of computer-supported collaborative inquiry learning in STEM education is well-documented in the literature. At the same time, research indicates that some students struggle to articulate relevant concepts, to make their reasoning explicit, and to regulate their learning—all of which are necessary for effective collaboration. In this study, 106 college students completed tasks related to Ohm's Law in a simulation-based, collaborative-inquiry learning environment. Using qualitative analysis, multilevel modelling, and data-mining techniques, we investigated the relationship between student engagement in transformative and non-transformative learning processes and learning outcomes. The results revealed that by using the appropriate feature engineering and algorithms, we could build accurate machine-learning models that could automatically identify transformative and non-transformative discussions on a large scale. Additional qualitative and quantitative analyses indicated that when groups engaged in additional interpretation and sustained mutual understanding, their members tended to have statistically better individual-learning outcomes. These analyses also indicated that when groups engaged in additional orientation and proposition generation, their learning outcomes were statistically lower. Approximately two-thirds of the students considered their group work helpful in completing inquiry tasks. Explanations of these results and research recommendations are provided.

1. Introduction

A significant body of research has revealed the effectiveness of using computer networks to promote collaborative learning in STEM education—for a review of this research, see Jeong, Hmelo-Silver, Jo, and Shin (2016). In Computer-Supported Collaborative Inquiry Learning (CSCIL), two or more learners use computers to collaboratively solve problems (e.g. problems in electrical engineering) and to co-construct knowledge (e.g. knowledge of the relationship between voltage and resistance), typically in a simulation-based learning environment (Stegmann et al., 2007). In doing so, learners may search for information, form hypotheses, experiment, and interpret, articulate and share ideas (Stahl, Koschmann, & Suthers, 2006, pp. 409–426). Research into collaborative-inquiry learning has consistently shown, however, that most students struggle to regulate their learning and to articulate the concepts required to make their reasoning explicit and to actively make sense of the subject matter—e.g. through hypothesis generation and data interpretation (Bell, Urhahne, Schanze, & Ploetzner, 2010; Gijlers & de Jong, 2009; Lazonder & Rouet, 2008; Popov, Xing, Zhu, Horwitz, & McIntyre, 2018). A key question in CSCL research—and in collaborative-inquiry learning more generally—is why some groups outperform other groups (Kapur, Voiklis, & Kinzer, 2008; Suthers, 2006).

This study seeks to identify practices that allow students completing inquiry tasks to reap the full benefits of their collaboration. According to Gijlers and de Jong (2009), when students complete typical inquiry tasks, they engage in transformative learning activities (i.e. activities directly related to knowledge construction) and regulative activities (i.e. activities necessary to coordinate and maintain mutual understanding). Transformative learning processes include *orientation, hypothesis generation, experimentation,* and *the drawing of conclusions* (Njoo & De Jong, 1993). Regulative learning processes include: *sustaining mutual understanding, planning,* and *monitoring* (De Jong, 2006; Gijlers & de Jong, 2009). Each learning process has specific objectives and requires specific types of collaboration—i.e. activities in which

* Corresponding author.
E-mail addresses: wanli.xing@ttu.edu (W. Xing), vpopov@sandiego.edu (V. Popov), gaoxia.zhu@mail.utoronto.ca (G. Zhu), phorwitz@concord.org (P. Horwitz), cmcintyre@concord.org (C. McIntyre).

https://doi.org/10.1016/j.chb.2019.04.022
Received 2 June 2018; Received in revised form 24 April 2019; Accepted 27 April 2019
Available online 29 April 2019
0747-5632/ Published by Elsevier Ltd.

information is transferred among participants—to be effective. In group problem-solving process, coordinating collaborative activities and appropriately distributing group efforts and resources are critically important (Rummel & Spada, 2005; Saab, van Joolingen, & van Hout-Wolters, 2012). Moreover, learning is more likely to occur when students engage in meaning-making processes, such as the use of prior knowledge to formulate hypotheses and the generation of new insights from experiences in the learning environment (Gijlers & de Jong, 2009).

There has been a push over the past two decades to unpack group processes and, in particular, to understand the complexities of interactional dynamics and their influences on group and individual performance (McKeown et al., 2017; Suthers, 2006; Popov et al., 2018). While previous CSCL studies—e.g. Kwon, Liu, and Johnson (2014) and Lee, O'Donnell, and Rogat (2015)—have generally taken a qualitative approach or relied on relatively small samples, this study investigated a relatively large number of groups and their collaborations. The intellectual merit of this paper stems from two unique contributions. First, it builds and expands upon previous research on collaborative-inquiry learning by examining the behavior patterns and transformative and non-transformative learning processes that are important for success when students engage in a simulation-based, collaborative-inquiry learning environment. Second, this study offers a practical example of applying large-scale CSCL analysis while maintaining human insights. We began by using human qualitative coding to identify transformative and non-transformative discussions, and we then used machine-learning algorithms to capture the human intelligence for automatic coding. Importantly, if group processes that improve individual learning are identified, teachers could improve student learning by encouraging specific types of activities.

2. Literature review

2.1. Inquiry learning

Although the existing literature on inquiry learning is extensive and considers various aspects of this form of instruction, a recent meta-analysis and literature review—Lazonder and Harmsen (2016)—identified two main strands of research. Studies into inquiry learning have focused mainly on 1) how different groups of learners (typically different age groups) engage in inquiry tasks and which processes promote successful learning, and 2) the effects of various types of guidance on learners' processes and outcomes (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011; Horwitz, Gobert, Buckley, & O'Dwyer, 2010; Linn, Lee, Tinker, Husic, & Chiu. 2006; Plass et al., 2012).

There are also two divergent attitudes towards inquiry learning: one is that inquiry learning could be very challenging for learners (especially young learners), and the other is that inquiry learning is a beneficial form of instruction. Kirschner, Sweller, and Clark (2006) argued that inquiry learning is so demanding that it uses up a learner's limited cognitive capacity and makes it difficult for new information to be stored in their long-term memory. Hmelo-Silver, Golan Duncan, and Chinn (2007) argued that inquiry-based methods are powerful approaches to learning, emphasising the importance of scaffolding in the learning process. In general, the focus has shifted from whether to use inquiry learning to the conditions under which it can benefit learners as much as possible.

Njoo and De Jong (1993) were among the first researchers to explore the learning processes that constitute inquiry learning. Specifically, Njoo and De Jong (1993) made the initial distinction between transformative processes (those that are directly related to knowledge construction) and regulative processes (those that are deemed necessary to complete the inquiry learning task). This distinction was made based on the overview of inquiry learning processes germane to discovery or exploratory learning with a computer simulation. Njoo and De Jong (1993) conducted a study in which they examined the learning processes (i.e., identifying variables, generating hypotheses, designing an experiment, predicting, interpreting data) of 91 students of mechanical engineering working on a simulation on control theory. They looked at how students engaged in the process of generating a hypothesis in terms of identification of variables and establishing a relation between them.

