

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/350777881>

What Do You Mean by Collaboration Analytics? A Conceptual Model

Article · April 2021

DOI: 10.18668/jla.2021.7227

CITATIONS

16

READS

691

6 authors, including:



Roberto Martinez-Maldonado

Monash University (Australia)

153 PUBLICATIONS 2,564 CITATIONS

[SEE PROFILE](#)



Dragan Gasevic

The University of Edinburgh

653 PUBLICATIONS 16,350 CITATIONS

[SEE PROFILE](#)



Vanessa Echeverria

Monash University (Australia)

43 PUBLICATIONS 443 CITATIONS

[SEE PROFILE](#)



Simon Buckingham Shum

University of Technology Sydney

335 PUBLICATIONS 8,629 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Temporal Process Mining in Learning Analytics [View project](#)



UTS ECR Grant 2017: High Performance Teamwork Analytics in Physical Spaces [View project](#)

What Do You Mean by Collaboration Analytics? A Conceptual Model

Roberto Martinez-Maldonado¹, Dragan Gašević², Vanessa Echeverria³, Gloria Fernandez Nieto⁴, Zachari Swiecki⁵, Simon Buckingham Shum⁶

Abstract

Using data to generate a deeper understanding of collaborative learning is not new, but automatically analyzing log data has enabled new means of identifying key indicators of effective collaboration and teamwork that can be used to predict outcomes and personalize feedback. Collaboration analytics is emerging as a new term to refer to computational methods for identifying salient aspects of collaboration from multiple group data sources for learners, educators, or other stakeholders to gain and act upon insights. Yet, it remains unclear how collaboration analytics go beyond previous work focused on modelling group interactions for the purpose of adapting instruction. This paper provides a conceptual model of collaboration analytics to help researchers and designers identify the opportunities enabled by such innovations to advance knowledge in, and provide enhanced support for, collaborative learning and teamwork. We argue that mapping from low-level data to higher-order constructs that are educationally meaningful, and that can be understood by educators and learners, is essential to assessing the validity of collaboration analytics. Through four cases, the paper illustrates the critical role of theory, task design, and human factors in the design of interfaces that inform actionable insights for improving collaboration and group learning.

Notes for Practice

- Collaboration analytics partially or fully automate the analysis of data captured from group/team settings. These include co-located, face-to-face groups and platform-mediated groups (synchronous and asynchronous).
- The results of such analytics should help the team, educators, researchers, or other stakeholders gain insight into the activity and inform action that improves collaboration processes and products.
- Collaboration analytics require the integration of three key inputs: (i) theoretically sound concepts, (ii) well-designed tasks that generate useful data traces to inform analytics, and (iii) human-centred design that empowers stakeholders to shape the analytics and produces effective user interfaces.
- Four cases demonstrate that when these elements are aligned, data traces from activity sensors can be designed in principled ways to serve as evidence of learners' capabilities, rendered via user interfaces that make sense to users.
- Collaboration analytics of this sort, which go beyond simply counting and graphing low-level events, are predominantly research prototypes and rare to find in products.

Keywords

Collaborative learning, teamwork, CSCL, CSCW, multimodal learning analytics, group work, learning design

Submitted: 29/05/20 — **Accepted:** 22/01/21 — **Published:** 09/04/21

Corresponding author ¹Email: roberto.martinezmaldonado@monash.edu Address: Faculty of Information Technology, Monash University, Melbourne, VIC3800, Australia. ORCID ID: <https://orcid.org/0000-0002-8375-1816>

²Email: dragan.gasevic@monash.edu Address: Faculty of Information Technology, Monash University, Melbourne, VIC3800, Australia. ORCID ID: <https://orcid.org/0000-0001-9265-1908>

³Email: vanechev@espol.edu.ec Address: Faculty of Electrical and Computer Engineering, Escuela Superior Politécnica del Litoral (ESPOL), Guayaquil, 09-01-5863, Ecuador. ORCID ID: <https://orcid.org/0000-0002-2022-9588>

⁴Email: gloria.m.fernandeznieto@student.uts.edu.au Address: Connected Intelligence Centre, University of Technology Sydney, NSW 2008, Australia. ORCID ID: <https://orcid.org/0000-0002-8163-2303>

⁵Email: zach.swiecki@monash.edu Address: Faculty of Information Technology, Monash University, Melbourne, VIC3800, Australia. ORCID ID: <https://orcid.org/0000-0002-7414-5507>

⁶Email: simon.buckinghamshum@uts.edu.au Address: Connected Intelligence Centre, University of Technology Sydney, NSW 2008, Australia. ORCID ID: <https://orcid.org/0000-0002-6334-7429>

1. Introduction

When learners work in small groups, some of their hidden mental processes are externalized to others through drawings, verbal communication, writing, and other forms of meaning representation (Stahl, 2006). The artefacts generated by the group and their dialogue enable group members to establish common ground, explain their personal views, reach consensus, or produce new group knowledge (Scardamalia & Bereiter, 2006; Stahl, 2004). For this reason, it has been proposed that small-group activity provides particular opportunities for *data-intensive learning analytics* innovations when traces of the meaning-making processes externalized by group members can be digitally captured (Dönmez, Rosé, Stegmann, Weinberger, & Fischer, 2005). These data traces can be analyzed and used to find recurrent behavioural patterns, make collaborative interactions visible, and broaden our understanding of collaborative learning in various domains (Wise & Schwarz, 2017). This can further accelerate computer-supported collaborative learning (CSCL) researchers' analysis cycles and developers' design iterations (Puntambekar, Erkens, & Hmelo-Silver, 2011).

Using data to generate a deeper understanding of CSCL activity goes back to before the establishment of learning analytics as a research field (Siemens & Long, 2011). For example, log data has been used to understand how groups work differently in synchronous and asynchronous settings (Reimann, Yacef, & Kay, 2011) and to characterize effective collaboration (e.g., Perera, Kay, Koprinska, Yacef, & Zaïane, 2008), argumentation (e.g., Rosé et al., 2008), and teamwork (e.g., Kay, Maisonneuve, Yacef, & Zaïane, 2006). Modelling interactions has also enabled the creation of mechanisms to adapt the support provided to groups (Kumar, Rosé, Wang, Joshi, & Robinson, 2007) and automatically adjust the collaboration script according to particular group needs (Rummel et al., 2008). Within other data-intensive educational technology communities, such as intelligent tutoring systems (ITS), educational data mining (EDM), and artificial intelligence in education (AIED), there has also been a sustained interest in providing intelligent support to learning in groups (ISLG). Researchers across these communities have organized thematic series of workshops (e.g., the ISLG workshop series organized in ITS and AIED conferences between 2012 and 2018; Kim & Kumar, 2012) and special issues in high-impact journals (e.g., Isotani, 2011; Kumar & Kim, 2014). In fact, the idea of mirroring group data traces back to students, which is a distinctive characteristic of learning analytics research (Clow, 2012), has also been explored in the CSCL (e.g., Jermann & Dillenbourg, 2008; Soller, Martinez, Jermann, & Muehlenbrock, 2005) and AIED (e.g., Bull & Vatrupu, 2011) literature. Moreover, data-intensive developments related to group activity are of interest in non-educational fields. For instance, at the Computer-Supported Cooperative Work (CSCW) 2002 conference, a wearable sensor (the sociometric badge) was introduced to collect real-time speech and motion data during face-to-face interactions to infer group dynamics (Choudhury & Pentland, 2002), while others have analyzed the use of collaborative tools using data-centric approaches (Wang, Jin, & Liu, 2010).

Nowadays, learning analytics research focused on supporting collaboration and group learning has diversified, and perhaps even fragmented, across various research communities, such as CSCL (e.g., Spikol, Ruffaldi, & Cukurova, 2017a), technology-enhanced learning (e.g., Praharaj, Scheffel, Drachler, & Specht, 2018), team science (e.g., Kim, Sottolare, Brawner, & Flowers, 2018), and human-computer interaction (e.g., Chandrasegaran, Bryan, Shidara, Chuang, & Ma, 2019). Novel analytic approaches have emerged and are applied to study collaboration with a wide variety of purposes, such as assessment and measurement of collaborative learning (Khan, 2017), theory building (Malmberg et al., 2018), orchestration support (Olsen, 2017), dashboards design (van Leeuwen, Rummel, & van Gog, 2019), and user modelling (Worsley, 2019). The emergence of accurate and inexpensive sensors is also enabling the exploration of multimodal aspects of group interaction, such as gaze tracking for measuring joint attention of students (Schneider & Pea, 2017), physiological sensing for identifying group synchrony (Haataja, Malmberg, & Järvelä, 2018; Schneider, Dich, & Radu, 2020), and gesture/motion tracking for inferring collaboration strategies in maker-spaces (Worsley & Blikstein, 2018). In short, although using log data to analyze and support collaboration is not new, and the research is fragmented across various communities, there is an emerging interest in considering the complexity of collaboration beyond the analysis of clicks and keystrokes on an online platform.

A new related term, *collaboration analytics*, has recently emerged in educational contexts (e.g., Anaya, González-Boticario, Letón, & Hernández-del-Olmo, 2015; Kasepalu, Santos, & Ley, 2019; Martinez-Maldonado, Kay, & Buckingham Shum, 2019b). Anaya and colleagues (2015) appear to be the first to use the term to our knowledge, but in a very limited sense, referring to the use of social network analysis outputs to alert teachers of students facing challenges in a MOOC. Martinez-Maldonado and colleagues (2019b) used the term in the context of co-located collaboration as a way to identify the principles and dilemmas commonly encountered when mining data from group settings. A recent workshop at the CSCL conference (Martinez-Maldonado, Worsley, Schneider, & Kharrufa, 2019c) attracted contemporary research work focused on the need for effective tools that provide automated assessment and meaningful feedback on group activity. However, two critical questions arise in the context of this literature: *what is meant by collaboration analytics?* and *how can collaboration analytics innovations go beyond previous work focused on modelling group interactions?*

This paper argues that a characterization of collaboration analytics for education is timely and distinctive. We propose a model of collaboration analytics to help researchers and designers identify opportunities enabled by such innovations to

advance knowledge in, and provide enhanced support for, collaborative learning and teamwork. Knight and Buckingham Shum (2017) argue that a distinctive feature of learning analytics is that it closes the feedback loop through automated analysis and feedback:

“The potential of learning analytics is arguably far more significant than as an enabler of data-intensive educational research, exciting as this is. The new possibility is that educators and learners — the stakeholders who constitute the learning system studied for so long by researchers — are for the first time able to see their own processes and progress rendered in ways that until now were the preserve of researchers outside the system. [...] So, for educators and learners, the interest turns on the ability to gain insight in a timely manner that could improve outcomes.”

