Bridging Learning Sciences, Machine Learning, and Affective Computing for Understanding Cognition and Affect in Collaborative Learning

Sanna Järvelä[1], Dragan Gašević[2], Tapio Seppänen[1] , Mykola Pechenizkiy[3] & Paul A. Kirschner[1&4]

[1]University of Oulu, [2]Monash University, [3] Eindhoven University of Technology, [4] Open University The Netherlands

Correspondence concerning this article should be addressed to Sanna Järvelä, sanna.jarvela@oulu.fi; Department of Educational Sciences and Teacher Education, Learning and Educational Technology Research Unit (LET), P.O. Box 2000, FIN-90014 University of Oulu, Phone +358 40 5777164

Biografical information

Sanna Järvelä is a professor of Learning Sciences and Educational technology at the University of Oulu. Her research interests are self-regulated learning, computer supported collaborative learning and multimodal methods

Dragan Gašević is professor of Learning Analytics at the Monash University. His research interests are learning analytics, educational data mining, self-regulated learning and collaborative learning.

Tapio Seppänen is professor of Biomedical Engineering at the University of Oulu. His research interests are intelligent signal analysis and pattern recognition in health and wellbeing applications.

Mykola Pechenizkiy is a professor in Data Mining at the Eindhoven University of Technology. His research interests are data science, knowledge discovery and data mining.

Paul Kirschner is Professor Emeritus at the Open University of the Netherlands and visiting professor at the University of Oulu. His research interests are learning and instruction, interaction in learning, and collaboration for learning (computer supported collaborative learning).

Abstract

Collaborative learning can be a powerful method for sharing understanding between learners. To this end, strategic regulation of processes, such as cognition and affect (including metacognition, emotion, and motivation) is key. Decades of research on self-regulated learning has advanced our understanding about the need for and complexity of those mediating processes in learning. Recent research has shown that it is not only the individual's but also the group's shared processes that matter and, thus, that regulation at the group level is critical for learning success. A problem here is that the "shared" processes in collaborative learning are invisible, which makes it almost impossible for researchers to study and understand them, for learners to recognize them, and for teachers to support them. Traditionally, research has not been able to make these processes visible nor has it been able to collect data about them. With the aid of advanced technologies, signal processing, and machine learning, we are on the verge of "seeing" these complex phenomena and understanding how they interact. We posit that technological solutions and digital tools available today and in the future will help advance the theory underlying the cognitive, metacognitive, emotional, and social components of individual, peer, and group learning when seen through a multidisciplinary lens. The aim of this paper is to discuss and demonstrate how multidisciplinary collaboration among the learning sciences, affective computing, and machine learning is applied for understanding and facilitating collaborative learning.

Keywords

Collaborative learning, regulation in learning, multimodal methods, multidisciplinary collaboration, learning sciences, affective computing, machine learning

Structured practitioner notes

What is already known about this topic

- Collaborative learning occurs when team-members systematically activate, sustain and regulate their cognition, motivation, emotions and behaviors towards the attainment of their goals.
- Socially shared regulation in learning contributes to success in collaborative learning.

What this paper adds

- ‘Shared’ processes in collaborative learning are hard for researchers to study and understand them, for learners to recognize them, and for teachers to support them.
- Multimodal data collection provides opportunities to study multiple aspects of student behaviors and regulation processes.
- With the aid of advanced technologies multidisciplinary collaboration between the learning sciences, affective computing and machine learning can help to study these complex phenomena.

Implications for practice and/or policy

- The case examples demonstrate how multidisciplinary collaboration can meet the challenges in understanding and facilitating collaborative learning.

- Multidisciplinary efforts with multimodal data will contribute to collaborative learning practice by providing theoretically-informed feedback and personalized support.

Introduction

In our complex interconnected world, we increasingly look to collaboration to solve the diffuse and wicked problems we face (Rittel & Webber, 1973). Collaborative learning (CL) affords us opportunities to migrate from individual learning to learning and working together in teams (Miyake & Kirschner, 2014). In collaborative teams, members attempt to solve problems together to make use of each other's knowledge and expertise.

Traditionally, people learning together in a team are unrelated to each other (Fransen, Kirschner, & Erkens, 2011; Fransen, Weinberger, & Kirschner, 2013). Though from the outside they may appear to fluently interact socially with each other, this often is not the case. This is because such learning requires socio-cognitive interactions with others to achieve individual and shared cognitive growth (Kirschner, Sweller, Kirschner, & Zambrano, 2018). Research has shown that successful CL only occurs when team members systematically activate and sustain their cognition, motivation, behavior, and emotions toward the attainment of their goals (Schunk & Greene, 2017).

The problem is that the core learning processes in the human mind (cognition, metacognition, emotions, and motivation), while not visible to others (team mates or researchers) are critical and need to be shared in collaboration (Järvelä & Hadwin, 2013). Strategic regulation of those invisible processes is key to successful learning (Järvelä, Hadwin, Malmberg, & Miller, 2018). A problem here is that the "shared" processes in CL (Roschelle & Teasley, 1995) are invisible and difficult for researchers to understand, for learners to recognize, and for teachers to support. We, thus, need to help people become able to "read each other's minds" and to make thinking, affective, and even metacognitive processes visible and accessible for better regulation.

Traditionally, research has not been able to make these processes visible nor has it been able to collect data about them. Currently, with the aid of advanced technologies for collecting and processing vast amounts of heterogeneous multimodal data with affective computing, signal processing, and machine learning, we are on the verge of "seeing" these complex phenomena. To unlock the full potential for these recent technological developments, we need fruitful multidisciplinary collaboration between the learning sciences, affective computing, and machine learning. The aim of this paper is to discuss and demonstrate through two examples how multidisciplinary collaboration between these fields is applied for understanding and facilitating CL.

Regulation in learning for overcoming challenging situations in collaboration

Research on CL has concentrated on different conditions of CL, processes of CL, and students' collaborative discourse and argumentation (e.g., O'Donnell & Hmelo-Silver, 2013). Collaboration is not spontaneous and is dependent on individual group members' capabilities to collaborate with each other. The interactive nature of collaboration (Isohätälä, Näykki, Järvelä, & Baker, 2017) has the potential to support the groups' learning success (Kirschner, Paas, Kirschner, & Janssen, 2011) but also to distract the collaboration (Janssen, Erkens, Kirschner, & Kanselaar, 2012) by causing new types of challenges. Those challenges are the result of external (e.g., task difficulty), internal (e.g., weak strategies), and/or social (e.g., socioemotional) conditions (Järvenoja, Volet, & Järvelä, 2013). Recent research has shown that not only self-regulation (SRL), but also co-regulation of learning (CoRL) and socially shared regulation of learning (SSRL) (Hadwin, Järvelä, & Miller, 2018; Järvelä & Hadwin, 2013) are especially critical.

