The Internet and Higher Education 41 (2019) 51–61

Contents lists available at ScienceDirect

The Internet and Higher Education

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

Uncovering the sequential patterns in transformative and non-transformative discourse during collaborative inquiry learning

Gaoxia Zhu[a], Wanli Xing[b,*], Vitaliy Popov[c]

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

ARTICLE INFO

Keywords:
Collaborative work
Transformative discourse
Sequential analysis
Learning analytics
Educational data mining

ABSTRACT

Many universities are using computer-supported collaborative-inquiry-learning (CSCiL) environments to develop their students' skills in collaboration, problem solving, and critical thinking. Diverse states of discourse during CSCiL occur in sequences, but we do not yet fully understand which patterns are beneficial to learning and when exactly to foster them. This study used transition-rate analysis, entropy-analysis, and sequential pattern mining to analyze the chat message of 144 students of two-year colleges. The participants worked on tasks related to Ohm's Law in a simulation-based collaborative-inquiry-learning environment. The results revealed that students in groups who completed tasks successfully tended to ensure that everyone in their group had a shared understanding of the relationship between the variables before they moved on to the next step. In contrast, students in groups who did not complete tasks successfully were more likely to regulate the process without reaching a shared understanding.

Gaoxia Zhu is a PhD candidate 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.

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.

1. Introduction

University educators are increasingly focused on developing students' competencies in productive collaboration and problem-solving (Binkley et al., 2009; Foster & Yaoyuneyong, 2016). A major reason for this is that many university graduates will work as members of interdisciplinary and cross-functional teams responsible for accomplishing cognitively complex tasks (Lovelace, Shapiro, & Weingart, 2001; Tancig, 2009), and functioning effectively on such a team requires a high level of professional competence (Mulder, 2014). For this reason, a crucial challenge in education is the development of instructional strategies that can effectively promote high-quality collaboration among learners (Fischer, Kollar, Stegmann, & Wecker, 2013). Computer-supported collaborative-inquiry learning (CSCiL)—in which two or more learners collaborate via computer to solve problems and/or co-construct knowledge—has promise as such a strategy.

Investigating how students engaged in CSCiL construct knowledge and solve problems and determining why some students outperform others could offer insights into how to better support CSCiL. Soller, Martínez, Jermann, and Muehlenbrock (2005) proposed a framework for describing the cycle process of collaboration management. The process consists of five phases: collect interaction data, construct a modal of interaction, compare the current state of interaction to the desired state, advise/guide the interaction, and evaluate interaction assessment and diagnosis (Soller et al., 2005). Based on results generated during the phases, mirroring tools, metacognitive tools and guiding tools can be developed (Soller et al., 2005). Mirroring tools aggregate data about students' interactions and raise their awareness

* Corresponding author.
E-mail addresses: gaoxia.zhu@mail.utoronto.ca (G. Zhu), wanli.xing@ttu.edu (W. Xing), vpopov@sandiego.edu (V. Popov).

https://doi.org/10.1016/j.iheduc.2019.02.001
Received 28 September 2018; Received in revised form 26 January 2019; Accepted 14 February 2019
Available online 17 February 2019
1096-7516/ Published by Elsevier Inc.

about their actions. Metacognitive tools display the current state of interactions and what the desired state may look like. Guiding tools provide pieces of advice and suggestions to users based on the difference between the desired state and the current state. Thus, this study can inform instructional design choices by suggesting what specific types of student activities should be promoted in order to improve the quality of collaboration, and thereby learning outcomes.

A significant number of studies have investigated the use of CSCiL in science education (see Donnelly, Linn, & Ludvigsen, 2014; Jeong, Hmelo-Silver, Jo, & Shin, 2016 or Lazonder & Harmsen, 2016 for reviews). When engaged in CSCiL in science environments, learners may search for information, form hypotheses, experiment, and interpret, articulate, and share ideas (Mäkitalo-Siegl, Kohnle, & Fischer, 2011; Stahl, Koschmann, & Suthers, 2006). However, research into collaborative-inquiry learning has consistently shown 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—for instance, by forming hypotheses and interpreting data (Bell, Urhahne, Schanze, & Ploetzner, 2010; Gijlers & de Jong, 2009; Lazonder & Rouet, 2008; Parkes, Stein, & Reading, 2015; Waugh & Su-Searle, 2014; Popov, Leeuwen, & Buis, 2017). Studies into collaborative problem solving and knowledge co-construction have mainly used different coding schemes to analyze student discourse and evaluated the frequencies with which students make different kinds of contributions (De Wever, Schellens, Valcke, & Van Keer, 2006).

CSCiL studies typically examine collaborative activities that last for weeks or months, which makes it important to consider variables such as time, discourse sequence, and mode of communication (synchronous versus asynchronous) when analyzing individual or group learning processes (Reimann, 2009). Several recent studies have examined the temporal characteristics of CSCiL, and some progress has been made in identifying relationships between specific sequences of learning processes and learning outcomes (e.g., Andriessen, Baker, & Suthers, 2003; Molenaar & Chiu, 2017; Schoor & Bannert, 2012; Tang, Xing, & Pei, 2018a, 2018b; Wise & Chiu, 2011; Popov, Biemans, Brinkman, Kuznetsov, & Mulder, 2013, Popov, Leeuwen, & Buis, 2017). There are multiple challenges to be addressed, including the absence of frameworks and theories that can be used to conceptualize the dimension of “time,” the complexity of translating theoretical conceptions of temporality into appropriate methods of analysis, and the difficulty of providing timely and impactful temporal analyses for use in physical or online classrooms given the labor and time required to collect the data and conduct the analyses (Knight, Wise, & Chen, 2017).

The goal of this study is to expand our understanding of the influence of specific sequences of learning processes on the performance of groups engaged in CSCiL in science contexts. We build upon our previous study (Popov, Xing, Zhu, Horwitz, & McIntyre, 2018), which investigated how transformative and non-transformative utterances made during CSCiL influence the problem-solving performance of small groups. The previous study and this current study categorizes student utterances as belonging to one of two categories of learning processes: “transformative” learning processes relate directly to the construction of knowledge, which include *orientation (O), proposition generation (P), experimentation (E), and interpretation & conclusion (I)*, and “non-transformative” learning processes relate to technical features of learning platforms and group management, which include *sustaining mutual understanding (S) and regulation (R)*. Recent advances in CSCL literature indicate that the efficacy of collaborative learning effort is influenced by the extent to which students can coordinate their collaborative activities (Erkens, Jaspers, Prangsma, & Kanselaar, 2005) and transact on each other's ideas (Kirschner, Beers, Boshuizen, & Gijselaers, 2008; Stahl, 2013). In other words, learners are required to coordinate both at a macro level (in terms of managing time, group relationship and task activities – “non-transformative”) and at a micro level (verbal exchanges of knowledge and beliefs – “transformative”) during discourse (Popov, Leeuwen, & Buis, 2017). Coordination at these two levels are essential to 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 micro level gives rise to cognitive activities that stimulate knowledge construction (Kirschner, Sweller, Kirschner, & Zambrano, 2018; Stahl, 2013).

Our previous study found that when a group of students successfully solved a task, their discourse contained a higher proportion of utterances related to *proposition generation*, *regulation*, and *the sustaining of mutual understanding* (Popov, Xing, Zhu, Horwitz, & McIntyre, 2018). In contrast, when a group of students failed to solve a task, their discourse contained a significantly higher proportion of utterances related to *orientation*, and *interpretation and conclusion*. Although that study demonstrated the importance of *proposition generation*, *regulation*, and *the sustaining of mutual understanding*, it remains unclear whether the sequence in which such processes occur influences students' success in collaborative inquiry learning environments. To determine this, this study performed a series of sequential analyses on students' transformative and non-transformative online utterances: a transition-rate analysis, an entropy analysis, and an instance of sequential pattern mining. We aim to investigate how the chat discourse differs in the successful and less successful conditions in terms of transition rate, entropy and sequences. The results imply that sequences such as “*interpretation and conclusion*” to “*experimentation*,” and “*proposition generation*” to “*sustaining mutual understanding”* should be enhanced via metacognitive or guiding tools or through students' reflection on their current state of the desired state of collaboration with or without teachers' support.