We argue, therefore, that *educational collaboration analytics* should refer specifically to methods for *automated data capture* from group settings (often considering multiple data sources), and their *automated analysis*, with the purpose of informing *actionable insights* for improving collaboration and group learning. These insights need to be effectively communicated to learners, groups of learners, educators, or other stakeholders for them to make decisions according to their values and factors unaccounted for in the analysis (Jørnø & Gynther, 2018). The challenge is that merely visualizing log data does not ensure that students and teachers will be able to interpret the data meaningfully (Ahn, Campos, Hays, & DiGiacomo, 2019; Corrin & De Barba, 2015; Gašević, Dawson, & Siemens, 2015b; Greller & Drachler, 2012; Lim, Dawson, Joksimovic, & Gašević, 2019) or discern the implications for action (Clow, 2014; Colvin et al., 2016; Tempelaar, Rienties, & Nguyen, 2017; van Leeuwen, 2019; Wise, 2014). This points to the need for (i) modelling from low-level group data to higher-order group constructs that are educationally meaningful and (ii) effectively representing such constructs in terms that can be understood by people without formal data science training.

2. Background

2.1. Learning Analytics and Collaborative Learning

Making evidence of team activity available for inspection and assessment has been proposed by the CSCL community as an effective way to encourage team members to reflect on and learn from their own experiences (Knipfer, Prilla, Cress, and Hermann, 2011). As a result, there has been a growing interest in embracing learning analytics methodologies and innovations in CSCL (Ludvigsen, Cress, Law, Stahl, and Rosé, 2017; Wise et al., 2015; Wise, Knight, & Buckingham Shum, 2021; Wise & Schwarz, 2017). For example, van Leeuwen, Janssen, Erkens, and Brekelmans have investigated the effects of various CSCL visualization designs on teachers' ability to diagnose students' participation within groups (2014) and to monitor and regulate multiple groups of students (2015). The same authors mapped the kinds of analytics support currently provided by CSCL tools, namely mirroring students' data, alerting teachers if some behaviours are automatically detected, or more proactively advising what actions teachers can take (van Leeuwen et al., 2019). Rodríguez Triana, Martínez-Monés, Asensio-Pérez, and Dimitriadis (2014) also investigated how to connect the pedagogical decisions made explicit in the learning design or the CSCL script with analytics extracted from group data. More recently, the applications of learning analytics in CSCL have diversified in the form of CSCL dashboards (see review by Liu & Nesbit, 2019), visualizations (Steier et al., 2019), and analysis tools (Rosé, 2018); cognitive assistants to support teachers monitoring CSCL tasks (Bae et al., 2019; Rodríguez-Triana, Prieto, Martínez-Monés, Asensio-Pérez, & Dimitriadis, 2018a); methodologies to co-design CSCL analytics with teachers and students (e.g., Martínez-Maldonado, Elliot, Axisa, Power, & Buckingham Shum, 2019a); and applications that automatically capture physical (e.g., Vujovic, Tassani, & Hernández-Leo, 2019) and physiological (e.g., Noroozi et al., 2019) aspects of learning to understand group regulation processes.

Similarly, numerous emerging analytics innovations are applied to collaborative (Rosé, 2017) and networked (Suthers, 2017) online learning contexts. For example, De Liddo, Buckingham Shum, Qunito, Bachler, and Cannavacciuolo (2011) proposed an approach for analyzing rhetorical moves in student discourse and visualizing these as social networks. Harrer (2013) proposed a rule-based system to guide students planning to work in small groups online. Xing, Wadholm, and Goggins (2014) proposed a clustering technique to assess participation of students across various CSCL tools, such as a chat, a virtual whiteboard, and a geometry learning system. Tan, Yang, Koh, and Jonathan (2016) investigated the effects of a student-facing dashboard in fostering collaborative critical reading skills. Swiecki, Ruis, Farrell, and Shaffer (2020) proposed the use of epistemic network analysis to automatically assess individual contributions in group problem-solving tasks. Gašević, Joksimović, Egan, and Shaffer (2018) proposed a network analytics approach to identifying low- and high-performing groups in a MOOC by combining properties of social and epistemic network analysis.

Moving beyond computer-mediated collaboration, Martínez-Maldonado and colleagues (2017) presented various analytics innovations applied to face-to-face group settings in which interaction was mediated by large shared displays and data was captured via depth cameras, eye-trackers, and audio recording devices. In fact, there is a growing interest in capturing traces of group multimodal interaction from physical learning spaces by combining sensor signals such as eye-trackers (Schneider, 2020; Vrzakova, Amon, Stewart, Duran, & D'Mello, 2020), computer vision systems (Spikol, Ruffaldi, Landolfi, & Cukurova,

2017b), indoor-positioning (Echeverria, Martinez-Moldonado, & Buckingham Shum, 2019), and electrodermal (Pijera-Díaz, Drachsler, Järvelä, & Kirschner, 2016) and electroencephalographic (Stevens, Galloway, & Willemsen-Dunlap, 2020) sensors. In sum, the analysis of collaboration in both the CSCL and the learning analytics communities is widening, and multiple data collection, analysis, and interface design approaches are being explored.

Beyond formal educational settings, Buckingham Shum and Ferguson (2012) argued for the need to engage social settings beyond those regularly targeted in CSCL research by considering non-academic contexts in which freely available, off-the-shelf online tools are commonly used, and neatly packed curriculum and learning designs are not common (i.e., home, social networking software, and the workplace). This points to work conducted in other areas aiming to support teamwork. For example, Kim and colleagues (2018) proposed a mapping between measures that can be derived from sensors and the learner states to create intelligent team training systems. This mapping includes the use of electroencephalography (EEG) headsets to identify team members' levels of engagement, anxiety, or excitement and the use of wristbands and face tracking to identify emotional and motivational traits. Some researchers are bringing these theoretical mappings into authentic teamwork settings, such as medical (Echeverria, Martinez-Moldonado, Power, Hayes, & Buckingham Shum, 2018c) and firefighting (Wake, Wasson, Bjørgen, & Heimsæter, 2019) training.

In short, the body of work presented is not aimed at providing an exhaustive review of works related to collaboration analytics. Instead, it is intended to show the variety of classic and emerging methods from other fields (i.e., CSCW, team science, AIED, ITS, and EDM) that the learning analytics community can use to provide data-informed support in teamwork training and collaborative learning situations. For example, this has been one of the key arguments to connect CSCL and learning analytics in recent vision papers (e.g., Ludvigsen et al., 2017; Wise & Schwarz, 2017), yet some concerns have been raised. Wise and Schwarz (2017) warned the community about the risk of creating sophisticated group models based on low-level data (e.g., clickstreams and eye-gaze saccades) that cannot be translated into forms comprehensible to teachers and learners. To avoid dystopian scenarios that strip agency from learners and teachers, Tchounikine (2019) argued that the role of analytics in CSCL should be limited to enhancing awareness and recommending potential actions. These concerns derive from the risk of making strong decisions about students' learning (by machine learning — ML — algorithms or humans) based on the vast quantities of low-level group data now available: data that is commonly incomplete (Kitto, Buckingham Shum, & Gibson, 2018), easily stripped from its context (Buckingham Shum, Echeverria, & Martinez-Maldonado, 2019a), and often disconnected from the learning design (Lockyer, Heathcote, & Dawson, 2013; Rodríguez Triana et al., 2014). Moreover, collaboration analytics user interfaces can be hard to interpret by educators and learners who commonly lack formal data science training (Maltese, Harsh, and Svetina, 2015; Mandinach & Gummer, 2013).

In response, there is a growing body of work focused on how to forge principled relationships between the kinds of data logged by software platforms; educational constructs; and, hence, the terminology needed to produce intelligible feedback that educators and learners can understand.

2.2. Mapping from Low-Level Group Data to Educational Constructs

The shorthand term *from clicks to constructs* was coined by Buckingham Shum and Crick (2016, p. 16) to capture the fundamental challenge of the level of description of events logged by a computational sensor being very different from how people conceptualize human activity. This reflects a growing body of work investigating how one maps from *clicks and keystrokes* in collaborative platforms, or *multimodal datastreams* in face-to-face collaboration, to constructs. This is important for researchers to make claims about learning processes (Cronbach & Meehl, 1955) but, equally, for stakeholders without a strong data analytic background who want to understand teaching and learning. Wise and colleagues (2021) provide a helpful diagram (adapted in Figure 1) to visualize this, which has an extensive precedent in the psychometric literature (Milligan, 2020; Stadler, Herborn, Mustafić, & Greiff, 2020). They provide examples of CSCL projects in which indicators extracted from forum postings provided evidence about learners' *cognitive presence* (Joksimovic, Gasevic, Kovanovic, & Hatala, 2014) and *uptake of ideas* between learners in online discourse (Suthers, Dwyer, Medina, & Vatrappu, 2010) and detected *promising ideas* in knowledge-building communities (Lee & Tan, 2017). Similarly, Shute and Ventura (2013) argue that the higher-order construct of student *conscientiousness* can be assessed based on *unobservable (sub)constructs* (such as learners' *persistence, perfectionism, organization, and carefulness*), which, in turn, can be modelled in terms of *observable indicators* extracted from system logs (such as number of attempts, time to solve a problem, and complexity of a solution) in the context of educational games.

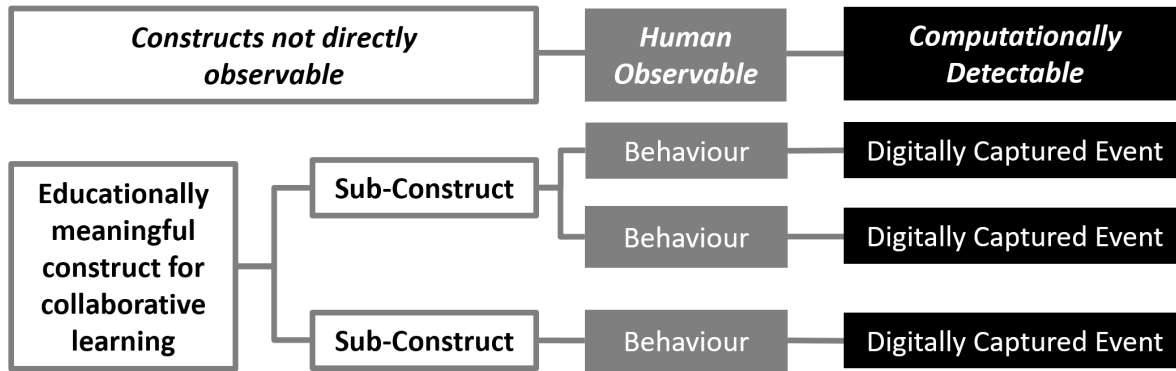


Figure 1. Mapping “from clicks to constructs” in CSCL (adapted from Wise et al., 2021).