Perspectives on forms of regulation are built on decades of research and theory development (Zimmerman & Schunk, 2011). Self-regulation is a cyclical complex metacognitive and social process that involves adapting thought, motivation, emotion, and behavior (Cleary & Zimmerman, 2012; Winne & Hadwin, 1998). Self-regulation is critical for successful individual and collaborative learning (Dignath & Büttner, 2008), but the problem is that the processes at the foundation of regulation are invisible and, thus, very challenging to understand, support, and influence. Regulation of learning is difficult at the individual level but is even more difficult when it is expected to occur collaboratively in groups where individual processes must be coordinated while planning and execution must be communicated between individuals (Kirschner et al., 2018).

Regulation of learning that takes place within collaborative groups presumes students have a precise understanding and awareness of the sources of the emerging regulation needs, “what is going wrong”, and of the *targets for regulation*. With strategic regulation, group members can influence various cognitive, motivational, and emotional processes taking place in the group and adjust their collaborative learning according to the task goals and shared standards. Socially shared regulation, particularly, assumes reciprocity in regulatory actions between the group members. SRL (regulating oneself), CoRL (supporting each other), and SSRL (regulating together) jointly form the relative regulated learning space in which individuals within the group regulate their own motivation, emotions, cognition, and behavior and simultaneously contribute to the groups’ shared regulatory processes (Hadwin et al., 2018).

Research has shown that groups do not recognize and react to challenging learning situations and, thus, cannot explicitly regulate their learning in such situations (Järvelä et al., 2016; Rogat & Adams-Wiggins, 2015). This means that there is a need for them to be alerted. If we were able to understand these challenging learning situations, targeted support for metacognitive monitoring and regulation of cognition, emotion, and motivation could be provided to help learners adapt their learning. If learners fail to optimize their strategic regulation, this failure can lead to nonfunctional groups and maladaptive individual and group learning performance.

Multimodal data for understanding regulation in a collaborative learning process

In attempts to understand the learning process, certain assumptions about learning become grounding principles in data collection, namely that the learning process is composed of discrete observable cognitive, metacognitive, motivational, and affective events that occur in a sequential temporal order (Bernacki, 2017). According to SRL theories (Zimmerman & Schunk, 2011), the learning process is organized into phases, and learners repeatedly progress through these phases in a cyclical and iterative fashion until the task engagement concludes and, ideally, the learning goal is achieved.

The use of multimodal data (e.g., psychophysiological, eye tracking, textual dialog transcripts, and interaction with technology) offers a significant promise for addressing the problem of understanding regulation in collaboration (see Järvelä & Bannert, 2019). We have been working on combining the powerful capabilities of affective computing and machine learning with the solid theoretical foundations from the learning sciences to make use of these data sources and make the invisible learning processes involved in collaboration visible for learners, teachers, and researchers. Until now, each of these fields has made progress in capturing learning instances from their own perspectives, but as this has been done in isolation from other fields, the desired progress has not been achieved. Our ability to connect the three fields can reveal hidden human mental processes, model mediating processes, and critical patterns of CL processes.

While learning is a complex process, it is also multimodal in terms of representations evidencing its cognitive, affective, and social aspects (Ochoa, Worsley, Weibel, & Oviatt, 2016). Multimodal research data are data that originate from different data channels that can be either subjective or objective (Ochoa, 2017). Multimodal approaches in learning science research can help tackle the constraints of typical single-channel data (e.g., self-report or protocol analysis) and help draw more valid and reliable inferences about the learning processes (Harley, Bouchet, Hussain, Azevedo, & Calvo, 2015; Pantic & Rothkrantz, 2003). Digital technologies offer a number of different opportunities or affordances to both record and facilitate multimodality in learning. These affordances can be (Kraut, Fussel, Brennan, & Siegel, 2002; Kreijns, Kirschner, Jochems, & van Buuren, 2011):

- audible—hearing sounds in digital environments;

- visible—seeing objects or people in digital environments;
- tangible—touching or clicking objects in digital environments;
- presence-related—sharing the same space with others and/or sensing each other's presence (i.e., social presence in digital environments);
- temporal—working with others in digital environments at the same time or at different times;
- reviewable—accessing messages in digital environments again and again (i.e., having a tangible history); and
- revisable—repeatedly updating the messages in digital environments.

Furthermore, a great deal of learning occurs in digital environments (Drysdale, Graham, Spring, & Halverson, 2013), where learners typically generate a great amount of cognitive, metacognitive, emotional, and motivational data on what is attended to and studied, the order in which this occurs, how much time was spent on what, when and where certain actions occurred, and so forth (Azevedo, Taub, & Mudrick, 2017). The new era of big data in education (Baker & Siemens, 2014) and fine-grained data that come from online learning and computer-based learning environments (Azevedo, Moos, Johnson & Chauncey, 2010) provide opportunities to study multiple aspects of student behavior and regulation processes that otherwise would have been difficult to gather with earlier existing methods (e.g., self-reports). These methods can significantly extend current knowledge on the sequential and temporal nature of the complex learning processes (Azevedo & Gašević, 2018). For example, the multimodal data captured by these technologies can provide new complementary data channels for making visible and understanding important phases of regulated learning as they occur (Harley et al., 2015).

Multidisciplinary collaboration for moving beyond the state of the art in collaborative learning

Since more complex and larger amounts of data are now available than ever in the past, we have wrestled with limitations in our own thinking as well as with problems related to data handling and analysis. The methods and tools for learning analytics have helped us trace and model SRL processes and understand CL in authentic learning contexts (Gašević, Dawson, & Siemens, 2015; Malmberg et al., 2018; Pijeira-Diaz, Drachsler, Järvelä, & Kirschner, 2018). We posit that the technological solutions and digital tools available today and in the future will help advance the theory underlying the cognitive, metacognitive, emotional, and social components of individual, peer, and group learning when seen through a multidisciplinary lens.

The learning sciences help us understand human learning processes (Miyake & Kirschner, 2014), while SRL, CoRL, and SSRL theories help us understand the strategic regulation of those processes in learning (Hadwin et al., 2018). Affective computing techniques, which aim to develop algorithms for automatic affect recognition from nonverbal human communication (Picard, 1997), are used to digitally capture these processes and factors in a minimally intrusive yet objective way. They allow us to extract information from physiological parameters related to autonomic nervous system activity, voice parameters, and facial expressions, complemented with observed behavior in video. Promising research in other sciences has illustrated how data science and machine learning in particular are transforming scientific discovery (e.g., Peek, Combi, Marin, & Pellazzi, 2014).