2. Literature review

2.1. Collaborative inquiry learning

CSCiL, the mix of inquiry learning and computer-supported collaborative learning (Bell et al., 2010), is a promising approach to preparing pre-college and college students for the workplace and life. In the information age, knowledge is growing and specializing at an ever-accelerating rate. Information and communication technology is changing social relationships, the types of work people are doing, and how work is done. The success of a professional increasingly depends on their abilities to locate and share useful information, to adapt to changing demands and circumstances, to collaborate with others to improve tangible and intangible artefacts and to create new knowledge (Binkley et al., 2009). In response to the growing demand for these abilities, many universities have introduced collaborative technologies that can better prepare their students for the new world of work (King & South, 2016).

CSCiL environments enable students to inquire the natural world like real scientists, by generating questions, conducting experiments, observing phenomena, proposing explanations, and so forth. CSCiL environments also allow students to engage with their peers in common endeavours to share information and to construct knowledge (Dillenbourg, 1999). In reviewing studies by 10 groups of investigators, Bell et al. (2010) identified nine “main inquiry process”: communicating, orienting and asking questions, making predictions, generating hypotheses, planning, investigating, analyzing and interpreting, modelling, and making evaluations and drawing conclusions. Similarly, De Jong (2006) identified seven cognitive processes constitutive of inquiry learning: orientating, generating hypotheses, experimenting, planning, monitoring, evaluating, and reaching conclusions. Some of these are processes directly contribute to the construction of knowledge—e.g. hypothesis generation and experimentation—while others are regulative activities—e.g. planning and monitoring.

Using new computer-based environments to promote collaborative learning in STEM education has several potential benefits. First, these environments can prepare students for participation in a networked, virtualized society (Belz, 2003). Second, these environments can help students to develop social and cognitive abilities (Fischer, Hmelo-Silver, Goldman, & Reimann, 2018; Lim & Liu, 2006). Third, these environments can give students experience in engaging in a variety of learning communities, regardless of physical or temporal distance (Scardamalia & Bereiter, 2014; Walther, 1997). Finally, these environments can facilitate the provision of timely and precise scaffolding that can help students to develop higher-order skills and to generate solutions to complex problems. After reviewing 144 experimental studies on the effects of computer-based scaffolding in STEM education, Belland, Walker, Kim, and Lefler (2017) found that computer-based scaffolding exhibited a consistently positive effect on students' cognitive achievement.

When students use computer-supported collaborative-learning (CSCL) environments to write messages and notes, these messages and notes make explicit their interactions and knowledge-construction processes (Macdonald, 2003). Indeed, discussion transcripts are a "gold mine of information concerning the psycho-social dynamics at work among students, the learning strategies adopted, and the acquisition of knowledge and skills" (Henri, 1992, p. 118). Discussion transcripts have been analyzed in several ways. Initially, students' participation was measured using various methods of quantitative frequency analysis, including counting the number of words or notes contributed by an individual student (Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2012). Because quantitative frequency analysis does not evaluate the quality of student contributions, then content analyses have recently been employed in investigating knowledge construction. These analyses have examined a number of different aspects of learning—including students' learning gains, skill acquisition, and attitudes towards different learning environments (e.g., Chen, Wang, Kirschner, & Tsai, 2018). For this purpose, different coding schemes have been employed (see De Wever et al., 2006's review for details). Given the recognized importance of interaction in collaborative learning, social network analysis and epistemic network analysis have also been employed (e.g., Shaffer et al., 2009; Zhang, Scardamalia, Reeve, & Messina, 2009). These analyses have generally overlooked the temporal information embedded in discussion transcripts, and for this reason have failed to give due attention to the importance of the order in which different types of learning processes occur (e.g. transformative and non-transformative learning processes).

2.2. Temporal analysis

CSCL researchers have developed methods to study the role of time issue in student knowledge construction process. For instance, to investigate the relationships among emotion, content, and interaction, Lund, Quignard, and Shaffer (2017) used epistemic network analysis to create temporal representations of students' conversations. To help identify meaningful connections between each discourse turn, Siebert-Evenstone et al. (2017) used epistemic network analysis to segment discourse data. Andrade, Danish, and Maltese (2017) used a hidden Markov model to investigate the influence of sequences of physical movements (especially hand movements) on learning. In addition, Molenaar and Chiu (2017) used Statistical Discourse Analysis (SDA) to study the effects of cognition sequences on group performance and the effects of social metacognitive actions on the occurrence of correct, new ideas and justifications (Chiu & Lehmann-Willenbrock, 2016). Wise and Chiu (2011) also used SDA to explore temporal patterns in knowledge construction. Chen, Resendes, Chai, and Hong (2017) used Lag-sequential Analysis and frequent sequence mining to identify sequential patterns of student online knowledge building discourse that distinguished productive and unproductive threads. Yang and his colleagues (Yang, Li, Guo, & Li, 2015; Yang, Li, & Xing, 2018) used lag sequential analysis and frequency analysis to examine patterns in how students constructed knowledge during online cooperative translation activities. However, no study to date has examined sequential patterns of transformative and non-transformative discourse processes in collaborative-inquiry problem solving.

We reviewed almost a dozen of studies which investigated elements of temporality in CSCL context. Most of these studies were conducted in asynchronous collaborative-learning environments, and they were split roughly 50/50 between K–12 settings and higher-education settings. For example, Epp, Phirangee, and Hewitt (2017) studied 90 graduate students taking online courses via Pepper, a discourse-based online learning environment. Lee and Tan (2017) analyzed 281 notes contributed by 13 in-service teachers via Knowledge Forum, an online environment designed to build knowledge. Siebert-Evenstone et al. (2017) developed approaches for segmenting the discourse contributed by 44 first-year engineering students via a portal for online work that featured email and a chat window. Author (2017) analyzed the activities of 74 first-year university students working on an ill-defined problem related to biodiversity collapse via a platform called the Virtual Collaborative Research Institute. Wise and Chiu (2011) analyzed the online discourse of 21 university students who were taking blended course via open-source LMS Moodle. Hmelo-Silver, Chernobilsky, and Jordan (2008) used frequency analysis to examine the activities of two groups of third-year university students participating in a blended course that utilized the STELLAR learning environment. Andrade et al. (2017) analyzed the sensorimotor coordination recorded in embodied simulation of 15 third and fourth graders. Lund et al. (2017) studied a conversation between two students during their lunch. Molenaar and Chiu (2017) analyzed 32,375 turns of talking recorded in a virtual workspace occupied by 54 primary school students. Chiu and Lehmann-Willenbrock (2016) analyzed 3296 turns by 80 9th grade students working in 20 small groups. Chen et al. (2017) analyzed 1101 online notes contributed by students between grades 1 and 6 via Knowledge Forum.

A number of studies have identified behavioural patterns. For example, Molenaar and Chiu (2017) found that high cognition (e.g., generating new ideas, elaborating them) often followed low cognition (e.g., reading, processing information from external sources), especially in groups that added more words to a lesson. Compared to low cognition, high cognition was more likely to trigger high cognition three, four, and five turns later, and planning turns tended to inspire high cognition up to four turns later. Chen et al. (2017) found that while "supporting discussion" was more likely to follow "obtaining information" or "theorizing" in unproductive threads, productive threads were more likely to include "sustained theorizing," "the integrated use of evidence and information," and frequent attempts to problematize proposed theories. Chiu and Lehmann-Willenbrock (2016) found that students who agreed or rudely disagreed with an idea tended not to express micro-creativity in the same turn, that rude disagreement tended to follow disagreement, and that a sequence of wrong ideas followed by rude disagreement likely led to micro-creativity. Wise and Chiu (2011) found that after a pivotal post, knowledge construction was higher and summaries were more likely to occur. They also found that the current post averaged lower knowledge construction if the previous post contained new ideas and that if a post was responsive, the following post was more likely to contain new ideas. Yang et al. (2018) found that higher-engagement groups were more likely than lower-engagement groups to move from sharing information and ideas to communicating emotions and negotiating meaning to negotiating meaning itself.