The examples presented above are derived from contexts where a single modality of a clickstream is associated with higher-order constructs. In contrast, co-located collaboration scenarios can generate multiple streams of data from individuals, making the task of giving meaning to low-level logs even more challenging. Recently, inroads have been made into mapping from *observable* multimodal sensor data to *unobservable* educational constructs. For example, Echeverria and colleagues (2019) proposed and implemented a modelling approach and data structure called the multimodal matrix to map from indoor positioning, physiological, audio, and system logging data to a holistic view of collaboration that includes *dimensions of collaboration* (epistemic, physical, social, and affective) as per a theoretical framework of collaboration (Martinez-Maldonado, Goodyear, Kay, Thompson, & Carvalho, 2016). A similar mapping was conceptually proposed by Worsley and colleagues (2016) to navigate from low-level sensor data (e.g., extracted from eye-trackers, video/depth cameras, and microphones) to *learning constructs* (e.g., conceptual change, emotional intelligence, mindset, and identity) using a range of *analytic techniques* such as classifiers, prosodic analysis, hidden Markov models, and constrained local neural fields.

A critical remaining question is: *how can we generate mappings from low-level data to educational group constructs in ways that are meaningful and valid?* Several authors in the learning analytics community have emphasized the critical role of educational theory in both driving applied research and facilitating interpretation of analytic outputs (Clow, 2012; Dawson, Mirriahi, & Gasevic, 2015; Gašević, Dawson, Rogers, & Gasevic, 2016; Gašević, Kovanović, & Joksimović, 2017; Knight & Buckingham Shum, 2017; Wise & Shaffer, 2015). In response to claims that the “big data revolution” signals the “death of theory” (Mazzocchi, 2015), Wise and Shaffer (2015) argue that when datasets are so large that spurious statistically significant patterns can be obtained easily, educational theory and expert knowledge are even more important to guide interpretation. Knight and Buckingham Shum (2017) propose that educational theory can provide a principled way to go from clicks to constructs by pointing at what aspects of learning are important to measure for particular stakeholders, and how to reliably generate indicators and analytics from data. Gašević and colleagues (2017) explicitly propose that the learning analytics innovations that are more effective and have the highest validity are those created by considering critical principles of educational theory, design, and data science. Analytics can thus be designed in a principled manner for higher-order group constructs, such as “effective collaboration” (Echeverria et al., 2019), “uptake of ideas from others” (Suthers et al., 2010), and “crowd-sourced capacity to learn” (Milligan & Griffin, 2016), by giving meaning to low-level data based on educational theory (Gašević et al., 2017; Wise & Shaffer, 2015), expert knowledge (Buckingham Shum et al., 2019a), or human-centred methodologies (Buckingham Shum, Ferguson, & Martinez-Maldonado, 2019b).

In tackling this challenge, an important foundation that learning analytics can build on is measurement science, which long predates the data/analytics revolution. The challenge of designing rigorous assessments has long required researchers to grapple with the challenge of assessing a construct or capability, for example, from student responses to test items. Researchers from those fields have developed methodologies that adapt the knowledge in this field for contemporary learning platforms and, by extension, collaborative learning. Milligan (2020) has coined the term *metrolytics* as a hybrid of metrics and analytics to encapsulate such approaches. Evidence-centred design (ECD) exemplifies such an approach, requiring the careful design of tasks to elicit student behaviours that can be detected digitally, from which data patterns contribute differential amounts of evidence regarding students’ capabilities (Behrens, DiCerbo, & Foltz, 2019; Mislevy, 2019; Rupp, Nugent, & Nelson, 2012). Milligan and Griffin (2016) document their mapping from clicks to constructs in a MOOC, operationalizing the higher-order construct “capability to learn in a scaled environment” in terms of four student capabilities: *recursiveness*, *risk taking*, *peer evaluation*, and *self-monitoring*. These were then mapped to *behavioural indicators*, such as peer responsiveness to conversation threads, reputation points, patterns of quiz practice, and patterns of quiz resubmission or video replay. These

indicators were obtained by applying rule-based algorithms to assess the log data based on thresholds (e.g., accessing documents > 20 times or resubmitting videos or quizzes < 50%).

To summarize, in order to provide intelligent interfaces that support collaboration and teamwork, designers of collaboration analytics require a methodology for imbuing low-level data with contextual meaning and interfaces that communicate insights at a higher-order level. These conceptual foundations serve to propose a conceptual model for designing collaboration analytics, which is presented in the next section.

3. What Do You Mean by Collaboration Analytics?

A preliminary definition of collaboration analytics was stated as “the particular techniques and approaches needed to automatically (or semi-automatically) capture, analyze, mine and distill data about collaborators’ interactions” (Martinez-Maldonado et al., 2019b). However, this definition does not articulate the phenomenon at the intersection of *collaboration* and *analytics*. In fact, the meaning of the word *collaboration*, and more specifically *collaborative learning*, has evolved over the years across different communities of research. In CSCL, an extremely broad definition of collaborative learning corresponds to those situations in which two or more people learn or attempt to learn together (Dillenbourg, 1998). This means that small groups working on a short classroom task can be considered as collaborative situations, as well as whole communities of practice working asynchronously for long periods of time. In practice, recent meta-analyses of the CSCL literature (Chen, Wang, Kirschner, & Tsai, 2018; Jeong, Hmelo-Silver, & Jo, 2019) and the foundational CSCL literature (Dillenbourg, 1999; Jeong & Hmelo-Silver, 2010; Stahl, 2017) indicate that CSCL research has mostly constrained its interest to small-group activity, with notable exceptions (e.g., Aviv, Erlich, & Ravid, 2003; Szwed et al., 2011). Other communities, such as networked learning (Goodyear, Banks, Hodgson, & McConnell, 2006) and computer-supported cooperative work (CSCW) (Grudin, 1991), have more explicitly focused on studying aspects related to collaboration at organizational and community levels, or group phenomena closely related to collaborative learning such as highly effective teamwork (Fiore, 2008) and cooperation in the workplace (Fitzpatrick & Ellingsen, 2013).

Although each research community evolves its own identity, epistemology, and definitions, numerous calls have been made to share or unify foundational theories about how people collaborate effectively using technology (Fransen, Weinberger, & Kirschner, 2013; Goggins, Jahnke, & Wulf, 2012; Hernandez-Leo, Villasclaras-Fernandez, Asensio-Perez, Dimitriadis, & Retalis, 2006; Stahl, 2013; Wise & Schwarz, 2017), extending further to fields not specialized in education (e.g., CSCW and Team Science). Researchers and developers within these communities have focused on developing or utilizing, to different extents, various kinds of *groupware*: digital technologies aimed at helping people work together while interacting directly or through a digital system (Ellis, Gibbs, & Rein, 1991). As a result, awareness of these other bodies of literature is critical for designing theory-informed learning analytics in order to support groups.

The word *analytics* has been defined as a body of knowledge or principles for using “data, statistical and quantitative analysis, exploratory and predictive models, and fact-based management to drive decisions and actions” (Davenport & Harris, 2007, p. 7) or, in other words, “transforming raw data into meaningful information to improve decision making” (Wilder & Ozgur, 2015, p. 180). These definitions of *analytics* are different from the data *analysis* commonly conducted by researchers in various areas (including educational research, CSCL, and teamwork science) with the purpose of generating academic knowledge. A key purpose of analytics is to help people make decisions and act upon meaningful forms of processed data. More specifically, the notion of analytics has been appropriated in the learning analytics field as “...the process of developing actionable insights...” (Cooper, 2012, p. 3).

The term *actionable insight* has been increasingly used in several learning analytics research papers (Drachler et al., 2015; Gašević et al., 2017; Pardo et al., 2016; Sergis & Sampson, 2017); however, it was only recently formally defined as “data that allows a corrective procedure, or feedback loop, to be established for a set of actions” (Jørnø & Gynther, 2018). This concept foregrounds the importance of designing for specific actors, with particular tasks and levels of expertise, rather than a vague “user.” This concept also enables us to more explicitly differentiate between collaboration analytics and those previous data-intensive AIED, CSCL, and ITS research works in which the purpose is often for the system to automatically make decisions to provide adapted actions (i.e., as in intelligent tutoring, recommender systems, and conversational agents) (Graesser et al., 2018), or to support researchers to gain new understanding of group communication processes (Dowell, Nixon, & Graesser, 2019). Although some of that work can also be considered within the umbrella of collaboration analytics, here we provide a conceptual model of collaboration analytics being strongly linked to the notion of actionable group insights. Figure 2 depicts a consolidated model in the form of a concept map.

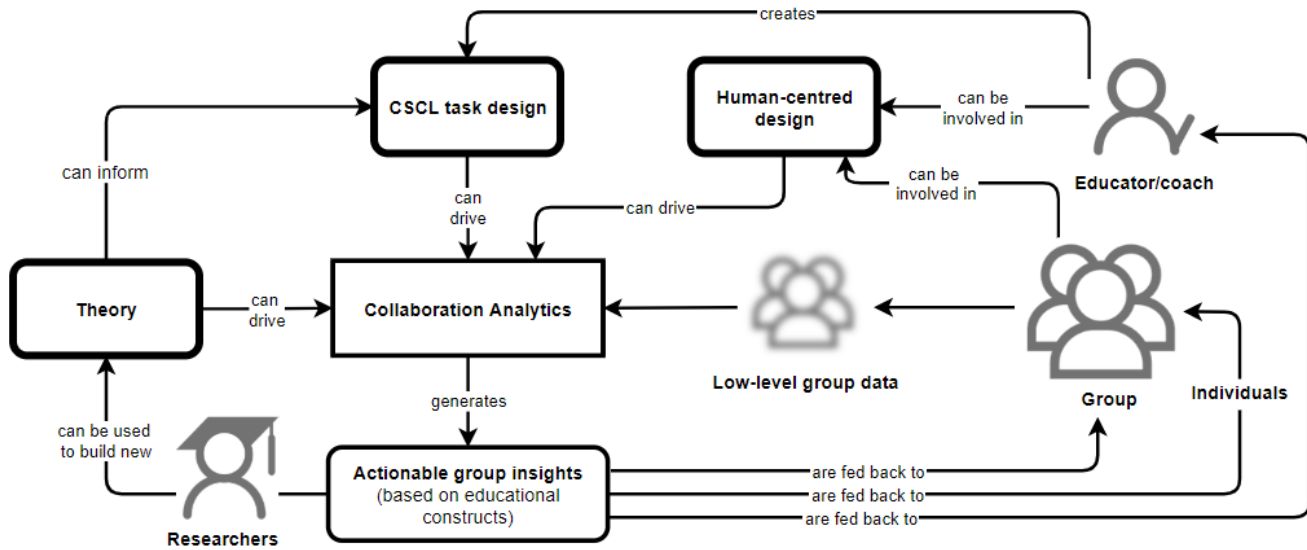


Figure 2. A conceptual model of collaboration analytics.