However, there are several fundamental problems regarding pure data-driven discovery, especially when dealing with an insufficient amount of representative (accurately annotated) training samples along with a large number of predictive and irrelevant features while the modeled phenomena show complex and nonstationary patterns of behavior that evolve over time. Traditional model selection and cross-validation procedures may fail to capture the true relationships and show misleading results, including spurious relationships. Thus, the hope is that theory-informed machine learning methods will allow for more accurate detection of learning processes by accounting for cognitive, affective, metacognitive, and behavioral processes of learning in solo and group learning activities by analyzing multichannel data. Through theory-informed machine learning, we assume that relevant theories from learning sciences drive (Wise & Shaffer, 2015) (i) the choice of data representation or measures used as input into machine learning and (ii) interpretation of the results of the application of machine learning. For example, Karpatne et al. (2017) illustrated the principles of theory-guided data science, including scientific consistency, as an essential component for learning generalizable and scientifically interpretable models facilitating the discovery of novel domain insights. Bauckhage et al. (2018) introduced the idea of treating data-driven learning as a problem of sequencing predefined operations that allow for incorporating expert knowledge and lead to traceable or explainable decision-making systems. In this paper, we show a case (Case 2) illustrating the promise of using mixed machine learning methods and combining social network analysis and epistemic network analysis in learning analytics.

Despite the growing availability of opportunities to collect and analyze learning process data, research has not yet fulfilled its potential to uncover the cognitive and affective aspects of SRL and CL. Machine learning might be the key here, as it can make use of existing data and develop new algorithms (Baker & Siemens, 2014), which ultimately will have a positive effect on the learner/teacher and their learning/teaching. Machine learning methods applied to SRL and CL focus primarily on detection of knowledge building processes from collaborative discourse (Kovanović et al., 2016) and identification of learning strategies, motivation, and affective states from logs of individual learners' interactions with technology (D'Mello, 2017; Jovanović, Gašević, Dawson, Pardo, & Mirriahi, 2017; Winne & Baker, 2013). A limitation is that most existing machine learning research has focused on analyzing single data streams (e.g., traces of interaction with technology). A second limitation is that existing approaches are often insufficiently connected with theoretical principles from the learning sciences, which may potentially lead to invalid interpretations and limited generalizability. Finally, there is little support for understanding the developmental progression of SRL and CL skills. Therefore, progress in machine learning is necessary to support our goal of understanding and supporting self-regulated learning and CL.

The limitations of existing machine learning approaches for the analysis of integrated multiple data streams in CL might be overcome by building on expertise in affective computing. In this paper, we show a case (Case 1) illustrating how different data streams can be visualized and navigated simultaneously. Data for extracting affective cues were collected with one or more measurement modalities, such as video cameras, microphones, and on-body sensors, extracting facial expressions, body positions, voice characteristics, autonomic nervous system activities, and central nervous system activities. Signal and image analysis-algorithms were used to characterize the data by extracting affect-specific features from each modality. Machine learning was then applied to model the feature sets of relevant affects to enable their detection in any new data.

Studies have shown very high performance levels for human accuracy in affect recognition in many applications, such as emotion recognition from voice detection (Väyrynen, Kortelainen, & Seppänen, 2013) and spontaneous micro-expression detection from facial videos (Pfister, Li, Zhao, & Pietikäinen, 2011). Affective computing, thus, may have the potential to measure learning processes using smart sensoring techniques and to transform them into visible information that can be interpreted by the teacher/trainer. For example, in our recent research, we demonstrated for the first time that affective cues can be measured in a group working scenario of students (Haataja, Malmberg, & Järvelä, 2018; Malmberg et al., 2018). In the following, we present two examples of how progress is being made in the integration of the three fields.

Example 1: Graphical user interface (GUI) for making complex multimodal CL data visible, measurable, and interpretable

Our first illustrative example of multidisciplinary collaboration is an effort to simplify the analysis and use of rich multimodal data gathered by learning scientists. This was done by making learners' regulation processes and their accompanying social and contextual reactions visible, measurable, and ultimately interpretable. To facilitate data visualization and processing with respect to the regulation of learning, a graphical user interface (GUI) tool known as SLAM-KIT was developed. SLAM-KIT reveals important features of complex learning environments by allowing users to travel through the learners' data and their statistical characteristics. This kit has been currently used by the researchers, but will have practical implications, as it simplifies complex information and data while making them available through visualization and analysis to both teachers and learners (for details, see Noroozi et al., 2018).

The SLAM-KIT is an integrated analysis tool for the learning sciences that simplifies the use and analysis of rich multimodal data recorded from CL situations. It enables users to dive into recorded interaction situations not only via the video and audio of the situation but also through a wealth of physiological information (i.e., body temperature, electrodermal activity [EDA], blood volume pulse [BVP], and heart rate) captured during the situation using unobtrusive sensors and cameras. The tool merges the diverse data sources and provides a unified navigable view of the entire interaction. The video and audio are aligned with sensor data, which is then presented as color-coded plots on the same timeline. The goal is to provide learning scientists with easy access to all of the physiological information relevant in the regulation of CL to analyze the learning situations.