The above-mentioned studies have explored the effects of time on group performance—e.g. essay quality, the quality of discourse, performance in problem solving, and performance in the construction of knowledge. For instance, Molenaar and Chiu (2017) found that groups with more low cognitions and groups that exhibited sequences of low cognition followed by high cognition added more words to their essays. They also found that the following groups differed in their uses of words and sequences: groups with high cognition, groups with sequences of low cognition followed by low cognition, and groups with sequences of high cognition followed by an action followed by low cognition (Molenaar & Chiu, 2017). Chen et al. (2017) characterized productive threads as those exhibiting more transitions between "questions," "obtaining information," "working with information," and "theorizing." They also found that threads that merely presented opinions were unlikely to be productive (Chen et al., 2017). Chiu and Lehmann-Willenbrock (2016) concluded, however, that rude disagreement may yield greater micro-creativity in groups that solve problems or when they occur after wrong ideas are presented. They also found that students who perform better academically tend to inspire more justification (Chiu & Lehmann-Willenbrock, 2016). Wise and Chiu (2011) argued that in most online discussions, a pivotal signal post separates a phase lower in knowledge construction from a phase higher in knowledge construction. They also found that students who played the roles of "synthesizer" and "wrapper" tended to contribute pivotal posts.

Although progress on temporal analysis in the CSCL field has been made, challenges remain. For instance, the Journal of Learning Analytics published two special issues on temporal analysis in late 2017 and early 2018, in which they outlined the latest studies in the field and the conceptual, methodological, and pragmatic challenges to temporally analyzing learning data. Frameworks and theories that can be used to conceptualize time are lacking, and translating theoretical conceptions of temporality into appropriate analytic methods is methodologically complex. Moreover, providing timely and impactful temporal analyses for use in physical or online classrooms is difficult given the labor and time required to collect the data and conduct the analyses (Knight et al., 2017). Similarly, Reimann (2009) argues that the dominant variable-centred variance theory limits the study of longer time scales in CSCL and that the event-centred view should instead be used to more tightly link quantitative and qualitative methods. Molenaar (2014) identifies six obstacles to temporal analysis: 1) the paradigm has shifted from variable-based approaches to event-based approaches; 2) a framework is needed for conceptualizing time and temporal characteristics; 3) methodological approaches are needed for appropriately answering time-related questions; 4) guidelines are needed for segmenting time units; 5) temporal analysis occurs at the micro level, but most theories are defined at the macro level; and 6) most research into time has been comparative or exploratory. To meet these challenges, the current study analyzes interactions among students in a synchronous computer-supported collaborative simulation inquiry environment to evaluate the impacts of sequences of transformative and non-transformative utterances on students' success in completing a task.

3. Methodology

3.1. Research context

The participants in this study were 144 students from five two-year colleges. All the 2-year colleges were located in the United States. The participants of the same class were divided into groups of three by the teachers. The students were asked to work with their group members to use their knowledge of Ohm's Law to solve four tasks of increasing complexity related to the relationship between resistance and voltage in series circuits. The interface of the Teaching Teamwork platform, a collaborative inquiry learning environment featuring a series of interactive STEM activities, is shown in Fig. 1. Each group member controlled one resistor—R1, R2 or R3. As shown in the upper and bottom of Fig. 1, in Level A task, both the E and R0 values are given, and R0 equals R1, R2 and R3. Therefore, the goal voltage across R1, R2, and R3 equals. As shown in Fig. 2, in Level B, both E and R values are given, R0 does not equal zero, and the goal voltage values across R1, R2 and R3 are different. In Level C, E is unknown, R0 is given and does not equal zero, the goal voltage values across R1, R2 and R3 are different. Finally, in Level D, both E and R values are unknown, and the goal voltage across R1, R2 and R3 are different.

The participants had to depend on one another to figure out the value of the resistor they controlled in series circuits, and they were only able to communicate through a chat box embedded in the Teaching Teamwork platform. The students had about 60 min to work on the teamwork tasks. A textual introduction and a video were also embedded in the platform to help the students understand their goals and learn how to use the simulation. For details regarding the participants, the settings, and the procedures, please see Authors (2018).

In our previous study (Popov, Xing, Zhu, Horwitz, & McIntyre, 2018), we characterized successful conditions as those in which the task was completed and less-successful conditions as those in which the task was not completed. We manually coded 1111 pieces of the chat messages generated in 15 solved and unsolved tasks as either "transformative" or "non-transformative," and these messages were randomly selected from the dataset generated by all of the participants with task as the unit. Table 1 displays the descriptions of the categories and related examples of the coding scheme. Then, a set of features was extracted from the manually coded messages and inputted into machine-learning algorithms. Supervised machine models were built, and the one with the best performance was used to automatically identify six different types of transformative and non-transformative discourses. As a result, every chat message in both the "successful" and "less successful" conditions was categorized as either "transformative" or "non-transformative" discussion. The labelled messages were the starting point of the present study.

3.2. Data analysis

In order to identify sequential patterns of transformative and non-transformative discussions in the "successful" and "less successful" conditions, a series of sequential analyses were conducted that included transition-rate analysis, entropy analysis, and sequential pattern mining. The focus of each of these analyses was different, but they complemented each other, providing a comprehensive picture of the sequential characteristics of the discourses. Below, we explain the different methods in detail.

3.2.1. Transition-rate analysis

The transition-rate analysis yielded information about the changes in discourse states that occurred most frequently among all of the group discourses. It has been widely applied in social- and learning-sciences research (e.g., Yang et al., 2015; Yang et al., 2018). In this study, we employed the basic procedure to compute the rates of transition between the different discourses (or "states"). For example, given a couple of states ("*orientation*" and "*regulation*") represented as (d_i, d_j), the transition-rate analysis calculates the probability of a change at a given position from "*orientation*" to "*regulation*." $n_t(d_i)$is defined as the number of sequences that end with "*orientation*" (d_i) at position t, and $n_t^{t+1}(d_i, dj)$ is the number of sequences with "*orientation*" (d_i) at position t and "*regulation*" (d_j) at position $t+1$. $M$is the maximum sequence length of the different group observations. Then, the transition rate $p(d_j|\, d_i)$ between "*orientation*" (d_i) and "*regulation*" (d_j) is:

$$p(d_j \mid d_i) = \frac{\sum_{t=1}^{M-1} n_t^{t+1}(d_i, d_j)}{\sum_{t=1}^{M-1} n_t(d_i)}$$

The outcome of the transition-rate analysis is a matrix in which each row describes a transition distribution from the originating discourse state at time t to the next discourse state at $t+1$. The sum of each row is one.

3.2.2. Entropy analysis

While the transition-rate analysis focused on each individual discourse state, the entropy analysis focused on the overall distribution of different discourse states of a group. This study used Shannon's entropy specifically, which is also known as the "entropy index" (Shannon &