The model features the notion of *collaboration analytics* as a central node that has, at least, four inputs and one main output. Evidently, the main input corresponds to the set of (typically incomplete, imperfect) *low-level data* traces obtained from a group situation (see the intentionally fuzzy icon). The group activity can be facilitated by an *educator or coach* for the cases in which there is a learning goal associated with such activity. The other three inputs for the collaboration analytics include (a) *theory*, (b) the *CSCL task design*, and (c) outputs from *human-centred design*. Some authors (Gašević et al., 2017; Knight & Buckingham Shum, 2017; Reimann, 2016; Rogers, Gašević, & Dawson, 2016) have argued that *theory* in learning analytics needs to play a pivotal role since purely data-driven approaches have been recognized as not sufficiently informative for educational purposes. CSCL theories (see recent review by Stahl & Hakkarainen, 2021) and team theory (Porter & Beyerlein, 2000; van Schuppen, 2015) are suggested as drivers to guide not only the selection of salient aspects of the group activity that need attention but also the selection of methods for analyzing the data and designing analytics tools.

Similarly, several authors argue for the importance of aligning the *task (or learning) design* with the analytics (Bakharia et al., 2016; Lockyer et al., 2013; Mor, Ferguson, & Wasson, 2015; Wise, 2014). For the case of collaboration analytics, the task design is commonly realized as a collaboration script (Rodríguez Triana et al., 2014), which should ideally be informed by some educational theory (Hernández Leo, Asensio-Pérez, Dimitriadis, & Villasclaras-Fernández, 2010). Scripts commonly give structure to collaborative interactions at a macro level (e.g., by defining a sequence of tasks or distributing roles explicitly) or a micro level (e.g., by establishing protocols of interaction or types of knowledge representation), facilitated by either an educator or a computer system (Kollar, Fischer, & Hesse, 2006). For the case of teamwork in professional settings, similar scripts or task designs can be found in the form of training programs (Fiore, 2008) or professional protocols, norms, and processes (e.g., for conflict resolution, decision making, and emergency response) to be followed during particular team situations (Johnson, Suriya, Yoon, Berrett, & La Fleur, 2002; Pauleen, 2004). Tuckman’s (1965) team development model and Salas, Sims, and Burke’s (2005) “big five” of teamwork are examples of theoretical models that describe what members of highly effective teams commonly do.

Regarding *human-centred design*, interest has recently been growing in the learning analytics field to collaborate with teachers and students to co-design the systems that they will end up adopting into their teaching and learning practices (Buckingham Shum et al., 2019b; Dollinger, Liu, Arthars, & Lodge, 2019). Prieto-Alvarez, Martínez-Maldonado, and Buckingham Shum (2020) argue that by involving multiple educational stakeholders in co-design sessions, it is possible not only to explicitly identify people’s data needs and requirements but also to deeply understand learning and assessment goals and interaction design aspects. While teachers are not usually experts in the co-design process and students most likely do not have pedagogical training, they can bring their expertise to the tasks they commonly need to accomplish (Gomez, Kyza, & Mancevice, 2018). For example, co-design sessions could help capture the details about the task design, pedagogical intentions, and theoretical underpinnings that may not be explicitly stated by educators and instructors in their learning designs or training programs (Echeverria, 2020). In sum, *theory*, *task design*, and *co-designing with educational stakeholders* can each play a key role in helping identify what aspects of collaboration are worth measuring (Kitto et al., 2018), clarify the meaningfulness of data associations discovered through statistical or ML algorithms (Wise & Shaffer, 2015), assess the applicability of findings

according to the particular context and pedagogical intent (Lockyer et al., 2013), and ultimately translate analytics outputs into educational constructs that educators and learners can understand.

The ultimate goal of collaboration analytics is to generate *actionable group insights* based on educational or teamwork constructs, shown as feedback loops. These insights can help *researchers* confirm or formulate new group theory as it has been proposed in both CSCL and teamwork research fields (Kim et al., 2018; Wise & Schwarz, 2017). Yet, a critical distinction discussed above is the opportunity to include groups, individual students, and educators in the feedback loop. This includes the design of effective interfaces for guiding *learners* or provoking reflection to improve the effectiveness of teamwork or learning. In role-based group situations, people could gain a better understanding of competencies that are particularly important to develop in certain roles, such as leadership (Echaluce, Fidalgo-Blanco, Esteban-Escano, Peñalvo, & González, 2018). *Educators and coaches* can also play a critical role by acting upon group insights by performing interventions, providing expanded feedback, or changing the task design or their pedagogical approaches. Other stakeholders (e.g., decision makers, groupware designers, and education providers) can also close the feedback loop by gaining group insights that can enable them to identify best practices or assess designed artefacts such as the collaborative space, groupware, or organizational strategies to improve collaborative learning and teamwork (omitted from Figure 2 for simplification).

While the components of the model (theory, CSCL task design, human-centred design, and actionable group insights) are clearly also relevant to non-collaborative phenomena, collaboration introduces elements beyond what is already considered by learning analytics in general. For example, collaboration may imply additional levels of analysis (individuals, groups, cohorts) and end-users (e.g., the group). Different theories and constructs of interest (e.g., from CSCL and teamwork literature) apply to collaborative work, and also different analytic techniques (e.g., those that account for social and cognitive phenomena, those that are sensitive to interdependence, and multimodal analytics concerned with face-to-face interaction). In fact, Stahl (2006) argued that even those tasks that seem completely individual feature at least very minimal collaboration. Learners need to gain access to others' resources and articulate and negotiate meaning, even if there is no direct dialogue involved. Nonetheless, below we focus on CSCL activities that have a strong social component and emphasize the need to consider literature from communities that study collaborative learning and teamwork from various perspectives.

In sum, we can consider collaboration analytics as the body of knowledge and principles for transforming group data, using descriptive and predictive modelling, into insights that educators, learners, researchers, and other stakeholders can act upon to enable the optimization of collaboration and learning, or the development of new designs and group theory. This definition directly connects to the efforts in mapping from clicks to constructs reviewed in Section 2.2. We use the term *higher-order group construct* to refer to the broad, unobservable, group-related concept to be assessed (e.g., effective teamwork and collaboration). These constructs can then be conceptually defined in terms of *sub-constructs*. For example, the higher-order construct effective teamwork has been characterized in terms of leadership capacity, adaptability, back-up behaviour, mutual monitoring, and team orientation (Salas et al., 2005). Sub-constructs (e.g., leadership capacity) can be refined further into distributedness, states of sharedness, and connectivity among team members (Day, Gron, & Salas, 2004). The next two concepts, *group behavioural/physiological markers* and *derived features*, are critical for guiding how the translation from *group data* to constructs can be implemented. Wise and colleagues (2021) explain that higher-order group constructs need to be mapped to human observable behaviours, giving the example of the observable marker *turn-taking* being a proxy for the unobservable construct *discourse distribution*. The markers are “observable” because they can more easily be modelled or assessed by deriving features from automatically captured behavioural or physiological group data (e.g., speech or chat traces). This mapping can be achieved through multiple means, ranging from very simple rule-based algorithmic approaches to more complex models.

4. Synthesis and Discussion

This section synthesizes insights from four collaboration analytics cases and relevant literature. The first subsection (4.1) briefly describes the learning situations of the four collaboration analytics cases. The following subsections (4.2–4.5) discuss and illustrate the importance of the four main elements of the proposed collaboration analytics conceptual model, namely *theory*, *task design*, *human-centred design*, and *actionable group insights*. The last subsection (4.6) presents a discussion of emerging opportunities and potential challenges to be addressed in this field. Table 1 summarizes the illustrative cases.

Table 1. Illustrative collaboration analytics cases

Collaboration analytics case	Cohort	Mode	Data sources	Higher-order construct	Target “user”	Modelling technique
1. Communities of inquiry	6 groups of 12–14 master students	Online	Discussion logs	Collaborative problem solving	Researchers and designers	LSA, prediction, ENA, linguistic analysis
2. Team-based healthcare simulation	16 teams of 3–5 nursing students	F2F	Physiological data Positioning data Audio/video Human/system logs	Clinical resuscitation protocol	Students	Rule-based algorithm Signal processing
3. Virtual internships	12 groups of 5 high school students	Online	Chat logs	Urban planning epistemic frame	Educators	NLP, ENA Prediction
4. Collaborative concept mapping	60 triads of undergraduate students	F2F	Speech data Groupware logs Group artefact	Transactivity and symmetry	Educators	ML Visualization

4.1. The Illustrative Cases

4.1.1. Case 1: Communities of Inquiry

This case was conducted in the context of an online master’s degree in software engineering in which groups of students engaged in a four-week task (Ferreira et al., 2020; Kovanović et al., 2016) based on principles of the community of inquiry model (Garrison & Arbaugh, 2007). In this case, the analytics were aimed at supporting researchers and developers to study the progression of learners developing collaborative problem-solving skills (Gašević, Adesope, Joksimović, & Kovanović, 2015a). Group data used in this case corresponded only to textual data from the online discussion messages and metadata about such messages (e.g., whether a message is a reply to another message). The purpose of the analytics was to model cognitive and social presence, from online forum data, using a combination of predictive, linguistic, and epistemic network analysis tools.

4.1.2. Case 2: Team-Based Healthcare Simulation

This case was conducted in the context of team-based nursing simulations (Echeverria et al., 2019) in which students enact different roles to act upon the needs of a patient manikin (e.g., suffering a cardiac arrest). Data is captured via the patient manikin (e.g., CPR compressions and checking vital signs), physiological wristbands (electrodermal activity — EDA), indoor positioning trackers (x and y coordinates of each nurse), lapel microphones, and a digital logging system. A reflection tool was designed to provide feedback on students’ performance by highlighting clinical errors after a simulation (e.g., Figure 3).

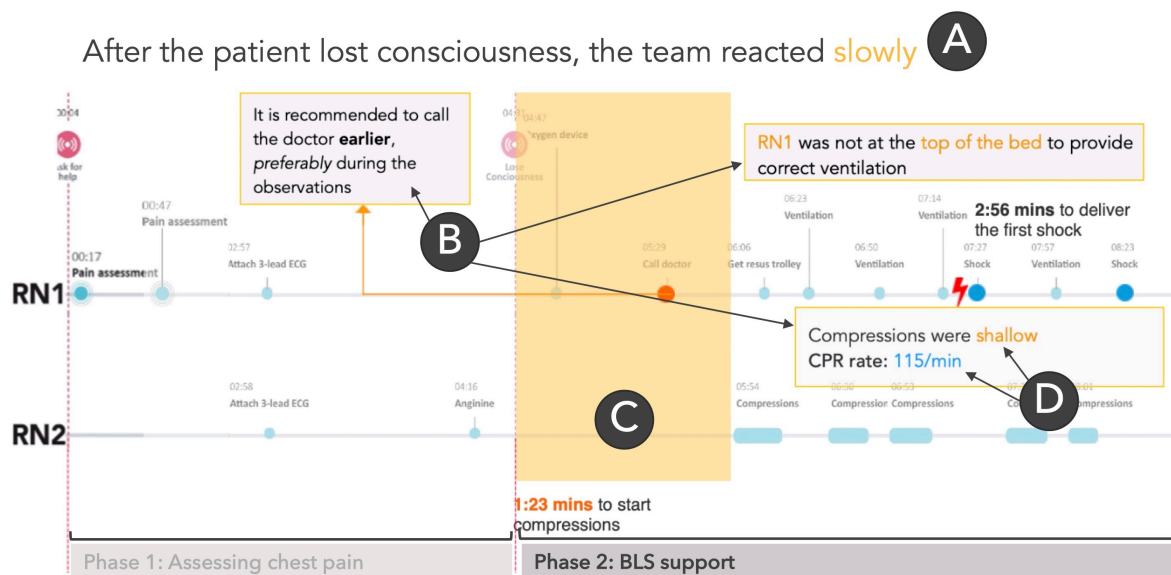


Figure 3. Team timeline for one team of nursing students. Errors are highlighted using visual elements such as (A) a descriptive title, (B) text annotations, (C) shaded areas, and (D) colour encoding.