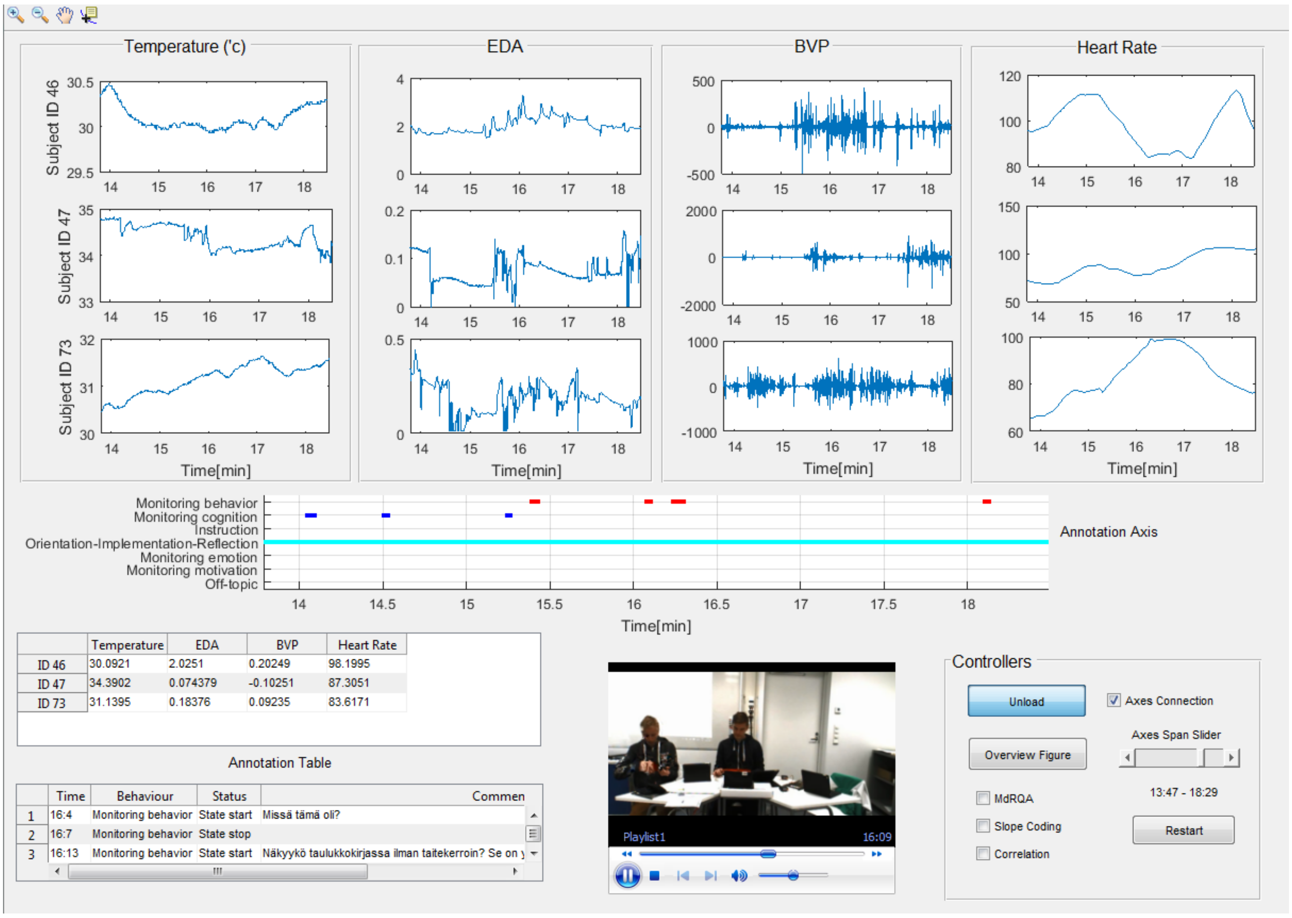


Figure 1. SLAM-KIT tool graphical user interface

Figure 1 presents the GUI of the SLAM-KIT tool in a subject view mode of a measurement session. The top rows of subfigures show the synchronously acquired physiological signals for three subjects. In the middle, an annotation axis plot shows the annotation labels for time, duration, and type. Additionally, more information about the labels is given in the annotation table (lower left-hand corner), and above this table, quantitative information is given on the analysis of the signals. The video playback panel is synchronized with the physiological signal subfigures and shows the recorded video with standard playback controls. The controller panel (lower right-hand corner) contains additional options for an advanced user. For example, this first top row indicates that there appear to be simultaneous activity changes in EDA, BVP, and heart-rate signal in the particular session.

We can shed further light on the simultaneously occurring changes in the biosignals with the SLAM-KIT by computing, for example, recurring plots of the data using multidimensional recurrence quantification analysis (MdRQA) (Wallot, Roepstorff, & Mønster, 2016), slope coding (Leite, Henriques, Martinho & Paiva, 2013), and the Pearson correlation coefficient among the students' signals. The tool can be used to study group behavior among subjects, as MdRQA analysis indicates any simultaneous activities in their biosignals and, thus, reflects possible interactions and synchronized affects. These algorithms can be run on the entire learning session or a selected shorter segment of the session by controlling the time axes.

Regulation in learning refers to an intentional process over which individuals exert agency by planning for and strategically responding to learning tasks and situations. It is a highly metacognitive process involving self-monitoring and self-control activated by regulatory processes that develop cyclically across phases of task engagement (Hadwin et al., 2018). The SLAM-KIT tool, as such, does not help analyze regulation in collaboration, but its functions have helped us explore the complexity of the dynamic nature of regulation (Winne, 2019). It has allowed us to contextualize physiological data, explore the characteristics and temporal aspects of the physiologically activating situations, triangulate different data channels, and identify challenging episodes in CL for further analyses that examine regulatory patterns and contingencies.

We have been working with collaborative science inquiries, and our multimodal data collection covers video recordings, physiological data, log data, situated self-reports, and learning outcomes (Järvelä, Järvenoja & Malmberg, 2019). For example, SLAM-KIT helped us identify the challenging episodes in collaborative work (videos) and each individual student's associated EDA signals. Figure 2 is an illustration of three group members' SCR (skin conductance response) peaks (levels -1–4) and the standardized SCR peaks (level 0) that occurred in a situation where group members confronted a challenge in their inquiry (evidenced in the video). The example shows that one student's strong frustration (right in a picture) created their own simultaneous physiological reaction (EDA), and in the current situation, the other group members followed that arousal instead of activating regulation. The situation as such gave us an opportunity to focus on the core mechanisms in regulation, such as adaptation (e.g., is the group facing challenges, are they aware of it, do they take the necessary steps to overcome it). In this vein, we have progressed further with larger multimodal data and explored physiological synchrony (Malmberg et al., 2019) and state transitions discovered by machine learning methods for understanding SSRL (Sobocinski et al., 2019).

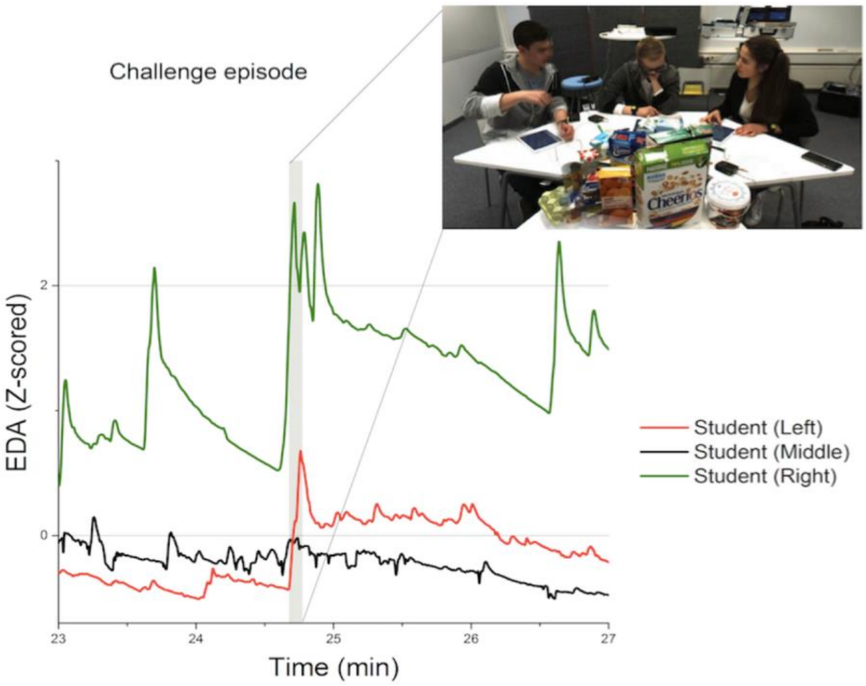


Figure 2. The challenge episode associated with EDA signals of three students.