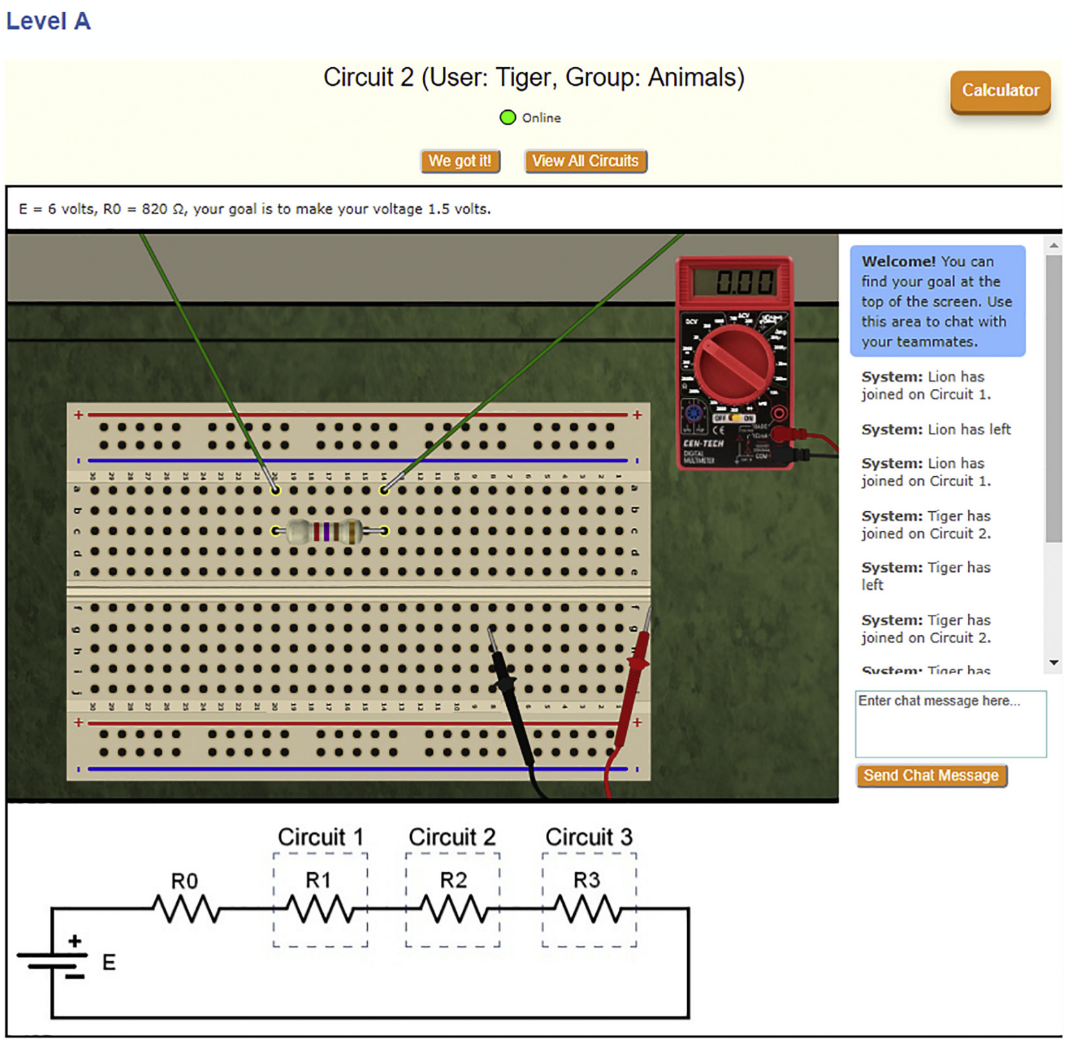

Fig. 1. Interface of the simulation (Level A task is shown as an example).

Weaver, 1998). It can be used to calculate the transversal discourse state distribution, and it was used as a metric in evaluating the diversity of the states in the chat messages of each group. Shannon's entropy has been used in many social-science studies (e.g., Bruno, 2010; Matei, Oh, & Bruno, 2006; Matei, Bruno, & Morris, 2015; Authors, 2017). Let p_i represent the proportion of the cases in the chat messages of a group in state *i*. Then, Shannon's entropy is calculated as:

$$e(p_1, p_2, \ldots, p_6) = -\sum_{i=1}^{6} p_i \log(p_i)$$

where there are six different discourse states. When all of the cases are in the same state, the entropy is zero. The entropy reaches its maximum value (one) when each of the six states has the same chance to occur in the chat messages of a group in a task. A *t*-test was used to compare the differences between the entropies of the discourses in the "successful" and "less successful" conditions.

3.2.3. Sequential pattern mining

Transition rate analysis and entropy analysis consider the discourse state and the discourse position separately. In contrast, sequential pattern mining considers the discourse and the position simultaneously. The main purpose of sequential pattern mining is to identify subsequences in a set of sequences that occur above a certain frequency level. In our case, each sequence was a group discourse consisting of different transformative and non-transformative discourse states. The sequences in Table 2 can be used as examples in illustrating the process of sequential pattern mining. The results of the sequential pattern mining are a series of subsequences consisting of the six different states. For instance, the sequence (E)–(O)–(S) is a subsequence of SID 2 and SID 3, and its support, i.e. its probability of occurrence, is 0.5 (1 in 4). The sequence (P)–(P) is only a subsequence of SID 4, and its support is 0.25. If we set 0.4 as the threshold, (E)–(O)–(S) is picked up as a subsequence of the four sequences in Table 1, but (P)–(P) is not. Sequential pattern mining can generate holistic insights, and it helped us to identify differences in the discourses of the "successful" and "less successful" conditions.

To contextualize some of the most popular sequences of chat discourses in the successful and less successful conditions, two researchers looked into the chat discourses of the two conditions. We selected excerpts that best represent the sequences in the related conditions from the manually coded results.

4. Results

4.1. Descriptive statistics

Before we conducted the advanced analysis, we calculated the descriptive statistics for the discourses in the "successful" and "less successful" conditions. The mean of the sequence lengths of the discourses in the "successful" condition was 39.74 (SD = 29.75, median = 31.5, range = 118). In contrast, the mean of the sequence lengths of the

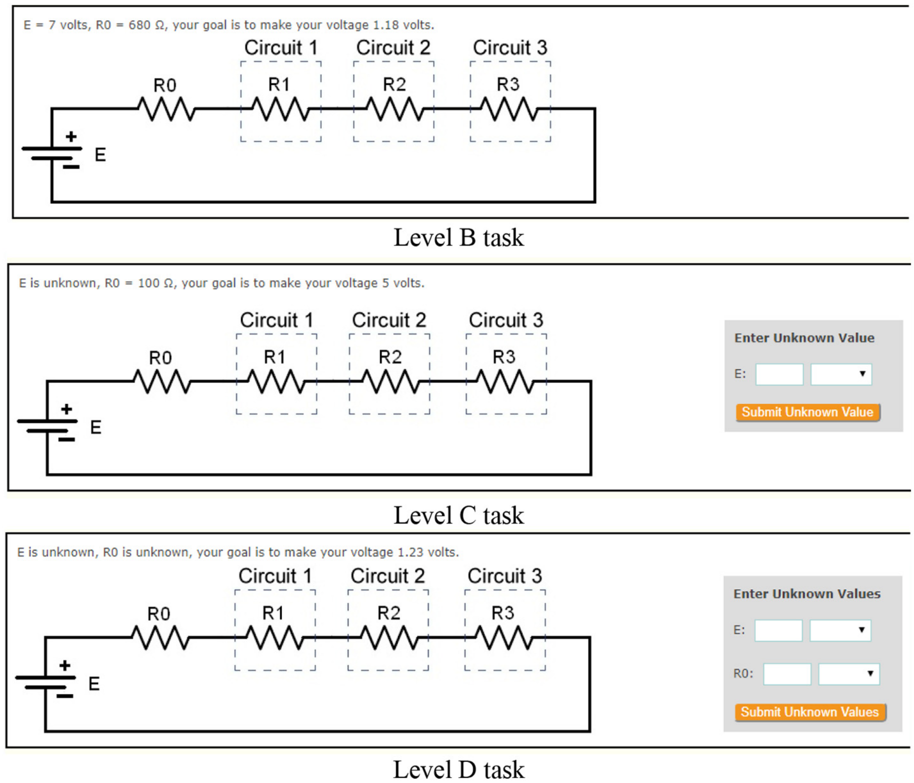

Fig. 2. The goals and submission boxes of Level B, C and D tasks.

Table 1
The coding scheme of the chat history during the collaborative inquiry learning process.

Contribution type	Category	Description	Example
Transformative (utterances that directly result in knowledge construction)	Orientation	Form an idea of the structure and the complexity of the task at hand by collecting goal-related information, the current values of the resistance or voltage and how these measures meet the individual or group goals.	*r* = 560 here; OK I need 6.69; I got my v; How many volts do you guys need? It's a series circuit.
	Proposition generation	Form a statement or a set of statements concerning the relations regarding the values of the resistance or voltage in order to solve the task.	You are going to need a higher resistance value then; So we can find the current at 6.71 mA going through the resistor, then we can find our r values using ohms law
	Experimentation	Include testing ideas and adjusting the resistors by making them lower or higher.	everyone set to 180 - see what that does; Let me get 1v hold up; Let me readjust; OK, let me go a little higher then; R1 at 680 Ω just lowered to 3.8 V; I will switch R1 to 680; Can you just change your resistor to 4.7 k ohms, I jsut want to see.
	Interpretation and conclusion	Review the proposition in light of the experimentation outcomes.	I still need to adjust; Well our total v of c1 -c3 is12.24 right; So there is a 3.76 drop across *r* = 560 totals to 16
Non-transformative (utterances relate to technical features, and time/group/task management, etc.)	Regulation	Manage time, group dialogue, the big procedures of solving tasks and so forth.	Let us get move on snow; One *sec*; Where is the other person; Let's try calculating; get to the goal; Don't we need the third person.
	Sustaining mutual understanding	Indicate shared/different understanding among different group members.	Ok; same here; yea; got it; now it does; that's good
	N/A	Information that cannot be coded with the six categories above.	Who this; who cares; and I am not trying to do this math, to be honest.

discourses in the “less successful” condition was 32.49 (SD = 26.76, median = 27, range = 99). These results show that the “successful” condition generally featured longer discourses than did the “less successful” condition. Given that the means of the discourse lengths in both conditions were < 40, we visualized the overall distributions of the discourse states in the first 40 positions. In Fig. 3, the x-axis represents the position and the y-axis represents the portion of each state among all the states. In both conditions, “*orientation*” dominated the

Table 2
An example sequence dataset.