4.1.3. Case 3: Virtual Internships

This case was conducted in the context of a virtual internship (Herder et al., 2018), in which students play the role of interns at a fictional urban planning firm. This case focuses on investigating how three teachers used an analytics dashboard — the *process tab* — that allowed them to monitor group chat discussions. The process tab (Figure 4), developed with the help of classroom teachers, presents a network model aimed at helping teachers identify the connections that are most indicative of the urban planning epistemic frame (emphasized in green colour) modelled from the chat logs.

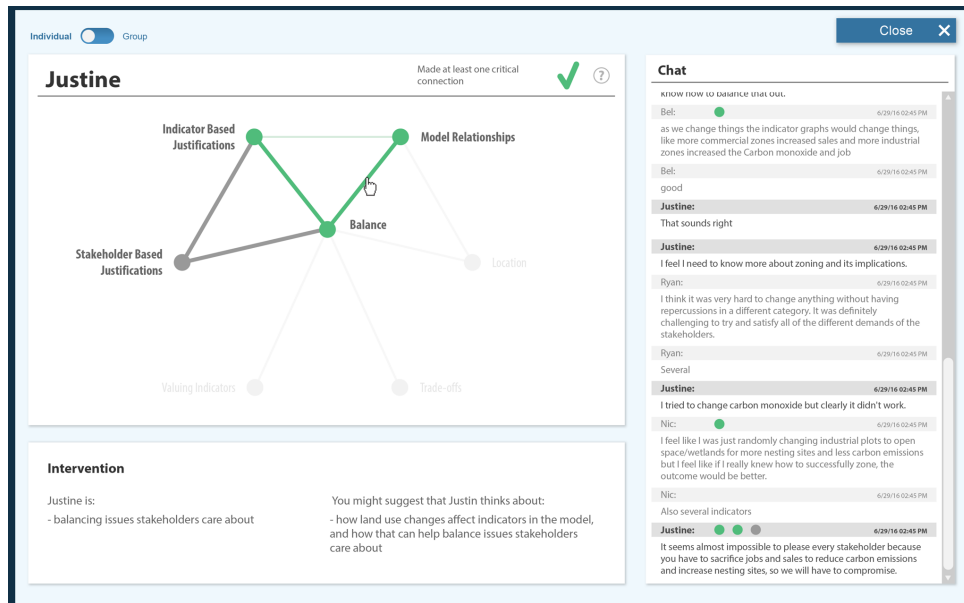


Figure 4. The process tab. Simplified network model at the top, chat record on the right, and intervention suggestions at the bottom.

4.1.4. Case 4: Collaborative Concept Mapping

This case was conducted in the context of a nutrition education initiative in which 60 undergraduate students organized in triads were asked to create concept maps collaboratively using a multitouch tabletop (Martinez-Maldonado, Kay, Wallace, & Yacef, 2011). Data was captured via the tabletop and a depth camera that differentiated student actions and a microphone array that recognized speech acts. Based on these group data, a dashboard was designed and evaluated with teachers as a post hoc reflection activity. This showed three visualizations. The *Collab-O-meter* (Figure 5, top) shows the output of an ML algorithm as a single value on a range of extent of collaboration. This algorithm uses the calculated number of actions and the index of dispersion of speech and touch activity from the triads to categorize each block of 30 seconds of group activity. The final output of the algorithm is a series of blocks of activity labelled according to the extent to which the students were collaborative (high, medium, or low level). The more blocks of activity are labelled as highly collaborative, the closer the dial points to the right extreme of the arc value axis.

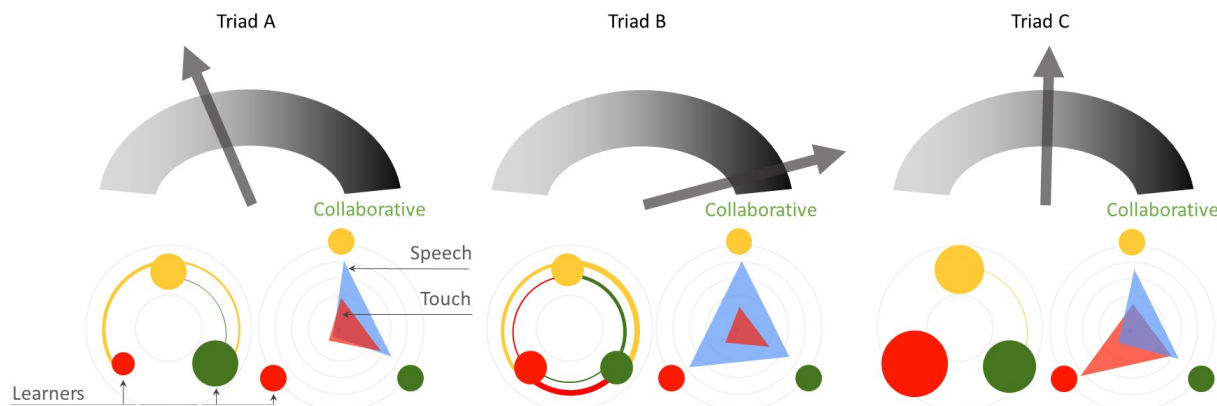


Figure 5. Top: Collab-O-meters for three triads A, B, and C. Bottom: transactivity charts (left) and radars of participation (right).

The *transactivity chart* (Figure 5, bottom left in each of the three pairs of images) quantifies connected ideas in the group concept map authored by different students. This is a bidirectional network in which the size of each node represents the total number of propositions created by each learner and the edges represent the extent to which they built on concepts created by others. Finally, the *radar of multimodal participation* (Figure 5, bottom right in each of the three pairs) quantifies the amount of speech and actions per student. This is a spider diagram with two dimensions, individual verbal and touch activity plotted in blue and red, respectively. The closer an edge of the triangle is to a coloured circle (node) the higher the number of speech acts or touch events for that particular learner.

4.2. Theory and Collaboration Analytics

Theory plays a key role in all four illustrative cases. *Educational theory* played a strong role in articulating multiple algorithms for the purpose of modelling the higher-order construct *collaborative problem solving* in Case 1 (Figure 6, A). This construct was conceptualized by building upon the CSCL theoretical model of communities of inquiry (Garrison & Arbaugh, 2007). This model posits that meaningful learning happens in online discussions when learners reach high levels of *cognitive presence* (Figure 6, B), which is underpinned by a model of practical inquiry that involves four main phases: *triggering event* (e.g., sense of puzzlement), *exploration* (e.g., brainstorming, sharing information, or leaping to conclusions), *integration* (e.g., synthesis of information and hypothesis formulation), and *resolution* (e.g., testing of hypothesis). The definition and detection of behavioural markers associated with these phases can thus serve as proxies for cognitive presence (Figure 6, C). Development of cognitive presence in collaborative problem solving is also supported by *social presence* (Hesse, Care, Buder, Sassenberg, & Griffin, 2015), which refers to the extent to which learners see each other as human beings. There are three *observable* categories of social presence: *affective*, *interactive*, and *cohesive*.

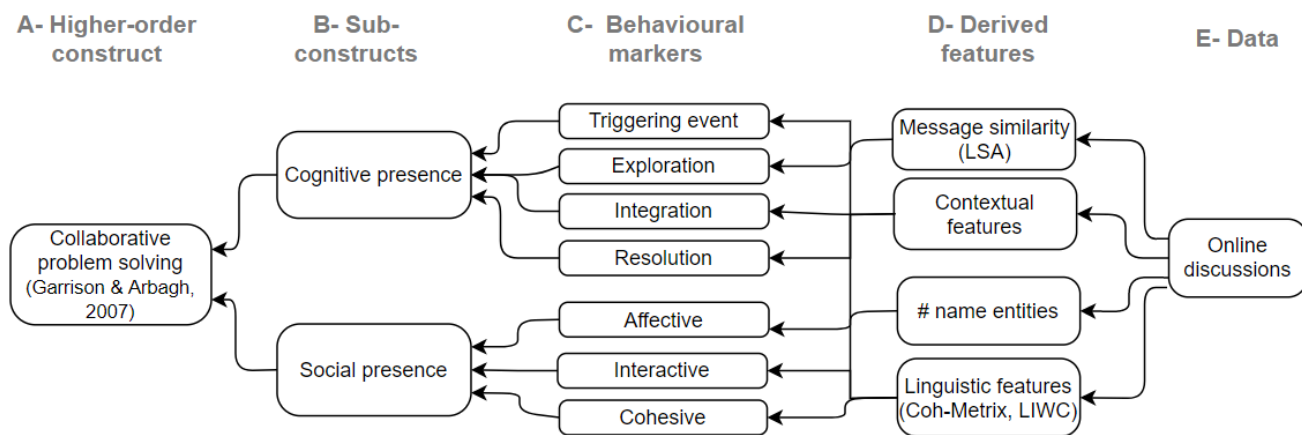


Figure 6. Mapping from online forum data to cognitive and social presence in collaborative problem solving in Case 1.