Example 2: Social and epistemic network signature (SENS) for combining social and cognitive dimensions in CL

Our second illustrative example of multidisciplinary collaboration is an approach that supports learning scientists in the use of machine learning and analytic techniques to advance understanding of connections between social and cognitive dimensions of CL (Gašević, Joksimović, Eagan, & Shaffer, 2019). This approach analyzes connections of social and cognitive dimensions by looking at how learners *regulate* their CL by playing different roles, whereby the learning science literature views these roles in terms of how learners interact with the relevant people about relevant content (Dillenbourg, Järvelä, & Fischer, 2009). Therefore, the approach assumes that relevant dimensions of CL can be modeled as networks. Network modeling is done as a combination of social network analysis (SNA) and epistemic network analysis (ENA) (Shaffer et al., 2016) to form the social and epistemic network signature (SENS) approach (see Figure 3).

Discourse is a typical source of data regarding the cognitive dimension of CL. Traditionally, learning scientists analyze discourse by following inductive or deductive coding approaches. The use of machine learning in SENS offers learning scientists techniques that can automate the coding process. This is performed by using supervised and unsupervised machine learning methods. *Supervised machine learning techniques* automate *deductive coding* by assigning codes to relevant analysis units of discourse according to a predefined codebook or coding scheme. *Unsupervised machine learning techniques*, such as latent Dirichlet allocation, *automate the inductive coding process*. That is, they enable the learning scientist to detect themes that emerge from data and interpret the themes against relevant theoretical frameworks.

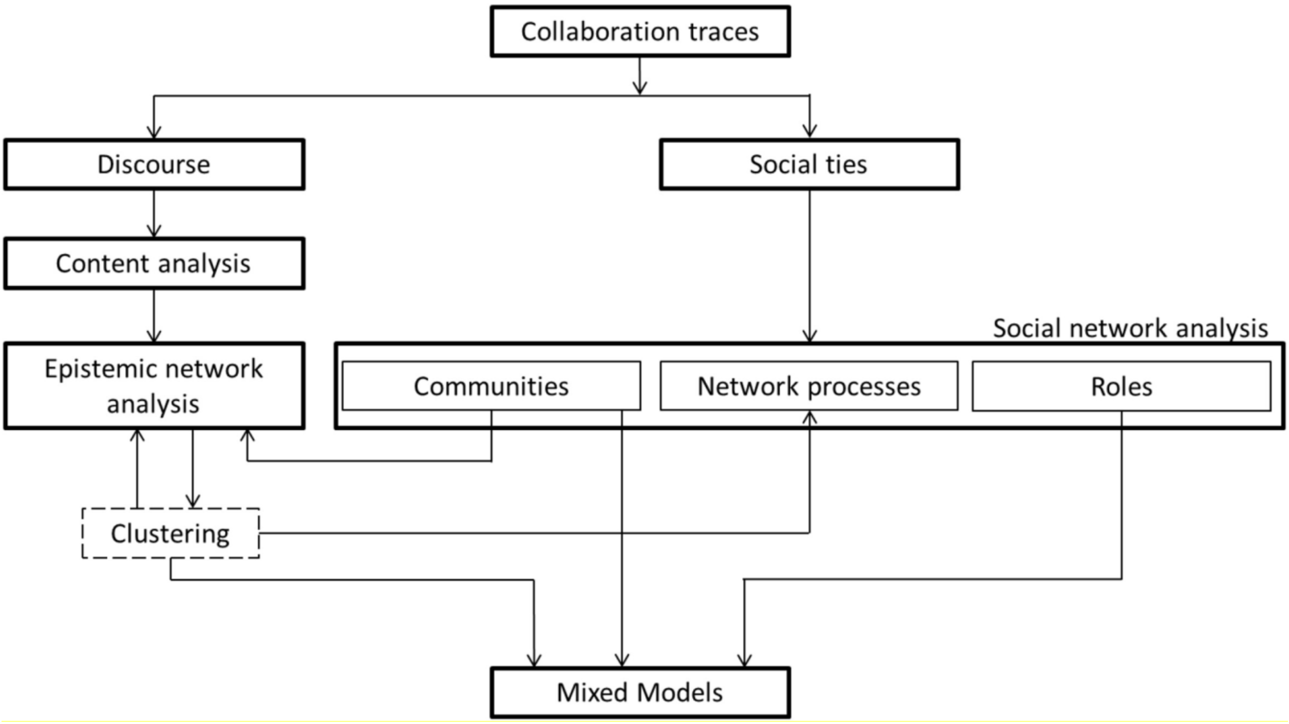
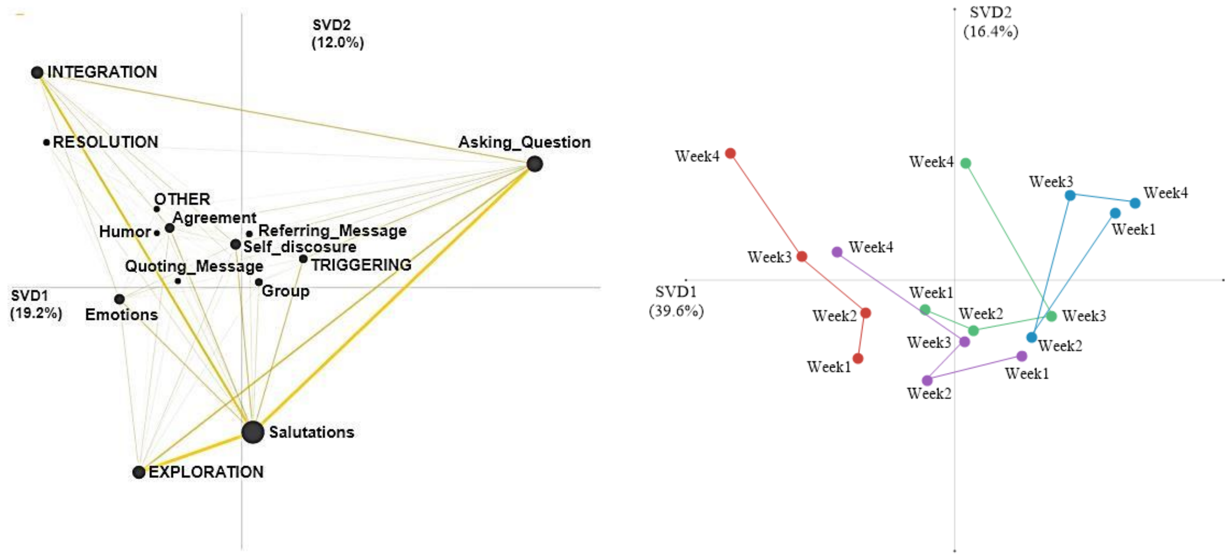


Figure 3: The combination of social network analysis and epistemic analysis with relevant machine learning and data analytic approaches to form the SENS approach.