SID	Sequence
1	(O)-(O)-(I)-(*E*)-(R)
2	(E)-(E)-(O)-(S)
3	(E)-(O)-(*S*)-(R)
4	(P)-(P)-(E)-(I)

conversation. There was a fair amount of "*experimentation*," "*sustaining mutual understanding*," and "*regulation*" in both conditions, but "*proposition generation*" and "*interpretation and conclusion*" were relatively rare—especially "*proposition generation*," which represented students' cognitive understanding of the relationship between resistance and voltage. The proportion of each state among the wholes in the "successful condition" seemed to be more stable than that in the "less successful" condition, indicating that the students in the "successful" condition were likely to constantly move between different states instead of focusing on certain states at specific times. The following analysis provides additional insights into the differences in discourse sequences observed for the "successful" and "less successful" conditions.

4.2. Results of the transition rate analysis

Table 3 shows the transition matrix describing the six transformative and non-transformative states in the "successful" and "less successful" conditions. The results show that the transition probability differed. In the successful condition, the probability of a state transitioning from "*proposition generation*" to "*orientation*" was higher; the probability of a state transitioning from "*interpretation and conclusion*" to "*experimentation*" was higher; the probability of a state transitioning from "*proposition generation*" to "*sustaining mutual understanding*" was 0.16, much higher than that in the "less successful" condition, which was zero; and the probability of a state transitioning from "*interpretation and conclusion*" to "*regulation*" was higher. In the less successful condition, the probability of a state transitioning from "*proposition generation*" to "*experimentation*" was almost six times as high as that in the "successful" condition; and the probability of a state transitioning from "*proposition generation*" to "*regulation*" was much higher than it was in the "successful" condition. The results indicate that in the "successful" condition, after the relationship between the voltage and resistance was proposed by a group member, the group tended to connect the

Table 3
The transition matrix for the six transformative and non-transformative states in the "successful" and "less successful" conditions.

	O		P		E		I		S		R	
	SC	LSC	SC	LSC	SC	LSC	SC	LSC	SC	LSC	SC	LSC
O	0.69	0.70	0.01	0.01	0.12	0.14	0.02	0.02	0.11	0.09	0.06	0.04
P	0.79	0.57	0.00	0.00	0.05	0.29	0.00	0.00	0.16	0.00	0.00	0.14
E	0.63	0.60	0.00	0.01	0.17	0.19	0.02	0.01	0.12	0.13	0.05	0.06
I	0.52	0.55	0.00	0.00	0.16	0.10	0.18	0.23	0.06	0.10	0.07	0.03
S	0.60	0.58	0.01	0.00	0.13	0.15	0.02	0.04	0.16	0.15	0.08	0.08
R	0.69	0.63	0.01	0.00	0.10	0.11	0.01	0.03	0.11	0.11	0.09	0.11

SC represents the "successful" condition, and LSC represents the "less successful" condition.
O: Orientation, P: proposition generation, E: experimentation, I: interpretation & conclusion, S: sustaining mutual understanding, R: regulation.

proposition with their task and ensure that everyone had a shared understanding of the relationship, while the students in the "less successful" condition tended to regulate the time, process or the system or to experiment immediately after generating propositions without making sure that everyone had a shared understanding.

4.3. Results of the entropy analysis

Entropy values were calculated for each task in the "successful" and "less successful" conditions. As a result, there were 86 values for the "successful" group and 39 values for the "less successful" group. The entropy value for each task is visualized in Fig. 4. We can see that the entropy values for both conditions fluctuated in different groups. The peaks suggest that all six states had almost the same chance to occur, indicating that the students made diverse utterances in attempting to complete the task. In contrast, the bottoms suggest that a few states dominated the conversation. For instance, for the "less successful" condition, the entropy value was close to zero at T2 and T35, suggesting that the likelihood of a single type of discourse (usually "*orientation*") was very high.

The average entropy value for the "success" condition was 0.53 (SD = 0.13, median = 0.52, range = 0.84), which was similar to that of the "less successful" condition, which had a mean of 0.51 (SD = 0.18, median = 0.51, range = 0.81). A *t*-test was conducted in which 39 tasks were randomly selected from the "successful" condition, and the result shows that entropy value for the "successful" groups is significant

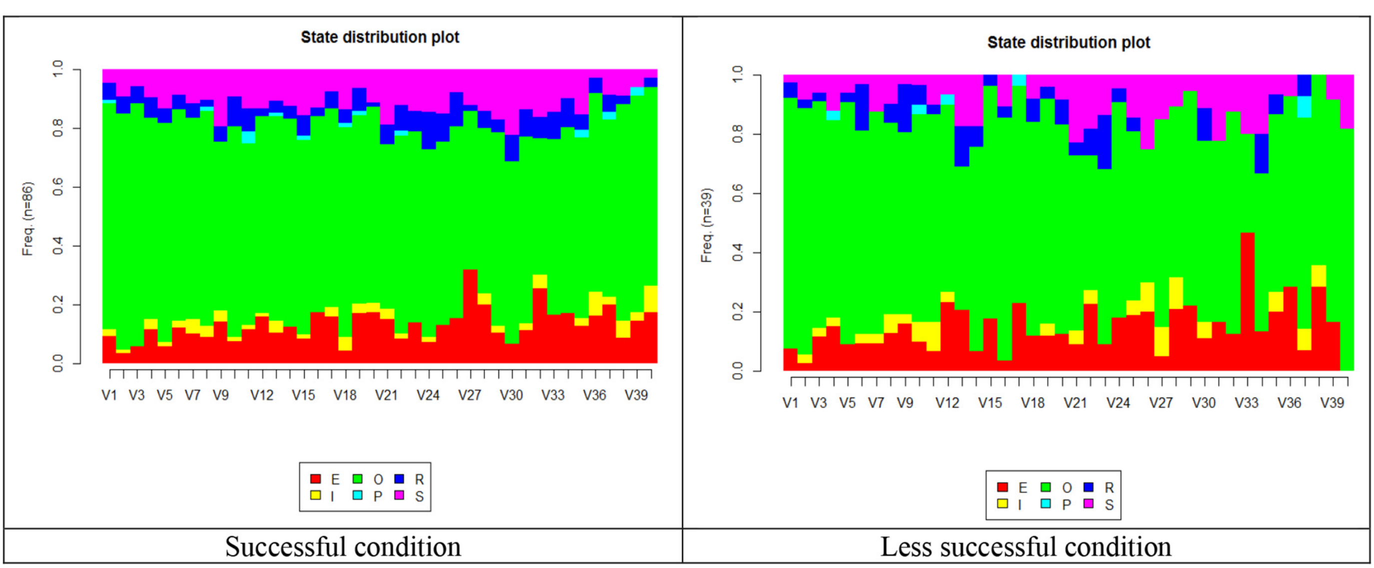

Fig. 3. Overall proportion of transformative and non-transformative states in the "successful" and "less successful" conditions.