Social presence was modelled based on linguistic features (Figure 6, D) extracted from the low-level textual data of online discussion. These include those representing text cohesion measured by the tool Coh-Metrix (McNamara, Louwerse, McCarthy, & Graesser, 2010), psychological and linguistic processes measured by the tool LIWC (Tausczik & Pennebaker, 2010), and general word frequency. In addition, the analysis of cognitive presence included *contextual features* (e.g., how far a message was from the start of the thread); *message similarity*, measured with latent semantic analysis (LSA) to determine how much a message builds upon content of previous messages (Landauer, McNamara, Dennis, & Kintsch, 2013); and *number of named entities* (i.e., how many domain-specific concepts are mentioned in each message). Once the features were extracted from transcripts of and metadata about discussion messages, supervised ML algorithms (random forest) were trained to automatically code messages for phases of cognitive presence and categories of social presence (Ferreira et al., 2020; Kovanović et al., 2016). Collaborative problem solving was operationalized by creating graphs based on co-occurrences of codes on the level of (i) individual messages to examine effects of the intervention of collaborative problem solving and (ii) daily messages for every student to determine progression in collaborative problem solving. The co-occurrence graphs were then processed using epistemic network analysis (ENA). An epistemic network quantifies connections between code elements in data and represents them in dynamic network models (Shaffer, Collier, & Ruis, 2016). Figure 7 shows one of the epistemic networks, aimed at supporting researchers’ analysis, resulting from this modelling. In this figure, each node corresponds to each of the theoretically grounded behavioural markers of cognitive and social presence automatically coded from students’ online discussions.

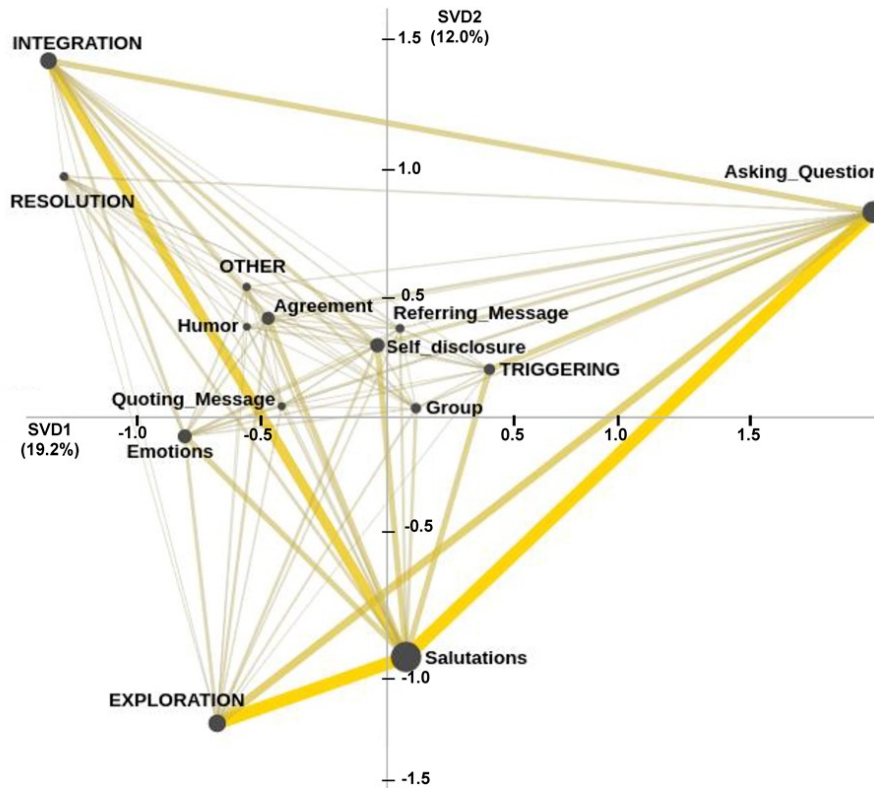


Figure 7. Epistemic network analysis (ENA) output used by researchers to investigate cognitive and social presence in Case 1.

This case illustrates how an ecology of analysis tools was articulated based on theoretical foundations to address critical questions that researchers had about collaborative problem solving. The discussion logs, after being modelled meaningfully, provided a rich source of information about cognitive and social presence, enabling researchers to measure the impact of student role allocation on group outcomes (Rolim, Ferreira, Lins, & Gašević, 2019). While *CSCW* theory played a critical role in this case, alternative theoretical underpinnings may be found in *team science* and *CSCW literature*, or in theoretical perspectives that define what effective teamwork is in a particular professional domain. For example, Tarmazdi, Vivian, Szabo, Falkner, and Falkner (2015) created a dashboard that automatically identified emerging team roles from online discussions based on a conceptual framework for self-organizing teams, while Wu, Waber, Aral, Brynjolfsson, and Pentland (2008) combined information richness and social network theories to model sensor data to predict team productivity in face-to-face workplace settings. Cases 2 and 3 (explained below) were informed by both educational theories and theoretical foundations relevant to the professional skills that students needed to develop in the contexts of healthcare and urban planning, respectively.

We turn next to the question of how collaboration analytics are integrated into the students’ task.

4.3. The Task Design and Collaboration Analytics

As discussed above, various authors have argued the importance of explicitly aligning learning analytics with the learning design, so that the data can be interpreted in context, that is, with insight into what the students were trying to accomplish (Bakharia et al., 2016; Hernandez-Leo, Martinez-Maldonado, Pardo, Muñoz-Cristobal, & Rodriguez-Triana, 2016; Lockyer et al., 2013; Persico & Pozzi, 2015; Rodríguez Triana et al., 2014). However, a recent review suggests that more work is needed for researchers to implement this alignment between metrics, learning processes, and pedagogy (Mangaroska & Giannakos, 2018). Rodriguez Triana, Prieto, Martínez-Monés, Asensio Pérez, and Dimitriadis (2018b) suggested one way to achieve this by enabling teachers to *customize* the analytics by involving them in the design process. Echeverria and colleagues (2018a) suggested *making teachers’ pedagogical intentions explicit* in the analytics interfaces themselves, by translating each intention into a visual data story presented to students. Similarly, Wiley, Dimitriadis, Bradford, and Linn (2020) proposed analytics reports, mostly based on text, generated based on teachers’ set learning goals.

Case 2 illustrates how the task design can drive the design and implementation of a collaboration analytics interface that supports students’ reflection. In this case, the higher-order construct targeted is the effective performance of basic life support

(BLS) during a group simulation created by an educator (Figure 8, A). Four sub-constructs (Figure 8, B) serve the educator to assess the effective execution of resuscitation (ANZCOR, 2016). These are (1) opening patient’s *airway*, (2) assessing whether the patient is *breathing* normally, (3) performing *compressions*, and (4) attaching a *defibrillator* as soon as available. Each sub-construct was partly modelled based on behavioural markers (Figure 8, C). For example, once the patient falls into a cardiac arrest, the airway nurse should be positioned behind the bedhead, performing a head tilt manoeuvre (*airway*), and providing ventilations (*breathing*) in synchrony with the other nurses performing compressions on the patient’s chest (*compressions*).

The mapping was implemented by defining rule-based algorithms that converted manikin logs, nurses’ critical actions, and positioning data into indicators of presence, timeliness, and/or correctness of specific events according to protocol performed by the team. The resulting interface highlighted correct procedures and errors made by the team using visual elements such as a descriptive title (A), text annotations (B) with colour encoding to show correct (blue) and incorrect (orange) actions (C), and shaded areas (D). The interface was designed so that it could be adapted to support the reflection task at the end of the simulation by presenting “data slices.” That is, each of the teachers’ pedagogical intentions was linked to how the data was presented. For example, Figure 3 highlighted the errors a team made while performing compressions and the defibrillation (sub-constructs 3 and 4). A different data slice would show other errors relevant to other sub-constructs for this or other simulations.

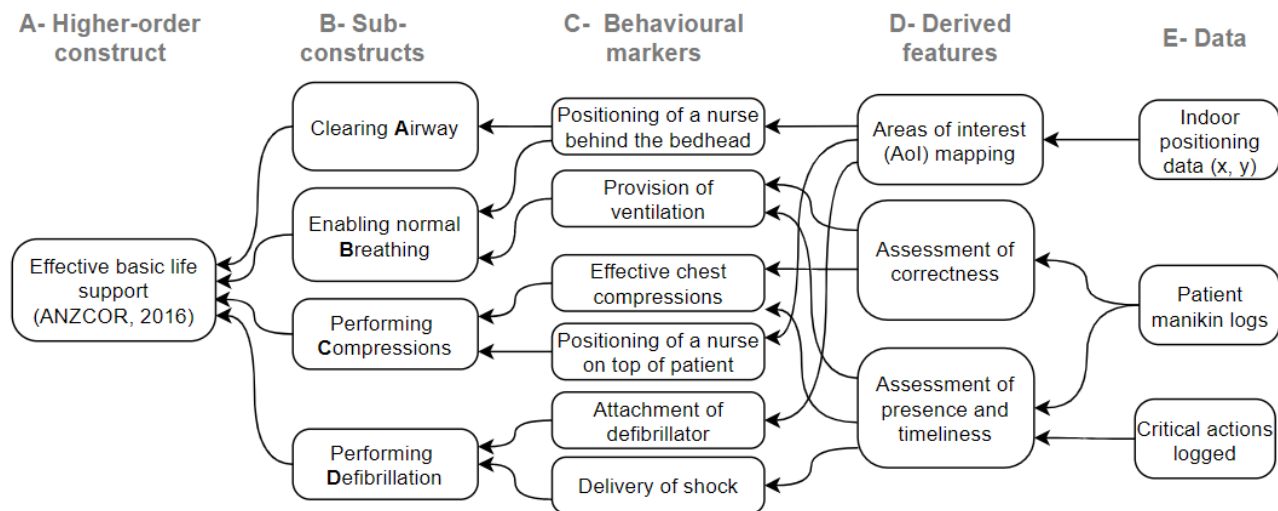


Figure 8. Mapping from indoor positioning and a combination of logs recorded manually and automatically to sub-constructs in the basic life-support protocol in Case 2.

This case presents a situation in which the collaboration analytics outputs are aimed at supporting educators and students directly rather than at informing researchers’ investigations. Moreover, each learning task (a clinical simulation) requires a particular mapping from data to higher-order constructs. This made it imperative to make teachers’ pedagogical intentions explicit both in the modelling and in the student-facing interface for both students and educators to make sense of the complex data underpinning the assessments suggested by the interface. This is thus an example of how collaboration analytics has to be aligned to the educators’ learning design, which is, at the same time, strongly connected to clinical theory in the form of guidelines that are relevant for this specific professional domain.

4.4. Human-Centred Design and Collaboration Analytics

Human-centred design approaches propose engaging with learners, educators, and other stakeholders to identify the kinds of meaningful educational constructs they value (Buckingham Shum et al., 2019b). For example, in Case 2, interviews and co-design sessions were conducted, with students, to identify the kind of feedback they needed the most (Prieto-Alvarez, Martinez-Maldonado, & Buckingham Shum, 2018) and, with teachers, to identify the kinds of data that could serve to create such analytics support according to each clinical simulation (Echeverria, 2020).