As stated above, transformative learning processes include *orientation, hypothesis generation, experimentation,* and *the drawing of conclusions* (De Jong & Van Joolingen, 1998; De Jong, 2006). During orientation students orient themselves towards the learning task, establish initial understanding of the problem at hand and activate their prior knowledge in relation to the variables and conditions. During hypothesis generation, students formulate a statement or a set of statements concerning the relations between the variables or certain values. Regarding experimentation, students typically begin with a period of designing an experiment, predicting its outcome, running the actual experiment. In drawing conclusions students strive to confirm or reject or reconsider their initially formulated hypotheses and come up with viable solution to the problem by weighing all options using the experimental data they gathered during the experimentation stage. These processes can be repeated and re-entered, i.e., students may go through the processes again or return to them if necessary.

Regulative activities, conceptualized as the application of metacognitive skills, are needed to regulate the transformative processes in both efficient and effective ways (Saab et al., 2012; Zheng, Xing, & Zhu, 2019). By engaging in regulative learning processes, students organize their work, devise a plan and continuously monitor progress and achievements.

From the field of inquiry learning the distinction made by Njoo and De Jong (1993) was then applied in a collaborative setting by Gijlers and de Jong (2009). Over the last decade, this framework has been utilized in a number of studies investigating a range of collaborative inquiry learning settings, tools and scenarios (e.g., Janssen & Bodemer, 2013, awareness tools; van Joolingen, de Jong, Lazonder, Savelsbergh, & Manlove, 2005, on collaborative discovery learning with the Co-Lab environment; Saab et al., 2012, on task and team regulation). There are several fundamental characteristics that distinguish the process of problem solving done by an individual from a collaborative work that involves multiple people working interdependently toward a common goal (Graesser et al., 2018). Most collaborative problem-solving frameworks (see Organisation for Economic Co-operation and Development, 2017 for review) indicate that the efficacy of collaborative learning effort is influenced a) by the extent to which students can ensure the consistency of the joint work product by coordinating their collaborative activities (Popov et al., 2018; Erkens, Jaspers, Prangsma, & Kanselaar, 2005), and b) by the extent to which students can take appropriate actions to solve the problem, establish and maintain shared understanding (Kirschner, Beers, Boshuizen, & Gijselaers, 2008; Stahl, 2013). In the collaborative problem-solving literature these are also sometimes called the teamwork and task work processes (Graesser et al., 2018). In the next section, we elaborate on the role of collaboration in the inquiry-learning process.

2.2. Peer collaboration as a source of support in the inquiry-learning process

Inquiry tasks could be highly challenging for learners if they are not sufficiently supported (De Jong & Van Joolingen, 1998; Makitalo-Siegl, Kohnle, & Fischer, 2011). Typical sources of support include teachers, small-group scripts (which provide scaffolding to guide students in deciding what to do, what roles to play, and what sequence of activities to perform during a learning task), experts (Mäkitalo-Siegl, Kohnle, & Fischer, 2011), and peers (Okada & Simon, 1997). In this study, we focused on peer support, exploring how learners find support by engaging in collaborative learning tasks. Okada and Simon (1997) found that students who worked in dyads during a molecular-biology inquiry-learning task outperformed students who worked alone, concluding that dyads can formulate greater numbers of alternative scenarios.

Piaget's (1926) ideas about sociocognitive conflicts and Vygotsky's (1978) "zone of proximal development" help explain why scaffolding—which can be provided by peers, teachers, or other instructional sources—can effectively foster students' cognitive abilities. A good scaffold can link a student's current understanding to the learning context. For example, a student may be good at explaining how chemical reactions relate to energy but have limited understanding of how to interact with dynamic visualisations (Gerard, Tate, Chiu, Corliss, & Linn, 2009). The student's partner, in contrast, might have expertise in constructing models but struggle to design consequential experiments. With peer scaffolding, each collaborative or knowledge-integration unit will enable students to take advantage of the expertise of their peers.

The Piagetian approach to socio-cognitive conflict claims that the efficacy of a collaborative-learning effort is influenced by the extent to which the students can identify and discuss conflicts in their knowledge and beliefs (De Lisi & Golbeck, 1999). For instance, participatory simulations may generate a type of collaborative learning in which each student's experience and contributions contribute to the collective understanding of the whole system that the students are studying. Collaborative learning efforts encourages students to make their thinking visible by explaining their reasoning to their partners in collaboration. Students working in groups reach shared understandings through the negotiation of meaning, which requires students to ask questions, have discussions, explain their thoughts and ideas, and support their viewpoints with additional information (De Lisi & Golbeck, 1999). Previous research supports the claim that learning is most likely to occur when students build on each other's reasoning by critiquing, challenging, and synthesising each other's contributions. This form of discourse triggers cognitive activities that stimulate knowledge construction (Andriessen, Baker, & Suthers, 2003).

This study explores a series transformative and non-transformative learning processes in a collaborative setting: orientation, proposition generation, experimentation, interpretation and conclusion, regulation, and the sustaining of mutual understanding (see Table 1). When orienting, collaborating students elicit ideas to become aware of their views of the situation, add new ideas to fill in missing information and to make sense of the topic, and strategise about how to approach and solve the problem at hand together. In proposition generation, students typically form a statement or a set of statements concerning the relations between the variables or certain values. Previous studies have devoted particular attention to the proposition-generation process and sought ways to support this process. This process is important because it prompts students to use their prior knowledge to connect the variables presented in a problem to explain the phenomena. During experimentation, students design, test, and run experiments and make sense of the outcomes. This process occurs very quickly in computer-assisted, simulation-based learning environments. When students conclude, they review their propositions/hypotheses by comparing them to their experimentation data/experiences (Gijlers & de Jong, 2009). In regulation, students allocate time to completing the tasks, discuss the major procedures required to complete the tasks, and so forth. Another important contributor to collaborative learning is the ability of group members to sustain mutual understanding throughout the entire process.

Learners are required to coordinate both at a teamwork level (in terms of managing time, group relationship and task activities) and at a task work level (the cognitive problem-solving aspects– "transformative") during discourse (Popov et al., 2018; Graesser et al., 2018). Coordination at these two levels may ensure that collaboration is both efficient and effective. These aspects of collaboration have separately been related to quality of group processes and outcomes; however, transformative form of discourse at a task work level has been suggested to give rise to cognitive activities that stimulate knowledge construction (Kirschner, Sweller, Kirschner, & Zambrano, 2018; Stahl, 2013). At the same time, regulative or non-transformative inquiry processes such as monitoring and coordinating, can support collaborating students to regulate the use and take advantage of the transformative processes (Njoo & De Jong, 1993).

We are interested in exploring the effects of transformative and non-transformative discourse on individual performance in collaborative-inquiry learning. To productively engage in a collaborative-inquiry task, students must spend enough time on and put sufficient effort into orientation, proposition generation, experimentation, interpretation and conclusion, regulation, and the sustaining of mutual understanding (Janssen & Bodemer, 2013; Lazonder & Harmsen, 2016). For example, how a group of collaborative partners go about reaching consensus or discussing their differences in opinion may vary and may ultimately affect their knowledge acquisition. When working on tasks together, collaborating students may not go beyond orientation phase and/or rush to experiments without identifying variables or establishing a relation between them, or may draw conclusions without sufficient experimental evidence (Gijlers & de Jong, 2009; Lazonder & Rouet, 2008). Groups may also suffer from processes losses resulting from coordination difficulties, which can hamper the students to regulate the use of the transformative processes.

Table 1
Example of "transformative" and "non-transformative" coding.