The SENS approach does not recommend counting of coded discourse data, which is commonly done in the learning science literature due to its grounding in the theory of epistemic frames (Shaffer, 2006). The theory posits that the cognitive dimension of CL is not a set of isolated processes, skills, and values but rather that the cognitive dimension of CL should be studied through connections among processes, skills, and values. Specifically, SENS creates co-occurrence networks of the codes used for coding collaborative discourse. Epistemic networks are formed based on the co-occurrence of relevant codes within a segment of discourse data, where the segment is determined by research questions the learning scientist has. An example research question could look at how learning progresses over several weeks of CL. In such a case, codes that co-occur in messages posted by a learner each day of CL could be connected to form an epistemic network. Such networks are then processed by a dimensionality reduction technique, as commonly used in machine learning and affective computing, to enable the learning scientist to project networks in an analytic space to study links between the codes or compare how different groups of learners regulated their CL under different conditions.

Finally, the use of SNA allows for performing graph-based and statistical analyses on social networks extracted from collaboration activities, including role analysis with graph multiplicity measures (e.g., degree or betweenness centrality), network formation processes (e.g., reciprocity of ties), and community detection. SNA can be combined with ENA by using 1) values of epistemic networks as attributes of social network nodes to predict whether students who share similar epistemic frames are likely to collaborate or reciprocate collaboration; 2) communities detected with SNA to check whether there are significant differences in epistemic networks across different communities; and 3) measures produced by ENA and SNA as predictors of performance or learning outcomes.

A notable example of the application of SENS is in research on learners' regulation of CL within communities of inquiry (see Figure 4). A supervised machine learning technique is applied to automate coding of discussion messages according to the coding schemes for social (i.e., 13 indicators categorized into three general categories—interactive, affective, and group cohesion) and cognitive (i.e., triggering events, exploration, integration, and resolution) presence constructs of communities of inquiry (Kovanović et al., 2016; Neto et al., 2018). The association between the phases of cognitive presence and indicators of social presence is then studied (Rolim, Ferreira, Gašević, and Lins, 2019).



a) b)

Figure 4. Epistemic network analysis of the association between cognitive and social presence in communities of inquiry: a) the epistemic network between phases of cognitive presence (capital letters) and indicators of social presence; and b) trajectory analysis of the students in the four conditions across four weeks of discussions – expert-control (red), expert-treatment (purple), practicing researcher-control (blue), and practicing researcher-treatment (green)

The epistemic network in Figure 4a shows that the lower levels of cognitive presence (i.e., the triggering event) were more connected with the indicators of the interactive category of social presence (e.g., asking questions or continuing a thread), while higher levels of cognitive presence (i.e., integration and resolution) were linked with the indicators of the affective category of social presence (e.g., use of humor or self-disclosure). ENA also enabled unveiling of the difference in the links between the social and cognitive presences of the students who were in different intervention groups (i.e., discussion scaffolded with self-regulated learning vs. externally facilitated regulated learning) and different roles assigned (i.e., experts and practicing researchers). The trajectory analysis diagram in Figure 4b indicates that the students who met the self-regulated learning condition and were in the role of researcher did not make much progress in their cognitive inquiry across four weeks of discussions; that is, they did not move toward the left on the ENA chart, showing a lack of progress toward the integration and resolution phases of cognitive presence. For the other three groups, evidence of progress was noted.

Conclusions

Our main message is that bridging learning sciences, affective computing, and machine learning will advance our understanding on affect and cognition, including metacognition, motivation, and emotion in CL, especially the phenomenon dealing with SSRL, and will, thus, help the theory progress and design better learning support. Multidisciplinary efforts will also offer specialized methods for affective computing and machine learning that can harness the wealth of multichannel data regarding CL. These methods will be beneficial for learning scientists in obtaining temporal dimensions of cognition and affect of collaboration and effective regulation practices. Still more work is needed to implement the multidisciplinary solutions for empirical research for collecting and analyzing big learning process data (Molenaar et al., 2019) that can be applied to future AI solutions in order to support human learning (Luckin & Cukurova, 2019). For example, future implications may be that teachers will be able to measure the level of learners' cognitive, affective, and metacognitive engagement and development of SSRL skills based on which they will be able to provide theoretically informed feedback. Then, learners can receive personalized support for the development of their SSRL skills at scale. Technology developers can apply sound theoretical principles and data analysis methods for the development of next-generation learning technologies. Progress in this issue will provide novel automated methods for empowering teachers to provide learner-personalized process feedback at scale (Azevedo et al., 2017).

Multidisciplinary collaboration in the three fields described in this paper will advance the CL theory underlying individual and group learning with respect to cognitive, metacognitive, social, and affective components of learning and how these components shape the results of the learning process. Various measurement modalities will signal more about the multifaceted phenomena (affect and cognition) than a single state, which will also contribute to the ongoing development of self-regulated learning theory (Järvelä & Bannert, 2019). Harnessing the three fields can also solve the problem in the learning sciences of the relatively small amount of data researchers have been able to collect due to the time and effort needed to analyze the data. The result of this limitation is that it is difficult to draw generalizable conclusions. An approach that uses the power of technologies to gather and analyze extensive and highly detailed information about learning and distributing learning analytics grounded in partitions of that big data has promise (Azevedo & Gašević, 2019). Such big data, appropriately analyzed, can inform us how to support learners to make adaptations, to self-regulate, and socially share regulation of learning in collaboration that would otherwise elude them or that adaptive technologies are not designed to make.

Acknowledgements

This study was supported by the Finnish Academy grant 275440.

Oulu University LeaF research infrastructure has been used in data collection.

Statements on ethics and conflict of interest

Data dealing with the Case example 1 has been collected at the Oulu University research infrastructure and the data collection has been accepted by the Oulu University Ethical committee. The dataset in Case example 2 was collected in the authentic classroom setting and the use of data for research was approved by Athabasca University's Research Ethics Board.

There is no conflict of interest.