Case 3 illustrates the importance of co-designing with teachers for the use and orchestration of collaboration analytics. The *process tab* (presented in Figure 4) was designed to measure the urban planning epistemic frame (Figure 9, A). Using principles of quantitative ethnography (Shaffer et al., 2016), the urban planning epistemic frame was operationalized as the relationships or connections among salient aspects of how urban planners think and act (codes). This process involved further operationalizing the codes in terms of how they are actually expressed in the data source. These expressions are referred to as codes and in this case consist of segments of chat utterances contributed by students. For example, the occurrence of the code

“balance,” which refers to the discussion of how to address multiple stakeholder concerns simultaneously, can be modelled by coding for segments like the chat utterance “since Keanu needs to increase birds, but Louis and Kayla need to decrease birds, we need to find a good balance.” Detailed definitions of other codes have been given by Herder and colleagues (2018). Once the codes have been identified in the data, which may be done manually or via automated techniques, connections among them are operationalized as the co-occurrence of codes within windows of data and modelled using ENA.

The mapping in Case 3 from chat data to the urban planning epistemic frame was implemented in real time by combining natural language processing with a predictive ENA model. Classifiers based on regular expressions were developed to label chat utterances for the presence of the codes. ENA was used to identify and count connections between codes within team discourse using a sliding window over the utterances of a given team (Figure 9, C). In sum, Figure 9 shows the seven relevant codes as derived features that occurred in chat utterances (D), as discourse markers when they co-occurred within predefined windows of utterances (C), and as sub-constructs of the urban planning epistemic frame when the connections between codes are modelled using ENA (B — lines between codes omitted for simplicity).

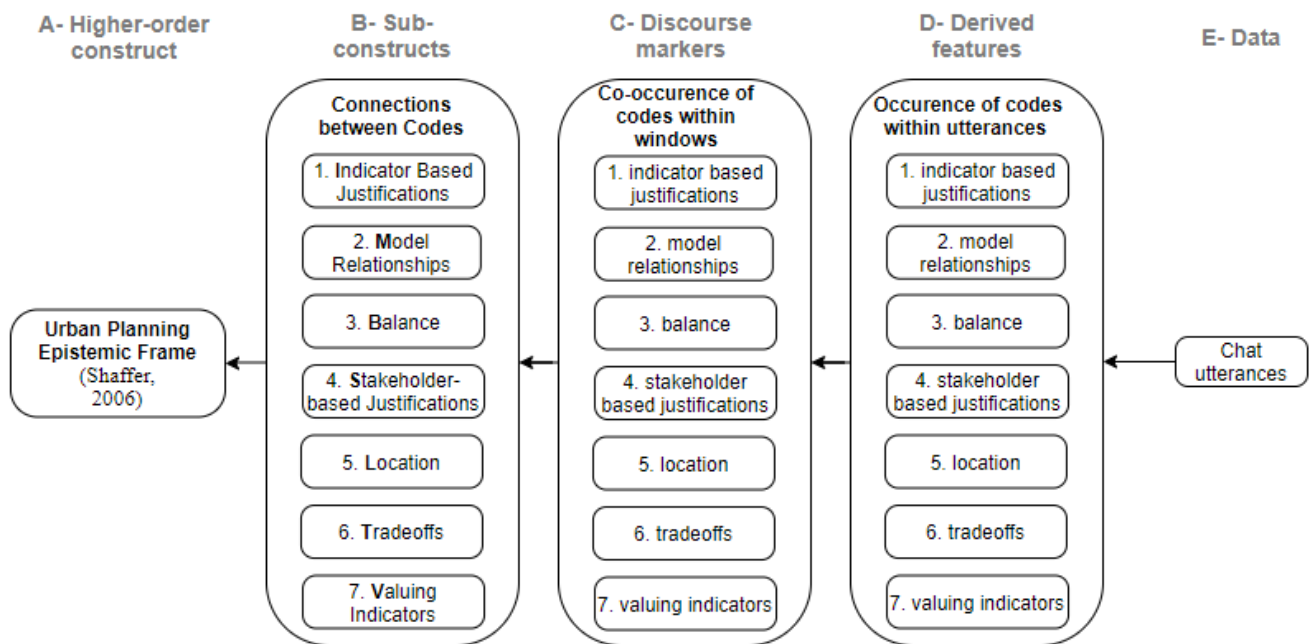


Figure 9. Mapping from chat logs to the urban planning epistemic frame in Case 3.

Results from this study demonstrated the potential of dashboards that enable teachers to monitor and investigate the *content* of collaborative interactions. However, although the process tab (Figure 4) was co-designed with teachers and the analytics were designed based on strong theoretical approaches, results also demonstrated the importance of empirical evaluation with users in authentic contexts. Specifically, teachers said that although they were able to use the process tab to point their attention to particular groups and provide feedback, in fact, their attention was more often directed to classroom orchestration tasks — answering questions, troubleshooting, and maintaining student engagement — which prevented them from using the dashboard as much as they had hoped. Thus, while this dashboard was co-designed with teachers following a human-centred design approach, their design feedback was more focused on what they *wanted* to do with the dashboard, rather than what they *could* actually do while teaching. These insights have motivated researchers to reconsider the design of the tool not only to consider the educational theory that is driving the technical aspects of the collaboration analytics innovation but also to consider the specific realities of the teaching environment in the co-design process.

Both Cases 2 and 3 illustrate how important it is to consider human-centred design practices in the same way we give importance to educational theory and the CSCL task design. An example of how these can converge is through the use of co-design to engage teachers, students, stakeholders, and experts in mapping from group data to higher-order constructs. This was achieved through two steps in the work of Echeverria (2020): (1) eliciting from teachers and students the meanings that low-level group data and analytics indicators should take on and then (2) using a well-established theory to fill the gaps in the mapping from low-level data to group constructs. A second role that human-centred design can play is in understanding the orchestration challenges faced by educators to maximize the adoption of collaboration analytics at a classroom level. For example, Prieto, Rodríguez Triana, Martínez Maldonado, Dimitriadis, and Gašević (2018) proposed a framework and

elicitation forms for conducting sessions with educators to fully understand how the innovations can actually be used under authentic conditions.

The cases and discussion presented in this and previous subsections illustrate how theory; the teachers’ pedagogical intentions; and involving teachers, students, and other stakeholders can all contribute to principled mappings from low-level group data and analytics outputs to derive meaningful semantics for the particular collaborative learning situation. We now discuss the expected outcome from effective collaboration analytics: informing actionable insights for improving collaboration and group learning.

4.5. Generating Actionable Group Insights

Although it is easy to display numbers and charts to learners and educators, this does not ensure that these will be meaningful to them (e.g., Gibson & Martinez-Maldonado, 2017; Macfadyen & Dawson, 2012; Matcha, Gasevic, & Pardo, 2019) and may not even be relevant for the task at hand (Kitto et al., 2018). In fact, Gašević and colleagues (2015b) warned about the danger of limiting system logging to certain types of actions performed by learners, and even finding significant yet spurious correlations with performance indicators, without fully understanding how these relationships can translate into sustained actionable impact on student learning and teaching practice. At the same time, Rummel, Walker, and Aleven (2016) warned about CSCL analytics systems that can feature great accuracy but still be pedagogically limited and opaque to educators’ scrutiny, which would hence hardly inform actionable insights. The visualizations of Case 4 illustrate these concerns.

In Case 4, the construct targeted by the first visualization was the extent to which a group of learners is collaborative (Figure 10, A). Each of the second two visualizations targets two sub-constructs. The first sub-construct, *transactivity* (Figure 10, B), originates from research in the learning sciences (Teasley, 1997) and refers to the extent to which learners build on each other’s reasoning. This was modelled as the times any group member created a new proposition building on concepts added by another member. The second sub-construct, *symmetry of action*, proposed in the classic CSCL literature as a strong criterion for determining a situation as collaborative (Dillenbourg, 1998; Dillenbourg & Baker, 1996; Kreijns, Kirschner, & Jochems, 2002), was modelled by measuring the extent and balance of verbal and physical activity of the group (Figure 10, C and D).

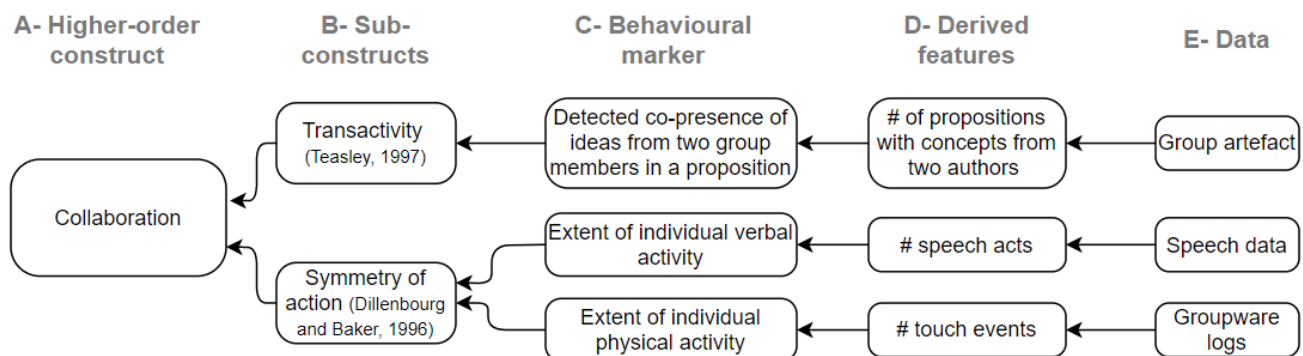


Figure 10. Mapping from speech data, multitouch data, and the group artefact to “collaboration” in Case 4.

As shown in Figure 5, the first (Collab-O-meter) visualization depicted the output of an ML model as a single value representing the level of collaboration. The other visualizations showed simple summaries of activity similar to those visualizations found in some CSCL dashboards (Liu & Nesbit, 2019; van Leeuwen et al., 2019) that mirror log data (e.g., counting clicks or speech acts). Results from this study suggested that although the Collab-O-meters were connected to the higher-order construct *collaboration*, and their minimalistic design was aimed at enabling the rapid interpretation of the ML algorithm output, teachers did not trust the interface. Since the underpinning model was not fully understood by teachers, interpretation and trust issues arose and teachers did not know what pedagogical action they could make even if the visualization reported “low levels of collaboration.” By contrast, the other two visualizations enabled teachers to identify potential meanings in the underlying low-level group data. For example, the transactivity charts enabled teachers to identify behaviours such as free-riding and learners working independently. The radars enabled teachers to identify how each learner contributed to the group by doing, talking, both, or neither.

This case, along with Case 3, illustrates how some visualizations, even if they are grounded on educational theory, may fail in supporting teachers in gaining actionable insights into students’ collaborative activity. For Case 4, although the transactivity charts and radars of participation imperfectly modelled sub-constructs of collaboration (i.e., not considering the dialogue content), teachers were able to use the evidence to gain a deeper understanding of potential groups and formulate

potential interventions. In Case 3, although the resulting collaboration analytics interface had the potential to facilitate the generation of actionable insights, the lack of consideration of the orchestration responsibilities became a threat to adoption.