Contribution type	Category	Description	Example
Transformative (utterances related to knowledge construction)	Orientation	Attempted to understand the structure and the complexity of individual and group goals by collecting goal-related information, such as the current values of the resistance or voltage.	r = 560 here I need 3.5 boys whats your R values now mines 330 total 11.7 for us I think
	Proposition generation	Formed a statement or a set of statements concerning the relationships between the values of resistance or voltage.	we need the v drop to be two across each one, ok so we all have to be 100
	Experimentation	Included testing their propositions and adjusting the resistors by making them lower or higher.	bear drop down one; let me get 1v hold up; let me readjust; put in a higher resistor; mike put yours at 70 Ω; let me get mine to 2v 1st, and then the 2nd person goes
	Interpretation and conclusion	Reviewed the proposition in light of the experimentation outcomes.	the bigger my resistor, the higher my voltage is; 150/.000671 = 149.03
Non-transformative (utterances related to technical features, time/group/task management, etc.)	Regulation	Managed time, group dialogue, the big procedures of completing tasks, and so forth.	Who opened the circuit; who is paper; we should start with the third guy in our group and just hope for the best; don't touch anything for a min
	Sustaining mutual understanding	Group members indicated shared/different understandings.	Not me; thanks; OK; Yes; I think we do too; no; done; yes I think so
	N/A	None of the six categories above were applicable.	Eassssyyyyyyyyy; hungry; these codenames have me feeling like I'm in Narcos

2.3. Research questions

The goal of the paper is to understand how the transformative and non-transformative discourse takes place at the group level and further impacts on the individual learning in collaborative inquiry learning environment. Correspondingly, two research questions were proposed as following:

1. Which transformative and non-transformative learning processes occur in more and less successful groups in collaborative-inquiry learning environments and how?
2. How do transformative and non-transformative processes impact the individual learning achievement in collaborative-inquiry learning environments?

3. Methodology

3.1. Participants

144 students were recruited from five two-year-college electronics classes in the United States. The students were instructed by teachers from their respective institutions. The classes were not reorganized; the students worked through the study in their original classes. In each class, the teacher first instructed the students to sit one-per-computer and then randomly assigned the students to different groups. Each student was given a class code so that they could enter the same class space as did their classmates. Each student was assigned a fake name, so the students could not know who their team members were unless they discussed the issue in the chat window.

106 students in 40 groups worked on the teamwork tasks and the post-survey; the other students either worked by themselves, did not participate in the teamwork tasks, or did not fill out the post-survey. Four of the 106 participants did not fill out the pre-survey, so their demographic information was not collected. Of the 102 participants whose demographic information was reported, 18.2% were female, 81.8% were male, 4% were American Indian or Alaska Native, 8% were Asian, 8% were Black or African American, 1% were Native Hawaiian or Other Pacific Islander, 66% were white, and 17 were “Other” (students could belong to multiple races). 43% utilized the free lunch program, while the others did not.

3.2. Environment and setting

This study used an online collaborative-inquiry learning environment called the “Teaching Teamwork,” which included a database of interactive STEM activities. The students worked in the “electronics” domain, which was designed to help them understand and apply Ohm's Law by exploring the relationship between resistance and voltage in series circuits. As Fig. 1 shows, the virtual “electronics” environment included a series circuit with supply voltage E and external resistances R0, R1, R2 and R3; the conditions and goal of the task; a digital multimeter (DMM) with black and red probes that students could use to measure the voltage, current, or resistance of the resistor; a calculator that would appear if the student clicked “calculator”; and a chat window that allowed the group members to talk about their goal, discuss how to complete their tasks, monitor their progress, and so forth.

The “electronics” domain was divided into five tasks of increasing complexity. Instead of requiring the students to complete tasks with full complexity from the very beginning, the domain gradually increased the complexity of the tasks to avoid overwhelming the students. The students were free to start at any level and to move back and forth between different levels. In practice, however, the students usually started with Level B and then moved on to Level C and to other, more complex tasks. Often, they did not move on to the next task until they had completed their current task, or they had difficulty in solving the current task. In Level A, both the E and R values were given. R was equal to zero, and the goal voltage was the same across R1, R2, and R3 (Level A was removed in the formal implementation of this study because it was too easy). The conditions in Level B were the same as those in Level A, except R equalled R1, R2, and R3 instead of zero. In Level C, both the E and R values were given, R did not equal zero, and the goal-voltages across R1, R2, and R3 were different. In Level D, E was unknown, R was given and did not equal zero, and the goal voltages across R1, R2, and R3 were different. In Level E, both the E and R values were unknown, and the goal voltages across R1, R2, and R3 were different. The log data confirmed the increasing complexity of the four tasks the groups worked on because none groups succeeded in solving a more complex task unless they solved all the previous easier ones.

3.3. Procedures

Each class participated separately in the study. In each class, the teacher first instructed the students to sit one-per-computer and then randomly assigned them to different groups. The members of a team were kept separated from one another and were only allowed to communicate via the computer-supported chat window. The group size was three in general. The unmatched students either worked in “solo” condition wherein they manipulated all three resistors themselves, or they did not fill the post questionnaire. These students were not included in the analysis. The teachers did not offer support to the students when they were working on their group tasks, and the researchers did not engage during implementation. This suggests that data collected reflected groups' natural online collaboration when they depended on one another to solve tasks.

Each session lasted for about 90 min and proceeded as follows: consent-form collection, pre-survey, four levels of teaching-teamwork tasks, and post-survey. The students first read and signed the consent form, which was not collected until the end. The students were then given approximately 10 min to fill out the pre-survey. Next, the students spent 60 min working on the teamwork tasks. When there were 10 min remaining, the students were reminded to stop their tasks and complete the post-survey.

The pre-survey included 13 items to collect demographic information (e.g. information on the participant's gender, age, race, and native language). The students joined their assigned online groups to complete the four tasks collaboratively. The electronics tasks began with textual slides introduction that explained the design of the tasks, how to get started, how to use the online chat window and the on-screen calculator, and how to submit results. The slides were embedded in the Teaching Teamwork web page. The students were also shown an introductory video to familiarise them with how the system worked and how to operate the tools. Then, the students in each class joined in groups they were assigned to by their teachers. Since students from the same groups might not sit next to each other, they could not talk to each other orally. The students selected fake names when joining in their groups, and ideally, they did not know who their group members were. All the tasks were mainly about achieving the correct voltage across the resistor an individual controlled by adjusting the resistance in a series circuit. Each student could control her stimulation, but meanwhile, they could not achieve their individual goals if other group members did not make the correct resistance since the current of a series circuit was influenced by the total resistance. The students had to collaborate with each other through the chatting window to achieve their goal. The students in each group initially attempted to complete the Level B task by working together to understand their goal, making plans to complete the task, doing experiments with the simulation, interpreting the results, and regulating the process. The students did not get support from their teachers or researchers when worked on the tasks. The main supports they got were text and video introductions and peer support through chatting box.

When a task was completed, the group was directed to the next-level task automatically. However, they did not have to follow the sequence.