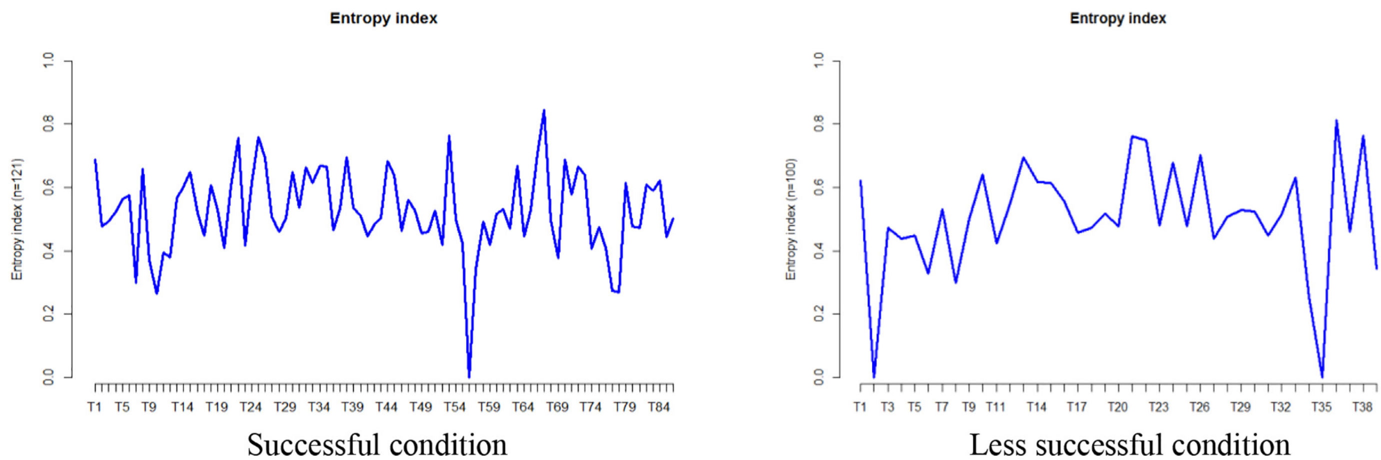

Fig. 4. The entropy value of the discourses in the "successful" and "less successful" conditions.

higher than that of the "less successful" groups (t = 2.17, df = 67, p < .05). This indicates that overall, the six types of transformative and non-transformative utterances have more equal chances to occur in the "successful" condition than in the "less successful" condition.

4.4. Results of the sequential pattern mining

The sequential pattern mining identified different subsequences with different support values for the "successful" and "less successful" conditions. 493 subsequences with a support mean of 0.46 (SD = 0.16, median = 0.41, range = 0.71) were identified in the "successful" condition, while 414 subsequences with support mean of 0.43 (SD = 0.14, median = 0.38, range = 0.69) were identified in the "less successful" condition. While the number of subsequences in the two conditions were comparable, the t-test of the support revealed a statistically significant difference between the two conditions (t = 2.56, df = 898, p < .05), indicating that the subsequences identified in the "successful" condition were more robust than the ones identified in the "less successful" condition. In both conditions, mixed transitions between "*experimentation*" and "*orientation*"—e.g. (O)–(O)–(O), (E)–(O)–(O), and (O)–(O)–(O)–(E)— dominated the subsequences, indicating that the students spent most of their time trying to understand their goals, doing experiments, and monitoring the results of their experiments. This result is consistent with those of the frequency and transition rate analyses. We removed the sub-sequences that consisted only of (O) or of (O) and only one other state, and we kept the sub-sequences that consisted of at least of two transitions other than (O). The ten most frequent sub-sequences and their related supports and counts in the "successful" and "less successful" conditions are shown in Table 4. The presence of the sub-sequence (S)–(S) in the "successful" condition suggests that the students tried to achieve shared understanding. In the "less successful" condition, the sub-sequence (E)–(E) was more prevalent, indicating that the students engaged in trial-and-error more frequently. Sequences such as (P)–(O) and (P)–(S) did not meet the threshold of support and were thus not picked up in by the sequence pattern analysis.

Table 5 shows excerpts of the chat discourses of a solved task. Sequences such as (O)-(S), (O)-(E)-(E)-(*R*)-(S), (O)-(*R*)-(S) and even (P)-(P) are displayed in this segment. This group of students first tried to understand the goal and the task and made it explicit that "*It's a series circuit*" by interacting with each other. The student B and A collaboratively formed their proposition of the relationship between voltage and resistance using Ohm's Law, indicated by "*We need the v drop to be 2 across each one*" and "*Ok so we all have to be 100.*" Then student B suggested to the group to test their proposition by adjusting the resistor each group member controlled. Both the logged data and the following chat utterance confirmed that the students did the experiment and the

Table 4
Examples of frequent sub-sequences in the "successful" and "less successful" groups.

Sub-sequence		Support		Count	
SC	LSC	SC	LSC	SC	LSC
(E)-(S)	(E)-(S)	0.69	0.64	59	25
(E)-(E)	(E)-(*S*)-(O)	0.66	0.62	57	24
(O)-(E)-(S)	(O)-(E)-(S)	0.66	0.62	57	24
(*S*)-(S)	(O)-(O)-(E)-(S)	0.65	0.62	56	24
(E)-(E)-(O)	(E)-(E)	0.64	0.59	55	23
(E)-(O)-(S)	(E)-(E)-(O)	0.64	0.59	55	23
(O)-(E)-(E)	(E)-(E)-(O)-(O)	0.64	0.59	55	23
(O)-(*S*)-(S)	(O)-(E)-(E)	0.64	0.59	55	23
(E)-(*S*)-(O)	(O)-(E)-(E)-(O)	0.63	0.59	54	23
(O)-(O)-(E)-(S)	(O)-(E)-(E)-(O)-(O)	0.63	0.59	54	23

SC represents the "successful" condition, and LSC represents the "less successful" condition.
O: Orientation, P: proposition generation, E: experimentation, I: interpretation & conclusion, S: sustaining mutual understanding, R: regulation.

proposition was correct. The excerpts indicated the students' efforts to elaborate on the relationship between resistance and voltage and to ensure group members' shared understanding.

Table 6 shows excerpts of the chat discourses of an unsolved task. Sequences such as (O)-(O)-(O) and (O)-(*E*)-(E)-(O) indicate that the students tried to orient themselves towards the learning task and understand their current values of the resistance or voltage. This group of students ran the actual experiment once. However, the lack of "*proposition generation*" suggested that no student revealed the relationship between voltage and resistance in the group, and the experimentation suggestion relied more on guesswork rather than on cognitive understanding of Ohm's Law. Although the students responded to each other, their communication was more about understanding the task itself rather than about distributing cognitive knowledge.

5. Discussion

Collaborative inquiry can promote the development of skills that students will need in the contemporary workplace, including collaboration, problem-solving, critical thinking, communication, and many others. A growing body of literature on CSCiL has paid particular attention to four elements: powerful visualizations, collaboration, student agency, and meaningful science contexts (Donnelly et al., 2014). The present study analyzed the temporal aspects of collaboration among college students in a simulation-based science-inquiry context in which they were given the agency to solve problems. Students in groups who

Table 5
Excerpt of the chat discourses of the successful conditions (typos have been corrected).

Students	Utterances	Coding
A	lol E = 8v	Orientation
B	They all need to be 120	Orientation
B	Never mind. I lied	Regulation
A	Don't we need the third person	Regulation
B	yea but you can figure it out	Regulation
B	It's a series circuit	Orientation
A	is your v drop 2v	Orientation
B	No	Sustaining mutual understanding
B	We need the v drop to be 2 across each one	Proposition generation
A	Ok so we all have to be 100	Proposition generation
B	Never mind. I have 1.98v	Orientation
B	Let me get mine to 2v first	Experimentation
B	and then the 2nd person goes	Experimentation
C	OK, just let me know when	Regulation
B	Ok I'm good	Sustaining mutual understanding
B	I used 100 Ω like you said	Orientation
C	You good, B?	Regulation
B	Yea	Sustaining mutual understanding
B	I have 2.09v	Orientation
A	Same	Orientation
C	Good	Regulation
A	now 2	Orientation
A	Done	Regulation
C	Done	Regulation

Table 6
Excerpt of the chat discourses of the less successful conditions (typos have been corrected).