References

Azevedo, R., & Gašević, D. (2019). Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Computers in Human Behavior, 96*, 207-210. https://doi.org/10.1016/j.chb.2019.03.025

Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational psychologist, 45*(4), 210-223. https://doi.org/10.1080/00461520.2010.515934

Azevedo, R., Taub, M., & Mudrick, N. V. (2017). Understanding and Reasoning about Real-Time Cognitive, Affective, and Metacognitive Processes to Foster Self-Regulation with Advanced Learning Technologies. In P. A. Alexander, S. Dale H., & J. A. Greene (Eds.), *Handbook of Self-Regulation of Learning and Performance* (pp. 275-292). London: Routledge. https://doi.org/10.4324/9781315697048.ch17

Baker, R., & Siemens, G. (2014). Educational Data Mining and Learning Analytics. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 253–272). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139519526.016

Bauckhage, C., Ojeda, C., Schücker, J., Sifa, R., & Wrobel, S. (2018). Informed machine learning through functional composition. *LWDA*, 33-37

Bernacki, M. L. (2017). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: trace data and their analysis. In Schunk, D.H. & Greene, J.A. (Eds.). *Handbook of self-regulation of learning and performance* (pp. 395–412). London: Routledge.

Cleary, T. J., & Zimmerman, B. J. (2012). A cyclical self-regulatory account of student engagement: Theoretical foundations and applications. In S. L. Christenson, A. L. Reschly, &

C. Wylie (Eds.). *Handbook of research on student engagement* (pp. 237–257). Boston: Springer. https://doi.org/10.1007/978-1-4614-2018-7_11.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The Evolution of Research on Computer-Supported Collaborative Learning. In D. N. Balacheff, D. S. Ludvigsen, D. T. de Jong, D. A. Lazonder, & D. S. Barnes (Eds.), *Technology-Enhanced Learning* (pp. 3–19). Berlin: Springer.

D'Mello, S. (2017). Emotional learning analytics. In: Lang, C., Siemens, A., Wise, G., and Gašević, D. (Eds.). *Handbook of learning analytics* (pp. 115-127). SOLAR, Society for Learning Analytics and Research. doi: 10.18608/hla17.010

Drysdale, J. S., Graham, C. R., Spring, K. J., & Halverson, L. R. (2013). An analysis of research trends in dissertations and theses studying blended learning. *The Internet and Higher Education, 17*, 90-100. https://doi.org/10.1016/j.iheduc.2012.11.003

Rolim, V., Ferreira, R., Gašević, D., Lins, R. (2019). A Network Analytic Approach to Uncovering the Relationship between Social and Cognitive Presences in Communities of Inquiry. *The Internet and Higher Education, 42*, 53-65. https://doi.org/10.1016/j.iheduc.2019.05.001

Fransen, J., Kirschner, P. A., & Erkens, G. (2011). Mediating team effectiveness in the context of collaborative learning: The importance of team and task awareness. *Computers in Human Behavior, 27*(3), 1103-1113. https://doi.org/10.1016/j.chb.2010.05.017

Fransen, J., Weinberger, A., & Kirschner, P. A. (2013). Team effectiveness and team development in CSCL. *Educational Psychologist, 48*(1), 9-24. https://doi.org/10.1080/00461520.2012.747947

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Gašević, D., Joksimović, S., Eagan, B. R., & Shaffer, D. W. (2019). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior*, *92*, 562–577. https://doi.org/10.1016/j.chb.2018.07.003

Haataja, E., Malmberg, J., & Järvelä, S. (2018). Monitoring in collaborative learning: Co-occurrence of observed behavior and physiological synchrony explored. *Computers in Human Behavior, 87,* 337-347. https://doi.org/10.1016/j.chb.2018.06.007

Hadwin, A. F., Järvelä, S., & Miller, M. (2017). Self-regulation, co-regulation and shared regulation in collaborative learning environments. In D. Schunk, & J. Greene, (Eds.). *Handbook of Self-Regulation of Learning and Performance* (2nd Ed, pp. 83-106). New York, NY: Routledge.

Hadwin, A. F., Oshige, M., Gress, C., & Winne, P. (2010). Innovative ways for using gStudy to orchestrate and research social aspects of self-regulated learning. *Computers in Human Behavior, 26*, 794–805. DOI:10.1016/j.chb.2007.06.007

Harley, J. M., Bouchet, F., Hussain, M. S., Azevedo, R., & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior, 48*, 615–625. http://doi.org/10.1016/j.chb.2015.02.013

Harley, J.M., Taub, M., Azevedo, R. & Bouchet, F. (2017). Let's set up some subgoals: Understanding human-pedagogical agent collaborations and their implications for learning and prompt and feedback compliance. *IEEE Transactions on Learning Technologies, 11*(1), 54-66. DOI: 10.1109/TLT.2017.2756629

Isohätälä, J., Näykki, P., Järvelä, S. & Baker, M.J. (2017). Striking a balance: socio-emotional processes during argumentation in collaborative learning interaction. *Learning, Culture and Social Interaction*, *16*(3), 1-19. DOI: 10.1016/j.lcsi.2017.09.003

Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2012). Task-related and social regulation during online collaborative learning. *Metacognition and Learning, 7*(1), 25-43. https://doi.org/10.1007/s11409-010-9061-5

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education, 33*, 74–85. http://dx.doi.org/10.1016/j.iheduc.2017.02.001

Järvelä, S. & Bannert. M. (2019, in press). Introduction to a special issue: Temporal and adaptive process of regulated learning - what can multimodal data tell? *Learning and Instruction.*

Järvelä, S. & Hadwin, A. (2013). New Frontiers: Regulating learning in CSCL. *Educational Psychologist, 48*, 25–39. DOI:10.1080/00461520.2012.74800

Järvelä, S., Hadwin, A.F. Malmberg, J. & Miller. M. (2018). Contemporary Perspectives of Regulated Learning in Collaboration. In: F. Fischer, C.E. Hmelo-Silver, Reimann, P. & S. R. Goldman (Eds.). *International handbook of the learning sciences* (pp. 127-136). New York, NY: Routledge. Taylor & Francis.

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

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

Järvenoja, H., Volet, S., & Järvelä, S. (2013). Regulation of emotions in socially challenging learning situations: An instrument to measure the adaptive and social nature of the regulation process. *Educational Psychology*, *33*(1), 31-58. https://doi.org/10.1080/01443410.2012.742334

Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly,A., Shekhar, S., Samatova, N. & Kumar, V. (2017). Theory-Guided Data Science: A New Paradigm for

Scientific Discovery from Data. *IEEE Transactions on Knowledge and Data Engineering, 29* (10), 2318-2331.

Kirschner, F., Paas, F., Kirschner, P. A., & Janssen, J. (2011). Differential effects of problem-solving demands on individual and collaborative learning outcomes. *Learning and Instruction, 21*(4), 587-599. https://doi.org/10.1016/j.learninstruc.2011.01.001

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

Kraut, R. E., Fussell, S. R., Brennan, S. E., & Siegel, J. (2002). Understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work. In P. Hinds & S. Kiesler (Eds.), *Distributed work* (pp. 137-162). Cambridge, MA, US: MIT Press.