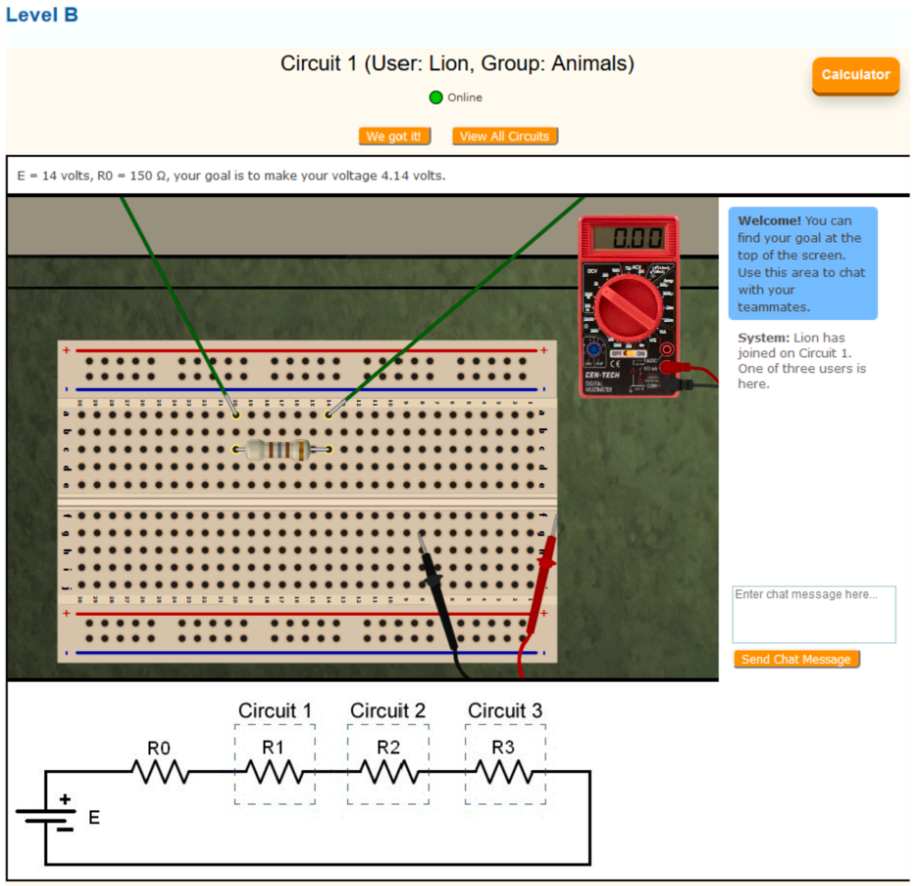

Fig. 1. Screenshot of a simulation of Level B task.

The post-survey consisted of six items to evaluate the students' cognitive understanding of Ohm's Law. Cronbach's alpha for the post-survey items reaches 0.72, and it can be considered acceptable according to van, Neudecker and Nel (2000). Some examples are as following:

What is Ohm's Law?

A circuit contains two resistors, R1 and R2, in series. The resistance of R1 is greater than the resistance of R2. Check all of the following statements that are true.

Use the schematic below to answer this question. Select the formula that gives the total resistance of this circuit.

Five of the questions were radio button questions for which students get one point if they selected the right answer. They were not taken away any point even if they selected wrong answer. One of the questions was multiple choice question with two correct questions. The student would get one point for selecting one of the correct answers and two points for selecting both correct answers. In the best case, they would get two points from this question. However, they would be taken away half point if they selected any one of the correct answers. In the worst case, they would get minus two points from the question.

3.4. Analysis

To answer the two research questions, a perquisite is to identify all the transformative and non-transformative utterances in the group discourse. Given the large scale of the chats, text classification was used to learn the human intelligence to automatically identify different transformative and non-transformative categories. Then for RQ 1, we used the overall frequency and qualitative case analysis to differentiate the transformative and non-transformative chat message of the successful and less successful groups. For RQ 2, multilevel modelling was used to quantify the effect of the transformative and non-transformative discourse on individual learning achievement.

3.4.1. Transformative and non-transformative discussion detection and analysis

Most previous research into the communication that occurs during small-group learning has relied on the manual coding of small samples of messages (Kwon et al., 2014; Lee et al., 2015). These qualitative- and content-analysis techniques are impractical, or at least would have been difficult to perform given the thousands of lines of chat data obtained in our study. Previous research has demonstrated, however, that it is possible to automate the analysis of conversations in CSCL (Mu, Stegmann, Mayfield, Rosé, & Fischer, 2012; Rosé et al., 2008). In this study, machine-learning models were built to automatically identify transformative and non-transformative messages in small group collaborations. Machine-learning models use statistical procedures to map sets of input features to targeted categories (Wang, Kraut, & Levine, 2012; Xing & Gao, 2018; Xing, Goggins, & Introne, 2018). Our inputs were various kinds of features, and our outputs were values representing the categories to which the utterances belonged. Each utterance was categorised as either "transformative" or "non-transformative."

Our machine-learning model was constructed in three steps. First, two senior researchers with extensive backgrounds in CSCL independently coded 1,111 pieces of utterances generated by 15 groups as either "transformative" or "non-transformative." These messages were randomly selected from the dataset generated by all the participants with task as the unit. In total, there were 5,521 pieces of utterances in the dataset we analyzed. A sample of this coding is presented in Table 1. The Cohen's kappa between the two researchers was 0.721, indicating intermediate to good reliability (Fleiss, 1981). The remaining disagreements were discussed and resolved to further improve the

Table 2
The sample feature set.

Regex Features:
Need (to be) + number; getting + number; number + away; Take …, divide … by …, E is + number; number + v set to number + ohm; r = number and number + volts.
LIWC Features:
Summary Dimensions: word count, sentence count, tone
Punctuation marks: period, comma colon, exclamation point, dash, quote, etc.
Function words: pronounce, article, adverb, negate
Other Grammar: verb, adjective, compare, number, quantity, etc.
LDA Topic Features (LDA):
Affect: anger, sadness
Social: family, friend, female, male
Percept: see, hear, feel
Relative: motion, space, time
Informal: swear, assent

consistency between the two coders. Second, three types of textual features (examples of which are shown in Table 2) were used to capture the different linguistic strategies employed by the students as they collaborated. The first type of textual features was regular-expression ("regex") features, which captured key elements of high-order concept propositions. The researchers identified these features through qualitative insights into which words and number combinations were more likely to belong to which category— "transformative" or "non-transformative." All of the regex features were extracted directly from the chats. The second type of textual features were functionally and linguistically relevant words from the Linguistic Inquiry and Word Count (LIWC) program (Pennebaker, Boyd, Jordan, & Blackburn, 2015). The third type of textual features were topical features. It was through the Latent Dirichlet Allocation (LDA) topic model to create specialised, small-group-collaboration-related dictionaries. Then, each LDA topic feature computed the number of words in a sentence that matched an entry in its corresponding topic-dictionary. Third, four different kinds of classic algorithms were used to optimise the machine-learning model: Naïve Bayes, Logistic Regression, Support Vector Machines (SVM), and Decision Tree. For detailed descriptions of these algorithms, one can refer to Kotsiantis, Zaharakis, and Pintelas (2007). The algorithms were evaluated using 10-fold cross-validation to determine their predictive capabilities.