Students	Utterances	Coding
D	I'm at 3.31v	Orientation
D	Can't get it to 3.5	Regulation
E	Everyone at 220	Orientation
D	220 OHMS at .01v currently	Orientation
D	I'm at 1.01 V and need to be 0.98	Orientation
F	Pen what you got	Regulation
D	Now 1.12	Orientation
E	Mike put yours at 70 Ω	Experimentation
F	Put in a higher resistor	Experimentation
D	.4v	Orientation
D	Where you guys at	Orientation
F	In the front	N/A
D	lol voltage	Orientation
F	what is everyone's voltage they need	Regulation
E	1.17	Orientation
D	I need 0.98. I'm currently at 0.37	Orientation
D	At 0.94 need to be 0.98	Orientation
D	I need to be at 1.45	Orientation
D	Currently at 2.74	Orientation
D	Currently at 1.57	Orientation
D	1.48v	Orientation

completed tasks successfully tended to ensure that everyone in their group had a shared understanding of the relationship between the variables before they moved on to the next step. The transition rate analysis of students' transformative and non-transformative chat discourse revealed that in the "successful" condition, there was a higher probability of transitions from "*proposition generation*" to "*orientation*," from "*interpretation and conclusion*" to "*experimentation*," and from "*proposition generation*" to "*sustaining mutual understanding*." Consistent with the transition rate analysis, the sequential pattern analysis shows that sequences such as (*S*)-(S) and (O)-(*S*)-(S) were more popular in the "successful" condition. The results suggest students' efforts to make connections between their understanding of the variables involved in the tasks and the goals of the tasks, to ensure that group members achieve shared understanding of the relationships between variables, to interpret experiment results and to conduct further experiments. These sequential actions associate with successful problem solving. In line with these results, Yang et al. (2018) found that higher-engagement groups had significantly more transitions from "*negotiating meaning*" to "*negotiating meaning*" and from "*sharing information/ideas*" to "*communicating emotion*."

In contrast, students in groups who did not complete tasks successfully were more likely to regulate the process without reaching a shared understanding. In the "less successful" condition, the transitions from "proposition generation" to "regulation" and from "proposition generation" to "experimentation" were much higher. Immediate experimentation and regulation followed by unelaborated proposition generation is likely to drive students away from solving group tasks or from improving their individual cognitive understandings.

The entropy analysis indicates that the students in the "successful" condition tended to use the six types of transformative and non-transformative utterances more equally. One possible explanation for this is that engaging in various types of utterances helps groups solve tasks while lacking some types of utterance might make students short of important element to succeed, for example, sustaining mutual understanding. As previous discussed, all these types, not matter directly related to knowledge construction or related to group, time or task management, are crucial, and their combination is more likely to lead to successful problem solving. Similarly, Chen et al. (2017) found that merely representing opinions tended not to results in improved explanations, while more transitions between diverse contributing types were more likely to lead to productive discourse.

While some temporal analytical methods had been applied in CSCL (e.g., Lund et al., 2017; Molenaar & Chiu, 2017; Siebert-Evenstone et al., 2017), no study had investigated collaborative-inquiry problem solving by examining sequential patterns that occurred in transformative and non-transformative discourse processes. The findings of this study shed new light on sequences in regulation and the co-construction of knowledge. The results suggest that teachers and platforms should not only encourage groups to make their cognitive understandings of the relationships between variables explicit, but also need to make efforts to ensure group members share common understanding. This finding shows the importance of both social aspects as well as cognitive aspects of collaboration. Socio-collaborative skills to establish grounding or mutual understanding are only beneficial when students build on each other's ideas continuously and keep checking their output for correctness. This delicate balance between support needed to provide and connect the social, emotional and cognitive dimensions of collaboration remains an important direction for research (Ludvigsen, 2016). This study also applied transition rate analysis and entropy analysis to the CSCL community. Because it focuses on single-state transability, transition rate analysis can provide sequential information that is much more granular than the information provided by lag-sequential analysis and epistemic network analysis. In addition, transition-rate analysis and entropy analysis are easier to implement than are statistical-discourse analysis and hidden Markov modelling, and their results are easier to understand. These methods can easily become new tools for CSCL researchers. The recent rise of educational technology and the power of multimodal learning analytics make this an ideal time to explore how to use temporal analysis to scaffold real-time collaboration and knowledge building. Since it is a demanding task for teachers to monitor and regulate multiple collaborating groups at the same time (Van Leeuwen, Janssen, Erkens, & Brekelmans, 2015), the automated moderation of discussions on both the socio-collaborative and cognitive dimensions of collaboration could be beneficial. Future research could develop a cloud-based computational tool that monitors collaboration behaviors that are not conducive of productive group work—e.g. "silent too long," "one collaborator is too over-bearing," "team stuck/not making a decision," "no sufficient elaboration of one's own or a partner's reasoning." On the basis of student inputs—e.g. their gazes or their verbal conversations (as recorded by automated speech-recognition technology)—such a system could intervene by prompting team members on an as-needed basis to get back on track.

This study has several limitations, however. First, the data were collected during a 60-min session in which students tried to solve four tasks related to Ohm's Law. Although there were materials to help the students understand the tasks and the simulation environment, the participants could not become experienced collaborators or skillful operators of the platform in such a short time. Their discourse patterns may have differed from those of students who have engaged in this environment for much longer—months or years. In addition, the only way for the members of the small teams of participants to communicate was via the chat window embedded in the platform, and the students had no idea who their team members were. This means of communication may differ from how people communicate via asynchronous platforms using multi-media, including pictures, audios, and videos. For these reasons, the results of this study should be interpreted with caution.

6. Conclusion

This study analyzed sequential patterns in the chat message generated by 144 two-year colleges students solving four tasks related to Ohm's Law. The tasks were categorized as "successful" or "less successful" depending on whether they were solved. All the chat messages were automatically coded as either "transformative" or "non-transformative." We used transition-rate, social entropy, and sequential pattern analyses. The results revealed no clear difference in the overall diversity of the discourse types observed in the two conditions. However, in the "successful" condition, there was a higher probability of transitions from "*proposition generation*" to "*orientation*," from "*interpretation and conclusion*" to "*experimentation*," and from "*proposition generation*" to "*sustaining mutual understanding.*" In the "less successful" condition, in contrast, the transitions from "*proposition generation*" to "*regulation*" and from "*proposition generation*" to "*experimentation*" was much higher.

The findings of this study have implications for instructional designers, teachers, and students. We attempted to explore what the desired collaboration may look like by analyzing the sequential patterns of transformative and non-transformative discourse of successful and less successful groups. The sequential patterns of successful groups may reveal the characteristics of the desired states. For example, to succeed, a group may need to achieve mutual understanding after a group member proposing the relationship between variables and need to elaborate on propositions to improve the group members' cognitive understanding. This mechanism can be applied at the computer level, helping designers to develop metacognitive tools and guiding tools to facilitate students' collaboration. This mechanism can also be applied at the human level. If the importance of these processes is conveyed to students who engage in collaborative problem solving, the students may try to ensure their mutual understanding and elaboration of propositions. Or teachers may try to provide supports to students if they notice their students are away from the desired states. Future work could focus on how to use these results to implement pedagogical and technical designs that can support students' problem solving and knowledge construction. For instance, such designs could encourage students to elaborate on their propositions. Another direction could be the collection and integration of multimodal data—e.g. log data, speech information, facial expressions, and bodily movements—into analyses to better understand a student's status and to better support them.

Declarations of interest

None.

Acknowledgements

Thanks to the Concord Consortium, especially Dr. Paul Horwitz and Cynthia McIntyre for supporting this work. We would also like to thank the children and teachers involving in this study for making it possible.