Kreijns, K., Kirschner, P.A., Jochems, W. & van Buuren, H. (2011). *Education and Information Technologies, 16*, 4, 365-381. https://doi.org/10.1007/s10639-010-9135-7

Kovanović, V., Joksimović, S., Waters, Z., Gašević, D., Kitto, K., Hatala, M., & Siemens, G. (2016). Towards automated content analysis of discussion transcripts: A cognitive presence case. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 15–24). ACM.

Leite, I., Henriques, R., Martinho, C., & Paiva, A. (2013). Sensors in the wild: Exploring electrodermal activity in child-robot interaction. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction* (pp. 41-48). IEEE Press Piscataway, NJ, USA.

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology, 50 (6)*, 2824-2838. doi:10.1111/bjet.12861

Malmberg, J., Järvelä, S., Holappa, J., Haataja, E., Huang, X., & Siipo, A. (2018). Going beyond what is visible: What multichannel data can reveal about interaction in the context of collaborative learning? *Computers in Human Behavior, 96,* 235-245. https://doi.org/10.1016/j.chb.2018.06.030

Miyake, N., & Kirschner, P. A. (2014). The social and interactive dimensions of collaborative learning. In K. R. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 418–438). New York: Cambridge University Press.

Neto, V., Rolim, V., Ferreira, R., Kovanović, V., Gašević, D., Dueire Lins, R., & Lins, R. (2018). Automated Analysis of Cognitive Presence in Online Discussions Written in Portuguese. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, & M. Scheffel (Eds.), *Proceedings of the 13th European Conference on Technology Enhanced Learning* (pp. 245–261). Springer International Publishing.

Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2018). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior*. (In Press). DOI: 10.1016/j.chb.2018.12.019

O'Donnell, A. M., & Hmelo-Silver, C. E. (2013). Introduction What is Collaborative Learning?: An Overview. In: Hmelo-Silver, C. E., Chinn, C. A., Chan, C. & O'Donnell, A. M. (Eds.). *The international handbook of collaborative learning* (pp. 13-28). New York, NY: Routledge.

Ochoa, X. (2017). Multimodal learning analytics. In: Lang, C., Siemens, G., Wise, A., Dragan, G. (Eds). *Handbook of learning analytics* (pp. 129-141). SOLAR, Society for Learning Analytics and Research.

Ochoa, X., Worsley, M., Weibel, N., & Oviatt, S. (2016). Multimodal learning analytics data challenges. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 498-499). New York, NY: ACM. DOI: 10.1145/2883851.2883913

Pantic, M., & Rothkrantz, L. J. (2003). Toward an affect-sensitive multimodal human-computer interaction. In *Proceedings of the IEEE*, *91*(9), 1370-1390. DOI: 10.1109/JPROC.2003.817122

Peek, N., Combi, C., Marin, R., & Bellazzi, R. (2015). Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes. *Artificial Intelligence in Medicine*, *65*(1), 61-73. https://doi.org/10.1016/j.artmed.2015.07.003

Pfister, T., Li, X., Zhao, G., & Pietikainen, M. (2011). Recognising spontaneous facial micro-expressions. In P*roceedings of the 2011 International Conference on Computer Vision* (pp. 1449-1456). IEEE Computer Society. DOI: 10.1109/ICCV.2011.6126401

Picard, R. W. (1997). *Affective computing*. Cambridge, MA, US: The MIT Press.

Pijeira-Díaz, H. J., Drachsler, H., Järvelä, S. & Kirschner, P. A. (2018). Sympathetic arousal commonalities and arousal contagion during collaborative learning: How attuned are triad members? *Computers in Human Behavior 92*, 188-197. https://doi.org/10.1016/j.chb.2018.11.008

Rittel, H. W. J. & Webber, M. M. (1973). *Policy Sciences, 4*, 155-169. https://doi.org/10.1007/BF01405730

Rogat, T. K., & Adams-Wiggins, K. R. (2015). Interrelation between regulatory and socioemotional processes within collaborative groups characterized by facilitative and directive other-regulation. *Computers in Human Behavior, 52*, 589–600. https://doi.org/https://doi.org/10.1016/j.chb.2015.01.026

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In: O'Malley, C. (Eds). *Computer-Supported Collaborative Learning* (Vol 128, pp. 69–97). Berlin, Heidelberg: Springer. http://doi.org/10.1145/130893.952914

Schunk, D. H. & Greene, J. A. (2017). Historical, contemporary, and future perspectives on self-regulated learning and performance. In: D. Schunk, & J. Greene, (Eds.). *Handbook of Self-Regulation of Learning and Performance* (2nd ed., pp. 1-15). London: Routledge.

Shaffer, D. W. (2006). Epistemic frames for epistemic games. *Computers & Education, 46*(3), 223–234. https://doi.org/10.1016/j.compedu.2005.11.003

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A Tutorial on Epistemic Network Analysis: Analyzing the Structure of Connections in Cognitive, Social, and Interaction Data. *Journal of Learning Analytics, 3*(3), 9–45. https://doi.org/10.18608/jla.2016.33.3

Väyrynen, E., Kortelainen, J., & Seppänen, T. (2013). Classifier-based learning of nonlinear feature manifold for visualization of emotional speech prosody. *IEEE Transactions on Affective Computing, 4*(1), 47–56. https://doi.org/10.1109/T-AFFC.2012.35

Wallot, S., Roepstorff, A., & Mønster, D. (2016). Multidimensional Recurrence Quantification Analysis (MdRQA) for the analysis of multidimensional time-series: A software implementation in MATLAB and its application to group-level data in joint action. *Frontiers in Psychology, 7*, 1835. http://doi.org/10.3389/fpsyg.2016.01835.

Winne, P. H. (2017). Cognition and metacognition within self-regulated learning. In D. Schunk, & J. Greene, (Eds.). *Handbook of self-regulation of learning and performance* (pp. 52-64). London: Routledge.

Winne, P. H., & Baker, R. (2013). The potentials of educational data mining for researching metacognition, motivation and self-regulated learning. *JEDM| Journal of Educational Data Mining*, *5*(1), 1–8.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In: D. Hacker, J. Zimmerman, B. (2000). Attaining self-regulation: A social constructive perspective: In: M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Self-regulation: Theory, research, and applications* (pp. 13–39). Orlando: Academic Press.

Wise, A., & Shaffer, D. W. (2015). Why Theory Matters More than Ever in the Age of Big Data. *Journal of Learning Analytics*, *2*(2), 5–13. https://doi.org/10.18608/jla.2015.22.2

Yang, D., Wen, M., Kumar, A., Xing, E. P., & Rose, C. P. (2014). Towards an integration of text and graph clustering methods as a lens for studying social interaction in MOOCs. *The International Review of Research in Open and Distributed Learning*, *15*(5). https://doi.org/10.19173/irrodl.v15i5.1853

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2011). *Handbook of self-regulation of learning and performance.* New York, NY: Routledge.