To provide detailed descriptions of what happened in different groups and how the collaborative-inquiry experience might have affected the students' performance on the post-test over Ohm's Law, we conducted a qualitative case analysis. First, we classified the groups into three categories based on how their members performed in the post-test. Groups were categorised as "low score" when all the three of their members scored no higher than 3.5 (the highest scores was 7). Groups were categorised as "mixed score" when some of their members scored high (6 or greater), but others scored no greater than 3.5. Groups were categorised as "high score" when all three students scored no less than 6. Then, two researchers examined the chat histories of the groups in the three categories to identify the excerpts that best represented the components and transitions of the students' contributions. Agreements were reached between the two researchers, and one excerpt was selected for each category. The two researchers manually coded the excerpts using the "transformative–non-transformative" coding scheme.

3.4.2. Multilevel analysis

In social science research, we often face datasets with multilevel structures. In the present study, individual students were nested within groups so that the student-level variables were confounded with the group variables. This clustering effect raised several major statistical problems—e.g. aggregation bias and heterogeneity of regression—-which are difficult to handle using traditional regression approaches. To solve these problems, multilevel modelling was employed, which can account for the non-independence of observations by addressing the variability associated with nesting effects (Du, Xu, & Fan, 2013). Multilevel modelling can more accurately estimate relationships between variables at the individual and group levels for within- and between-group variances (Raudenbush & Bryk, 2002).

In this study, a two-level analysis was employed to determine the impacts of group-level transformative and non-transformative discussions on the cognitive understandings of individual students. Since there were no prior hypotheses regarding the between-group differences in the independent variables, a random-intercept model was used instead of random-slope model. The random part of the intercept was freely estimated to manifest the between-group differences in cognitive understanding, and the slopes of the model were considered fixed. All of the variables were standardized with Mean = 0 and SD = 1 before the multilevel analysis was conducted to increase the interpretability of the regression coefficients.

The first-level outcome variable was students' test scores for cognitive understanding after collaboration. It can be described using notation common in multilevel analysis:

$$Y_{ij} = \beta_{0j} + r_{ij}$$

where Y_{ij} is the test score for student i nested under group j, β_{0j} represents the group intercept, and r_{ij} is the unexplained individual residual.

The second level was a random intercept model of group discussions with "transformative" and "non-transformative" as variables. It can be illustrated as follow:

$$\beta_{0j} = r_{00} + r_{01}(W_1^G) + r_{02}(W_2^G)...+r_{06}(W_6^G) + e_0$$

where W_1^G through W_6^G are group-level variables quantified from the chats, and e_0 is the unexplained group residual. Maximum likelihood was used for model estimation. Before we conducted the complete multilevel analysis, a fully unconditional or "null" model was built to partition the variance in the students' cognitive understandings into between-group and within-group components.

4. Results

4.1. Results of transformative and non-transformative discussion detection and analysis

Using different features and various algorithms, machine-learning

Table 3
Machine-learning model performance.

	Naïve Bayes		Logistic Regression		SVM		Decision Tree	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
Regex	43.0%	40.90%	72.3%	47.24%	72.2%	50.00%	79.8%	53.36%
LIWC	20.2%	32.70%	35.2%	34.31%	41.6%	31.76%	44.1%	43.23%
LDA	38.5%	32.41%	61.8%	35.73%	54.4%	28.01%	52.2%	38.92%
Regex, LIWC	39.1%	41.75%	69.4%	49.99%	64.5%	31.79%	76.0%	50.90%
LIWC, LDA	34.6%	35.60%	61.8%	43.67%	60.2%	30.60%	61.9%	45.22%
Regex, LIWC, LDA	42.9%	38.64%	66.5%	50.51%	60.2%	30.60%	72.6%	53.10%

Table 4
Descriptive statistics of group level variables and students' test scores.

	N	Mean	Median	SD	Min	Max
Orientation	40	67.62	49.0	55.59	1	253
Proposition generation	40	0.70	0	1.30	0	7
Experimentation	40	12.99	11.5	9.61	0	42
Interpretation and conclusion	40	1.27	0.5	1.77	0	7
Regulation	40	6.60	4.0	6.43	0	27
Sustaining understanding	40	11.18	9.0	10.13	0	46.0
Test Score	40	4.06	4.0	1.86	0	7.0

models were built to automatically identify transformative and non-transformative discussion chats. Table 3 shows the performance results for the machine-learning models. The performance of each machine-learning model was measured using precision and recall measures, classical metrics for evaluating supervised models. When the various algorithms and different feature sets are compared, the results reveal that the decision tree with the regex features performed best, reaching a precision of 79.8% and a recall rate of 53.3%. The prediction performance of the machine-learning model that used the decision tree algorithm was comparable to those of other machine-learning models built in similar collaborative contexts (Dalal & Zaveri, 2011). One possible reason for the better performance of decision tree using regex features is that the utterances in our study is relatively short and the data size is not that large. While the regex features are based on human insights with few dimensions, LIWC expands the feature space explosively. Given the data size, the high dimensional feature space from LIWC may compromise the performance of the machine learning algorithms. Topic features from LDA may not be that obvious in such short texts.

The Decision Tree constructed was applied to the remaining chats. Table 4 reports the descriptive statistics of group level variables and students' test scores. The low occurrences of proposition generation and interpretation and conclusion further showed the reliability of the automatic coding system given that in the initial 1,111 utterances, only 45 proposition generation and 16 interpretation and conclusion. By looking at the average percentage of utterances of group members, we found that the students in the successful groups had more equal participation (Range = [0.1%, 81.5%], SD = 13.0%) compared to the students in the less successful groups (Range = [0.1%, 100%], SD = 22.8.0%). We then conducted a multilevel analysis to examine the effects of transformative and non-transformative discourse on the cognitive understandings of individual students.

We conducted a qualitative case analysis to show what transformative and non-transformative learning processes looked like in more- and less-successful groups. This was part of research question 1. As is shown in Table 5, the three students took a while to orient themselves to the assignment and the platform and to understand the individual and group goals. The three students scored 3, 3, and 3.5 on the section in the post-survey on understanding Ohm's Law. These results could be explained by the students' inability to make their reasoning explicit and to articulate concepts that would have enabled them to make sense of the relationship between resistance and voltage in series circuits.

Table 6 shows the excerpts from a group of students with mixed low and high scores on the post-survey. Student D scored 7, student F scored 4.5, and student E scored 3.5 on the section in the post-survey on understanding Ohm's Law, indicating that one of them had a good understanding, one of them had a medium understanding, and one of them had an unsatisfying understanding of Ohm's Law after having worked through the collaborative-problem-solving tasks. The mixed result could be attributed to the fact that student E proposed the right answer but did not explain the law and made decisions for the other two students. Since the other two students only experimented with what student E proposed and did not ask her to elaborate on her reasons, student E's rationale in generating the proposition was never made explicit to the group. As a result, it is not clear if student E understood Ohm's Law or simply provided guesswork. Although the post-survey results indicated that student D had a good understanding of Ohm's Law, she did not get an opportunity to share her knowledge with the other two members before they solved the task.

As is shown in Table 7, the three students communicated their goals, formed statements about the relationship between resistance and voltage and discussed these statements, made plans to solve the problem,

Table 5
Excerpts from a group of students with low scores on the post-survey.