References

Andrade, A., Danish, J. A., & Maltese, A. V. (2017). A measurement model of gestures in an embodied learning environment: Accounting for temporal dependencies. *Journal of Learning Analytics, 4*(3), 18–46.
Andriessen, J., Baker, M., & Suthers, D. (Eds.). (2003). *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments*. Dordrecht: Kluwer.
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.
Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research, 87*(2), 309–344.
Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., & Rumble, M. (2009). Developing 21st century skills and assessments. *White paper from the assessment and learning of 21st century skills project*.
Belz, J. A. (2003). Linguistic perspectives on the development of intercultural competence in telecollaboration (1). Language. *Learning & Technology, 7*(2), 68.
Bruno, R. J. (2010). *Social differentiation, participation inequality and optimal collaborative learning online (unpublished doctoral thesis)*. West Lafayette, USA: Purdue university.
Chen, B., Resendes, M., Chai, C. S., & Hong, H. Y. (2017). Two tales of time: Uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interactive Learning Environments, 25*(2), 162–175.
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.
Chiu, M. M., & Lehmann-Willenbrock, N. (2016). Statistical discourse analysis: Modeling sequences of individual actions during group interactions across time. *Group Dynamics: Theory, Research, and Practice, 20*(3), 242.
De Jong, T. (2006). Technological advances in inquiry learning. *Science, 312*(5773), 532–533.
Dillenbourg, P. (1999). What do you mean by "collaborative learning"? In P. Dillenbourg (Ed.). *Collaborative learning: Cognitive and computational approaches* (pp. 1–16). Amsterdam, NL: Pergamon, Elsevier Science.
Donnelly, D. F., Linn, M. C., & Ludvigsen, S. (2014). Impacts and characteristics of computer-based science inquiry learning environments for precollege students. *Review of Educational Research, 84*(4), 572–608.
Epp, C. D., Phirangee, K., & Hewitt, J. (2017). Talk with me: Student pronoun use as an Indicator of discourse health. *Journal of Learning Analytics, 4*(3), 47–75.
Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior, 21*(3), 463–486.
Fischer, F., Hmelo-Silver, C. E., Goldman, S. R., & Reimann, P. (2018). *International handbook of the learning sciences.* Routledge.
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.
Foster, J., & Yaoyuneyong, G. (2016). Teaching innovation: Equipping students to overcome real-world challenges. *Higher Education Pedagogies, 1*(1), 42–56.
Gijlers, H., & de Jong, T. (2009). Sharing and confronting propositions in collaborative inquiry learning. *Cognition and Instruction, 27*(3), 239–268.
Henri, F. (1992). Computer conferencing and content analysis. *Collaborative learning through computer conferencing* (pp. 117–136). Berlin, Heidelberg: Springer.
Hmelo-Silver, C. E., Chernobilsky, E., & Jordan, R. (2008). Understanding collaborative learning processes in new learning environments. *Instructional Science, 36*(5–6), 409–430.
Jeong, H., Hmelo-Silver, C. E., Jo, K., & Shin, M. (2016). CSCL in STEM education: Preliminary findings from a meta-analysis. *2016 49th Hawaii international conference on system sciences (HICSS)* (pp. 11–20). IEEE.
King, J., & South, J. (2016). Reimagining the role of technology in higher education: A supplement to the national education technology plan. http://static.politico.com, Accessed date: 31 July 2018.
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., Kirschner, F., & Zambrano, J. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning,* 1–21.
Knight, S., Wise, A. F., & Chen, B. (2017). Time for change: Why learning analytics needs temporal analysis. *Journal of Learning Analytics, 4*(3), 7–17.
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. V. Y., & Tan, S. C. (2017). Promising ideas for collective advancement of communal knowledge using temporal analytics and cluster analysis. *Journal of Learning Analytics, 4*(3), 76–101.

Lim, J., & Liu, Y. (2006). The role of cultural diversity and leadership in computer-supported collaborative learning: A content analysis. *Information and Software Technology, 48*(3), 142–153.

Lovelace, K., Shapiro, D. L., & Weingart, L. R. (2001). Maximizing cross-functional new product teams' innovativeness and constraint adherence: A conflict communications perspective. *Academy of Management Journal, 44*(4), 779–793.

Ludvigsen, S. (2016). CSCL: Connecting the social, emotional and cognitive dimensions. *International Journal of Computer-Supported Collaborative Learning, 11*(2), 115–121.

Lund, K., Quignard, M., & Shaffer, D. W. (2017). Gaining insight by transforming between temporal representations of human interaction. *Journal of Learning Analytics, 4*(3), 102–122.

Macdonald, J. (2003). Assessing online collaborative learning: Process and product. *Computers & Education, 40*, 377–391.

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.

Matei, S., Oh, K., & Bruno, R. (2006 Nov). *Collaboration and communication: A social entropy approach.* San Antonio: Paper presented at the National Communication Association Annual Convention.

Matei, S. A., Bruno, R., & Morris, P. (2015). Visible effort: Visualizing and measuring group structuration through social entropy. In S. A. Matei, M. Russell, & E. Bertino (Eds.). *Transparency in social media—Tools, methods and algorithms for mediating online interactions*. New York: Springer.

Molenaar, I. (2014). Advances in temporal analysis in learning and instruction. *Frontline Learning Research, 2*(4), 15–24.

Molenaar, I., & Chiu, M. M. (2017). Effects of sequences of cognitions on group performance over time. *Small Group Research, 48*(2), 131–164.

Mulder, M. (2014). Conceptions of professional competence. *International handbook of research in professional and practice-based learning* (pp. 107–137). Dordrecht: Springer.

Noroozi, O., Weinberger, A., Biemans, H. J., Mulder, M., & Chizari, M. (2012). Argumentation-based computer supported collaborative learning (ABCSCL): A synthesis of 15 years of research. *Educational Research Review, 7*(2), 79–106.

Parkes, M., Stein, S., & Reading, C. (2015). Student preparedness for university e-learning environments. *The Internet and Higher Education, 25*, 1–10.

Popov, V., Biemans, H. J. A., Brinkman, D., Kuznetsov, A., & Mulder, M. (2013). Facilitation of Computer-Supported Collaborative Learning in Mixed- versus Same-Culture Dyads: Does a Collaboration Script Help? *The Internet and Higher Education, 19*, 36–48.

Popov, V., van Leeuwen, A., & Buis, S. C. A. (2017). Are you with me or not? Temporal synchronicity and transactivity during CSCL. *Journal of Computer Assisted Learning, 33*(5), 424–442.

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 June.

Reimann, P. (2009). Time is precious: Variable-and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 4*(3), 239–257.

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. *Cambridge handbook of the learning sciences* (pp. 397–417). .

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.

Shaffer, D. W., Hatfield, D., Svarovsky, G., Nash, P., Nulty, A., Bagley, E., ... Mislevy, J. R. (2009). Epistemic network analysis: A prototype for 21st century assessment of learning. *The International Journal of Learning and Media, 1*(2), 33–53.

Shannon, C. E., & Weaver, W. (1998). *The mathematical theory of communication.* Urbana: University of Illinois Press.

Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2017). In search of conversational grain size: Modelling semantic structure using moving stanza windows. *Journal of Learning Analytics, 4*(3), 123–139.

Soller, A., Martínez, A., Jermann, P., & Muehlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal of Artificial Intelligence in Education, 15*(4), 261–290.

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. In R. K. Sawyer (Ed.). *Cambridge handbook of the learning sciences* (pp. 409–426). Cambridge, UK: Cambridge University Press.

Tang, H., Xing, W., & Pei, B. (2018a). Time really matters: Understanding the temporal dimension of online learning using educational data mining. *J. Educ. Comput. Res.* 0735633118784705.

Tang, H., Xing, W., & Pei, B. (2018b). Exploring the temporal dimension of forum participation in MOOCs. *Distance Education, 39*(3), 353–372.

Tancig, S. (2009). Expert team decision-making and problem solving: Development and learning. *Interdisciplinary Description of Complex Systems: INDECS, 7*(2), 106–116.

Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2015). Teacher regulation of cognitive activities during student collaboration: Effects of learning analytics. *Computers & Education, 90*, 80–94.

Walther, J. B. (1997). Group and interpersonal effects in international computer-mediated collaboration. *Human Communication Research, 23*(3), 342–369.

Waugh, M., & Su-Searle, J. (2014). Student persistence and attrition in an online MS program: Implications for program design. *International Journal on E-Learning, 13*(1), 101–121.

Wise, A. F., & Chiu, M. M. (2011). Analyzing temporal patterns of knowledge construction in a role-based online discussion. *International Journal of Computer-Supported Collaborative Learning, 6*(3), 445–470.

Yang, X., Li, J., Guo, X., & Li, X. (2015). Group interactive network and behavioral patterns in online English-to-Chinese cooperative translation activity. *The Internet and Higher Education, 25*, 28–36.

Yang, X., Li, J., & Xing, B. (2018). Behavioral patterns of knowledge construction in online cooperative translation activities. *The Internet and Higher Education, 36*, 13–21.

Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge-building communities. *The Journal of the Learning Sciences, 18*(1), 7–44.