Students	Utterances	Coding
A	idk what were doing	Orientation
B	trying to get the voltage to 2.75	Regulation
B	are all of ours together supposed to make 2.75	Orientation
A	i got 2.70	Orientation
C	me too	Orientation
B	wait we got different resistances?	Orientation
A	how do we know what our voltage is	Orientation
B	Isn't it your E?	Orientation
B	like at the topish of this thing it says E 11	Orientation
B	and I aint trynna do this math tbh	Orientation
A	so we have to change 11 to 2.75 in the top left corner	Orientation
B	Yes	Sustaining mutual understanding
B	Well you got 11 right?	Sustaining mutual understanding
A	Yeah	Sustaining mutual understanding
B	what the equation for the law tbh	Orientation
A	v = iR	Orientation
B	I'm so confused	Regulation
A	Yeah	Sustaining mutual understanding
B	What are we trying to do?	Regulation
B	Okay so find I	Orientation
B	cause if V = IR	Proposition generation
B	Honestly Idk	Orientation
A	i have no clue	Orientation
B	All of you have E = 11 and R = 150 right to start?	Orientation
C	Yes	Sustaining mutual understanding
A	yeah but idk what that means	Orientation

Table 6
Excerpts from a group of students with mixed low and high scores on the post-survey.

Students	Utterances	Coding
D	Does your voltage read 0	Orientation
E	Now it does	Sustaining mutual understanding
E	What do you have your resistor set to?	Orientation
E	2.92 v	Orientation
D	470	Orientation
E	Same here	Orientation
F	My voltage is 4.23v	Orientation
D	Triangle?	Regulation
E	I think if everyone has 470 Ω it will be 3.25	Proposition generation
D	Is everyone reading 3.25?	Orientation
F	No the voltage across my resistor is 3.25v	Orientation
E	I am	Sustaining mutual understanding
F	Now*	Orientation
E	We got it	Regulation
D	Okay so we got it	Regulation

Table 7
Excerpts from a group of students with high scores on the post-survey.

Students	Utterances	Coding
G	Hey guys, I have an idea	Sustaining mutual understanding
H	What is that?	Sustaining mutual understanding
G	Everyone set resistors to 100 Ω and say what voltage you get, so we can determine the other resistors value unless there's a better way	Proposition generation
I	The other resistor is 100 isn't it?	Orientation
G	Oh yeah. Never mind	Sustaining mutual understanding
I	Voltage divider?	Proposition generation
H	I'm thinking so.	Sustaining mutual understanding
I	That only works with 2 resistors though?	Proposition generation
G	It should work with multiple. I know my resistance needs to be 25% of the total to meet my voltage	Proposition generation
G	What about yall?	Orientation
G	We should be able to figure out the resistor values with that	Proposition generation
I	circuit 1 what does your resistance have to be. i mean voltage	Orientation
G	3.5 V	Orientation
I	I am assuming pliers is the same?	Orientation
H	Yes, mine is also at 3.5 V	Orientation
G	What about you wrench	Orientation
I	Same	Orientation
G	So everyone should be 100 Ω	Proposition generation
G	Can everyone try 100 and see what voltages we get	Experimentation
I	They are in series so the R at each resistor should have to be same to have the same voltage drop	Interpretation and conclusion
G	Mine's at the right voltage now	Orientation
H	Okay, I've got mine set to 100 and I got −3.5	Orientation
I	Me too	Orientation
H	Never mind. I had the leads switched, it's at 3.5	Orientation
G	Cool we got it	Regulation
I	I guess we got it	Regulation

and explained their reasoning by stating that "it should work with multiple. I know my resistance needs to be 25% of the total to meet my voltage." In accordance with their plan, they experimented together and completed the task. Throughout the process, they responded to and referred to one another to sustain a mutual understanding, which was essential to solving the group task collaboratively. In addition, they interpreted their results by making it explicit that "they are in series so the R at each resistor should have to be same to have the same voltage drop." Student G scored 7, student H scored 6, and student I scored 6 on the post-survey, indicating that the students shared a good understanding of Ohm's Law after having worked through the collaborative-problem-solving tasks. The students' good understanding could be attributed to some extent to their explicit discussion of the relationship between resistance and voltage in series circuits, which they conducted in making a plan and in reaching a conclusion.

4.2. Results of the multilevel analysis

The null model indicated a variance of 67.7% in the students' cognitive understanding scores at the individual-student level and a variance of 32.5% at the group level. The nested data had an ICC value of 0.32. The deviance was 293.3. According to Kreft, Kreft, and de Leeuw (1998), cluster effects should be controlled by multilevel modelling when the ICCs are as low as 0.02. It was thus important and legitimate to conduct such an analysis.

The complete modelling results are shown in Table 8. "Orientation" (-0.56, $p < 0.05$) and "proposition generation" (-0.32, $p < 0.05$) had statistically significant negative influences on students' cognitive understandings. "Orientation" tended to have a larger effect than "proposition generation." In contrast, "interpretation" (0.29, $p < 0.05$) and "sustaining understanding" (0.53, $p < 0.05$) had statistically significant positive influences on students' cognitive understandings. The influence of "sustaining understanding" was larger than that of "interpretation" and "conclusion." The results showed that "experimentation" and "regulation" had no statistically significant effects on students' cognitive understandings.

An analysis of random effects estimated the variance accounted for by the within- and between-group factors. R is the variance of the first-level residual r_{ij}, and U_0 is the variance of the second-level residual e_0. When the second-level variables were controlled for, significant between-group differences were observed in the students' cognitive understanding scores ($p < 0.05$).

Table 8
lts.

	Fixed Effect			
	Symbol	Coefficient	Standard Error	*P*
Orientation	W_1^G	−0.56*	0.24	0.028
Proposition generation	W_2^G	−0.32*	0.14	0.027
Experimentation	W_3^G	0.08	0.19	0.696
Interpretation and conclusion	W_4^G	0.29*	0.13	0.029
Regulation	W_5^G	−0.05	0.13	0.700
Sustaining understanding	W_6^G	0.53*	0.23	0.028
	Random Effect			
	Variance	*Df*	Chi-Square	*P*
Intercept, U_0	0.20	1	4.63	0.031*
Level-1 effect, *R*	0.67			

* < 0.05, ** < 0.01, *** < - 0.001.

The deviance and the number of estimated parameters for the full model were 291.7 and 9, respectively. A likelihood-ratio test was conducted to compare the null model and the full model, and the results revealed that the full model fit the data significantly better than did the null model ($\chi^2_{(df=6)} = 17.27$, $p < 0.01$). Overall, the full model explained 67.7% of the variance in the students' cognitive understanding at the student level, 19.8% of the variance at the group level, and 49.1% of the total variance. In social science research, these effect sizes would generally be considered medium to large (Fan & Konold, 2010).

5. Discussion

This study was conducted in an authentic setting and on a large scale, which supports its ecological validity. The intellectual merit of this study derives from its contribution to our understanding of the role that both transformative and non-transformative learning processes play in collaborative-inquiry learning and of how important they are to individual performance. Specifically, we discovered the extent to which transformative and non-transformative group-learning processes contributed to individuals' understandings of Ohm's Law. For example, although the importance of orientation, as well as of interpretation and conclusion, is widely recognized in the literature (Gijlers & de Jong, 2009), we found that when a group of students did not go beyond orientation and drew too many conclusions without sufficient experimental evidence or without seriously examining the relationship between the variables by engaging in proposition/hypothesis generation, they were less likely to improve their understanding of Ohm's Law. Moreover, groups were more likely to understand Ohm's Law when they made more frequent contributions to sustaining mutual understanding. This finding reveals the importance of the social aspect (which contributes significantly to group cohesion), coordination efforts, and interactional group dynamics in general (Kreft, Kreft, & de Leeuw, 1998).

The results reveal the importance of gaining topic-related cognitive knowledge, forming it into a statement and conveying it to one's group members, managing time, engaging in dialogue and problem-solving procedures, and maintaining shared understanding among group members. In contrast, when a group focused too much on discussing their goal and reporting their resistor and voltage statuses but did not really discuss the relationship between resistance and voltage or interpreted and concluded without sufficient evidence or explanation, they were unlikely to achieve a complete understanding of Ohm's Law. Previous studies have also examined differences between high-performing and low-performing groups. For instance, Malmberg et al. (2015) found that high-performing groups focused mainly on regulating the cognitive, motivational, and social aspects of their collaboration, while low-performing groups focused more on external challenges, such as the environment and time management. Sinha et al. (2015) showed that social, cognitive and metacognitive, and emotional indicators could increase students' engagement in collaborative learning and thereby improve their learning performance. In their study, the low-engagement group developed vague and incomplete plans and did not elaborate, back up with evidence, or further discuss their ideas. Moreover, their task monitoring focused mainly on the spelling of components instead of on content, and a focus on individual thinking and individual activity was revealed by the words and phrases they used, including “I think,” “I am going to,” and “my turn.” In contrast, the high-engagement group used words that referred to the collective (e.g. “we”).

This paper also has methodological implications. In this study, we built accurate machine-learning models that can identify transformative and non-transformative discussions in small-group collaborative learning environments. While previous CSCL studies (e.g., Kwon et al., 2014; Lee et al., 2015) have generally used qualitative analysis to examine a few collaborative processes, this study examined a relatively large number of groups and their collaborations. Large-scale endeavours like this one can provide more generalizable results than can qualitative case studies. Even though this study found that the decision tree model performs best using regex features, it does not necessarily such combination will perform well in all the contexts. When applying text classification models in other studies, the algorithms and features proposed in this algorithm should also be explored fully to optimise the classification performance.

In addition, this study was not purely quantitative and reliant upon data mining; we began by using human qualitative coding to identify transformative and non-transformative discussions and then used automatic machine-learning algorithms to capture the human insights for automatic coding. By using a diverse set of features for various algorithms, we optimised the automatic-coding process. The extracted features can be easily used to construct prediction models in other group-learning contexts. While the LIWC features can be readily applied in other scenarios, the rule-based features and LDA-topic features may need to be adapted to new contexts. These two steps usually require little effort, however. In addition, instead of using a simple linear regression to examine the nested-data structure, this study used multilevel modelling to determine the effects of group-level transformative and non-transformative discussions on students’ learning outcomes. This type of modelling can estimate effects in CSCL research more accurately than can traditional statistical methods. On the whole, this study offers a practical example of applying large-scale CSCL analysis while maintaining human insights.

One of the major limitations of this study has to do with assessing students' prior knowledge. Since these students did not have much background in physics and electronics, they were assumed to be starting their investigation of Ohm's Law at the same point. Still, the students may have differed in their general cognitive abilities and understanding of basic physics concepts. Similarly, the colleges students attend, and the roles of the instructors and other environmental factors may also influence students' collaboration and development of cognitive understanding. These variables were not considered in this study. In addition, the only supports the students got were introductions of how to use the Teaching Teamwork simulation and peer collaboration among group members though textual chatting. The purpose of this design was to maximize students' dependence on team members to solve the tasks and to investigate how students' transformative and non-transformative knowledge and manipulation were shaped and spread among group member. It was not because this design represented the best classroom practice of learning Ohm's Law. Ideally, to support students' learning and collaboration, teachers may provide timely instructions and scaffolding embedded may be triggered under certain conditions of students' learning. Finally, all of the findings were produced by studying collaborative investigation into a single physics concept. For this reason, they should be applied with caution in other subjects and studies.

There are several directions for future research. First, in future studies, students' prior knowledge of the learning content should be pretested, controlled, and taken into consideration during grouping to reduce its effects on the collaborative-learning process and subsequent analyses. Second, future studies employing similar collaborative-inquiry learning environments and content areas could test the validity and generalizability of our findings. Third, the current findings may help students and teachers understand the importance of actions such as interpretation and sustained mutual understanding to successful problem solving. Therefore, teachers can try to enhance this kind of contributions, and students can better understand what is desired from them and may be more capable to monitor their progress. The findings may also guide design of automatic scaffolding strategies to monitor collaboration behaviors that are not conducive to productive group work—e.g. remaining silent for too long, spending too much time to understand tasks, or failing to sufficiently elaborate on one's or one's partner's reasoning.

Conflicts of interest

The authors declare that they have no conflict of interests.

Acknowledgements

Thanks to the Concord Consortium for supporting this work. We would also like to thank the children and teachers involving in this study for making it possible.

References

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*, 1–18.
Andriessen, J., Baker, M., & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. *Arguing to learn* (pp. 1–25). Dordrecht: Springer.
Bell, T., Urhahne, D., Schanze, S., & Ploetzner, R. (2010). Collaborative inquiry learning:

Models, tools, and challenges. *International Journal of Science Education, 32*(3), 349–377.
Dalal, M. K., & Zaveri, M. A. (2011). Automatic text classification: A technical review. *International Journal of Computer Applications, 28*(2), 37–40.
De Jong, T. (2006). Technological advances in inquiry learning. *Science, 312*(5773), 532–533.
De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179–201.
De Lisi, R., & Golbeck, S. L. (1999). *Implications of Piagetian theory for peer learning*.
Du, J., Xu, J., & Fan, X. (2013). Factors affecting online groupwork interest: A multilevel analysis. *Journal of Educational Computing Research, 49*(4), 481–499.
Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior, 21*(3), 463–486.
Fan, X., & Konold, T. R. (2010). Statistical significance versus effect size. *International encyclopedia of education: vol. 7*, (pp. 444–450). Oxford: Elsevier.
Fleiss, J. L. (1981). *Statistical methods for ratios and pro- portions* (2nd ed.). New York: John Wiley & Sons, Inc.
Gerard, L. F., Tate, E., Chiu, J., Corliss, S., & Linn, M. C. (2009). Collaboration and knowledge integration. *Paper presented at the International conference on Computer supported collaborative learning*.
Gijlers, H., & de Jong, T. (2009). Sharing and confronting propositions in collaborative inquiry learning. *Cognition and Instruction, 27*(3), 239–268.
Graesser, A. C., Fiore, S. M., Greiff, S., Andrews-Todd, J., Foltz, P. W., & Hesse, F. W. (2018). Advancing the science of collaborative problem solving. *Psychological Science in the Public Interest, 19*(2), 59–92.
Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to kirschner, sweller, and. *Educational Psychologist, 42*(2), 99–107.
Horwitz, P., Gobert, J. D., Buckley, B. C., & O'Dwyer, L. M. (2010). Learning genetics from dragons: From computer-based manipulatives to hypermodels. In M. Jacobson, & P. Reimann (Eds.). *Designs for learning environments of the future: International perspectives from the learning sciences* (pp. 61–89). Berlin: Springer.
Janssen, J., & Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist, 48*(1), 40–55.
Jeong, H., Hmelo-Silver, C. E., Jo, K., & Shin, M. (2016). January). CSCL in STEM education: Preliminary findings from a meta-analysis. *System sciences (HICSS), 2016 49th Hawaii international conference on* (pp. 11–20). IEEE.
van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., & Manlove, S. (2005). Co-lab: Research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior, 21*(4), 671–688.
Kapur, M., Voiklis, J., & Kinzer, C. K. (2008). Sensitivities to early exchange in synchronous computer-supported collaborative learning (CSCL) groups. *Computers & Education, 51*(1), 54–66.
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.
Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.
Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, J. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning,* 1–21.
Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering, 160*, 3–24.
Kreft, I. G., Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling.* Sage.
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.
Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research, 86*(3), 681–718.
Lazonder, A. W., & Rouet, J.-F. (2008). Information problem solving instruction: Some cognitive and metacognitive issues. *Computers in Human Behavior, 24*(3), 753–765.
Lee, A., O'Donnell, A. M., & Rogat, T. K. (2015). Exploration of the cognitive regulatory sub-processes employed by groups characterized by socially shared and other-regulation in a CSCL context. *Computers in Human Behavior, 52*, 617–627.
Linn, M. C., Lee, H. S., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and assessing knowledge integration in science. *Science, 313*(5790), 1049–1050.
Malmberg, J., Järvelä, S., Järvenoja, H., & Panadero, E. (2015). Promoting socially shared regulation of learning in CSCL: Progress of socially shared regulation among high-and low-performing groups. *Computers in Human Behavior, 52*, 562–572.
Mäkitalo-Siegl, K., Kohnle, C., & Fischer, F. (2011). Computer-supported collaborative inquiry learning and classroom scripts: Effects on help-seeking processes and learning outcomes. *Learning and Instruction, 21*(2), 257–266.
McKeown, J., Hmelo-Silver, C. E., Jeong, H., Hartley, K., Faulkner, R., & Emmanuel, N. (2017). *A meta-synthesis of CSCL literature in STEM education.* Philadelphia, PA: International Society of the Learning Sciences.
Mu, J., Stegmann, K., Mayfield, E., Rosé, C., & Fischer, F. (2012). The ACODEA framework: Developing segmentation and classification schemes for fully automatic analysis of online discussions. *International journal of computer-supported collaborative learning, 7*(2), 285–305.
Njoo, M., & De Jong, T. (1993). Exploratory learning with a computer simulation for control theory: Learning processes and instructional support. *Journal of Research in Science Teaching, 30*(8), 821–844.
Okada, T., & Simon, H. A. (1997). Collaborative discovery in a scientific domain. Cognitive Science, 21(2), 109-146.ocesses and learning outcomes. Learning and Instruction, 21(2), 257-266.
Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*.
Piaget, J. (1926). Language and though of the child. London: Paul Kegan. (original work published 1923).
Plass, J. L., Milne, C., Homer, B. D., Schwartz, R. N., Hayward, E. O., Jordan, T., et al. (2012). *Science Teaching, 49*, 394–419.
Popov, V., Xing, W., Zhu, G., Horwitz, P., & McIntyre, C. (2018). The influence of students' transformative and non-transformative contributions on their problem solving in collaborative inquiry learning. *The13th international conference of the learning sciences* (pp. 855–862). London: International Society of the Learning Sciences.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*, *Vol. 1*. Sage.
Organisation for Economic Co-operation and Development. (2017). PISA 2015 assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving (Rev. ed.), Retrieved from the OECD website: https://doi.org/10.1787/9789264281820-en.
Rosé, C., Wang, Y. C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., et al (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), 237-27.
Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *The Journal of the Learning Sciences, 14*(2), 201–241.
Saab, N., van Joolingen, W., & van Hout-Wolters, B. (2012). Support of the collaborative inquiry learning process: Influence of support on task and team regulation. *Metacognition and learning, 7*(1), 7–23.
Sinha, S., Rogat, T. K., Adams-Wiggins, K. R., & Hmelo-Silver, C. E. (2015). Collaborative group engagement in a computer-supported inquiry learning environment. *International Journal of Computer-Supported Collaborative Learning, 10*(3), 273–307.
Stahl, G. (2013). Transactive discourse in CSCL. *International Journal of Computer-Supported Collaborative Learning, 8*(2), 145–147.
Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. Cambridge handbook of the learning sciences, 2006, -.
Stegmann, K., Kollar, I., Zottmann, J., Gijlers, H., De Jong, T., Dillenbourg, P.,., & Fischer, F. (2007). July). Towards the convergence of CSCL and inquiry learning: Scripting collaborative inquiry learning. *Proceedings of the 8th iternational conference on Computer supported collaborative learning* (pp. 831–832). International Society of the Learning Sciences.
Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning, 1*(3), 315–337.
Vygotsky, L. S. (1978). Mind and society: The development of higher mental processes. (Cambridge).
Wang, Y. C., Kraut, R., & Levine, J. M. (2012, February). To stay or leave: The relationship of emotional and informational support to commitment in online health support groups. *Proceedings of the ACM 2012 conference on computer supported cooperative work* (pp. 833–842). ACM.
Xing, W., & Gao, F. (2018). Exploring the relationship between online discourse and commitment in Twitter professional learning communities. *Computers & Education, 126*, 388–398.
Xing, W., Goggins, S., & Introne, J. (2018). Quantifying the effect of informational support on membership retention in online communities through large-scale data analytics. *Computers in Human Behavior, 86*, 227–234.
van Zyl, J. M., Neudecker, H., & Nel, D. G. (2000). On the distribution of the maximum likelihood estimator of Cronbach's alpha. *Psychometrika, 65*(3), 271–280.
Zheng, J., Xing, W., & Zhu, G. (2019). Examining sequential patterns of self-and socially shared regulation of STEM learning in a CSCL environment. *Computers & Education, 136*, 24–48.

Wanli Xing is an Assistant Professor in Instructional Technology at Texas Tech University, USA with background in learning sciences, statistics, computer science and mathematical modelling. His research interests are educational data mining, learning analytics, and CSCL.

Vitaliy Popov is a research associate in Jacobs Institute for Innovation in Education at University of San Diego. Dr. Popov areas of interest include: computer supported collaborative learning, mobile learning, learning sciences, learning analytics, and technology-enhanced learning across cultures.

Gaoxia Zhu is a PhD student and research assistant in the Institute for Knowledge Innovation & Technology, Ontario Institute for Studies in Education (OISE), University of Toronto with background in Educational Technology and Curriculum Studies. Her research interests include Knowledge Building, learning analytics, and CSCL.

Paul Horwitz is a theoretical physicist with an undergraduate degree from Harvard, a master's degree from Columbia, and a doctorate from New York University. His educational research interests center around the creation of interactive computer models to help students understand and apply scientific concepts, the analysis of performance data generated by such models, and the design of collaborative problem-solving environments for teaching and assessing the performance of teams.

Cynthia McIntyre has worked at the Concord Consortium since 1997 on a variety of projects. In her early days, she worked on one of the first online teacher professional development courses, which led to the development of online “metacourses” (online courses about online courses) for the Virtual High School. She co-authored a short book on the subject, too.