4.6. Opportunities and Challenges

The four cases discussed above illustrate the diversity of analytics methods that can be applied to study or to support collaboration. At the same time, they served to emphasize the importance of the various facets of design and theory to extract and communicate meaningful insights from group data. In other words, the cases illustrate divergence in collaboration analytics methods and convergence in design/theory aspects. In this subsection, we synthesize three opportunities related to such aspects, and associated challenges, that emerge from the proposed conceptual model of collaboration analytics and the illustrative cases.

Opportunity 1: Returning meaningful group information to teachers and students. There is evident potential for collaboration analytics to serve as teachers' assistants in CSCL classrooms (Bae et al., 2019). The support and evidence provided by the analytics can augment teachers' pedagogical abilities by, for example, embedding effective feedback processes to support CSCL tasks (An, Holstein, d'Anjou, Eggen, & Bakker, 2020; Martinez-Maldonado, 2019), conducting data-informed team debriefings (as in Case 2), assessing higher-order group constructs (all cases), or redesigning the CSCL task. With adequate guidance, the opportunity is there to help students engage in socially shared regulation of their collaborative learning based on representations of their own data (e.g., Järvelä, Malmberg, Haataja, Sobocinski, & Kirschner, 2019). Case 2 and subsequent work (Martinez-Maldonado, Fernandez, Echeverria, & Buckingham Shum, 2020) also used collaboration analytics to provide "mirrors" to students in order to provoke productive reflection and dialogue. In sum, our call to actionable analytics is aimed at supporting students and educators in ways they can comprehend and, in doing this, also accelerating learning science.

The challenge: The trade-off between incompleteness and complexity of group models. All learning analytics are based on models that are an approximation to the actual learning activity (Kitto et al., 2018), and collaboration/teamwork introduces more modelling complexity than individual learning tasks. Winne (2017) has warned the community about the challenges in attempting to help learners using predictive models that are based on incomplete data. He suggests that it can be "quite hazardous to predict what will happen to a particular learner" (p. 8) based on big data. We argue that for the case of collaboration analytics, it would be even more risky to try to find causal relationships in incomplete data about such a complex phenomenon. Yet, even if we are able to create better predictive models based on more data traces, this would also increase (i) the complexity of the underlying collaboration models and (ii) the complexity of user interfaces for educators and students (Martinez-Maldonado et al., 2020).

Instead, "mirrors for reflection" could be used to address phenomena that are currently too complex to model in sufficient detail to assess individual or team progress. *Explanatory* analytics interfaces have been proposed to address this problem by enhancing visualizations with narratives that explain the meaning behind specific data points and trends in the context of group learning (Echeverria et al., 2018b). However, making automated interpretations of data on behalf of educators and students brings its own trade-offs, such as opaque black boxes (cf. Section 4.5), and unintentional pedagogical commitments (Tchounikine, 2019). Future collaboration analytics innovations will need to navigate between embracing very incomplete group models that may still provide simple feedback that helps human judgments and embracing the complexity of more complete models, more advanced automated analysis, and visualizations that intentionally hide that complexity, while maintaining the trust of users.

Opportunity 2: Creating collaboration analytics centred on authentic needs. A recent CSCL review (van Leeuwen & Rummel, 2019) points to the need for analytics tools tailored to teachers' actual needs. We argued that robust connections between higher-order group constructs and theory, the CSCL task design, and human-centred design should increase the chances of collaboration analytics being adopted by educators. Adopting human-centred design approaches can make it possible to understand how people can interact with group data representations that are tailored to their data analytic literacy, expertise (e.g., the clinical background of students in Case 2), and the particular task at hand (Buckingham Shum et al., 2019b).

The challenge: Educators and students lack a meaningful voice in analytics design. A critical challenge for vendors, and the CSCL and learning analytics communities, will be to navigate obstacles to the participation of students and educators as equal partners in the design of analytics innovations. This is not a new challenge but instead a continuous issue experienced in design practice in general (Sanders & Stappers, 2008). In academia, both the CSCL and the learning analytics communities are calling for greater learner/educator agency (Prinsloo & Slade, 2016; Tchounikine, 2019), and the lack of students' voice in the development of learning analytics is being questioned (West, Luzeckyj, Toohey, Vanderlelie, & Searle, 2020). Yet, even if teachers and students are invited to be part of a learning analytics project as partners, further challenges can emerge in integrating potentially divergent stakeholder views or conflicts between theory and teachers' pedagogical beliefs. The role of the designer is thus critical to address such differences and identify what needs to be prioritized for a particular learning context or analytics project. Nonetheless, as is widely recognized in the professional and academic communities focused on user-

centred design, a constant campaign must be fought to ensure that end-users' needs maintain visibility compared to often competing research, engineering, marketing, and other priorities. To pick a systemic challenge for academia, most university research groups lack the incentives, resources, and skills to drive their research prototypes through to 24/7 “enterprise products” in their universities. It takes concerted organizational will to build such capacity (Buckingham Shum & McKay, 2018).

Opportunity 3: Development of new theory. Tracking collaborative learning used to depend solely on hours of observational (often video) analysis conducted by trained researchers. Those researchers used their deep analytical and interpretative skills, drawing on the qualitative tradition, but this was resource intensive, necessarily limited to relatively small data samples. The embedding of multimodal sensors into the physical environment, and the automation of that analysis, means that collaboration analytics offer, in principle, an extraordinary new array of quantitative tools to generate more precise measurements of interaction. Collaboration analytics infrastructures thus represent a new form of e-research platform (akin to a “fourth paradigm” shift: Hey, Tansley, & Tolle, 2009), accelerating the pace at which scientists can engage in exploratory analyses as well as theory-driven hypothesis testing. Excitement about collaboration analytics as a theory-building tool is not, however, a rejection of the richness of the qualitative tradition. As we have argued in this paper, the fundamental challenge is to draw on theory to help forge principled links from quantitative data to the higher-order constructs that serve as the language of theory. Moreover, the emergence of hybrid methodologies such as quantitative ethnography (Shaffer, 2017) point to new ways to harmonize the complementary strengths of the qualitative and quantitative traditions when working with collaborative activity traces.

The challenge: Maturity and validity of collaboration analytics instruments. Collaboration analytics can be thought of as scientific instruments, and from an instrumentation perspective, there has been tremendous variability in the technologies used to capture interaction data from CSCL and team scenarios (Fiore, 2019). This, naturally, further complicates linking these to higher-order constructs. For example, while the study of joint attention using dual-eyetracking technologies is one of the most established areas at the intersection of learning analytics and CSCL, most developments so far do not scale beyond groups of more than two (Schneider, Worsley, & Martinez-Maldonado, 2021). A key requirement for the maturing of this field is, therefore, the rigorous validation of these new scientific instruments, so that (as in any science), researchers can make clear arguments on what they are measuring, and on the limitations and biases of a given tool. Since education is clearly not the same as the more deterministic sciences, the debate about how to claim progress in this instrumentation challenge is significant.

5. Concluding Remarks

A key contribution of the conceptualization of collaboration analytics proposed in this paper is that it blends foundations of educational measurement and the growing interest of the learning analytics community in automatically *developing actionable insights* for the particular purpose of supporting and improving collaborative learning, teamwork, and/or professional group practices. We have built on the notation for mapping from clicks to constructs from recent advances at the intersection of CSCL and learning analytics (Wise et al., 2021) to encourage collaboration between analytics researchers and designers to make explicit how they are forging principled mappings between observable low-level group data, derived indicators, and the typically unobservable higher-order constructs that are the language of educators and researchers. We offer the conceptual model of collaboration analytics, and the exemplar mappings, to help researchers rigorously theorize how collaboration constructs relate to activity traces, and to help collaboration analytics designers produce software tools that make sense to educators and students.

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Roberto Martinez-Maldonado's research is partly funded by Jacobs Foundation.

References

- Ahn, J., Campos, F., Hays, M., & DiGiacomo, D. (2019). Designing in context: Reaching beyond usability in learning analytics dashboard design. *Journal of Learning Analytics*, 6(2), 70–85. <https://doi.org/10.18608/jla.2019.62.5>
- An, P., Holstein, K., d'Anjou, B., Eggen, B., & Bakker, S. (2020). The TA framework: Designing real-time teaching augmentation for K–12 Classrooms. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '20), 25–30 April 2020, Honolulu, HI, USA (pp. 1–17). New York: ACM. <https://doi.org/10.1145/3313831.3376277>

- Anaya, A. R., González-Boticario, J., Letón, E., & Hernández-del-Olmo, F. (2015). An approach of collaboration analytics in MOOCs using social network analysis and influence diagrams. *Proceedings of the Eighth International Conference on Educational Data Mining (EDM '15)*, 26–29 June 2015, Madrid, Spain (pp. 492–495). <https://www.educationaldatamining.org/EDM2015/proceedings/short492-495.pdf>
- ANZCOR. (2016). *ARC Resuscitation Guidelines*. Melbourne, Australia: Australian and New Zealand Resuscitation Council. Retrieved from <https://resus.org.au/guidelines/>
- Aviv, R., Erlich, Z., & Ravid, G. (2003). Cohesion and roles: Network analysis of CSCL communities. *Proceedings of the IEEE International Conference on Advanced Technologies (ICAT '03)*, 9–11 July 2003, Athens, Greece (pp. 145–149). IEEE. <https://doi.org/10.1109/ICALT.2003.1215045>
- Bae, H., Glazewski, K., Hmelo-Silver, C., Lester, J., Mott, B. W., & Rowe, J. (2019). Intelligent cognitive assistants to support orchestration in CSCL. *Proceedings of the International Conference on Computer Supported Collaborative Learning (CSCL '19)*, 17–21 June 2019, Lyon, France (pp. 947–948). ISLS. <https://repository.isls.org/handle/1/1743>
- Bakharia, A., Corrin, L., de Barba, P., Kennedy, G., Gašević, D., Mulder, R., . . . Lockyer, L. (2016). A conceptual framework linking learning design with learning analytics. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (LAK '16)*, 25–29 April 2016, Edinburgh, Scotland (pp. 329–338). New York: ACM. <https://doi.org/10.1145/2883851.2883944>
- Behrens, J. T., DiCerbo, K. E., & Foltz, P. W. (2019). Assessment of complex performances in digital environments. *The Annals of the American Academy of Political and Social Science*, 683(1), 217–232. <https://doi.org/10.1177/0002716219846850>
- Buckingham Shum, S., & Crick, R. D. (2016). Learning analytics for 21st century competencies. *Journal of Learning Analytics*, 3(2), 6–21. <https://doi.org/10.18608/jla.2016.32.2>
- Buckingham Shum, S., Echeverria, V., & Martinez-Maldonado, R. (2019a). The multimodal matrix as a quantitative ethnography methodology. *Proceedings of the International Conference on Quantitative Ethnography (ICQE '19)*, 19–22 October 2019, Madison, WI, USA (pp. 26–40). Springer. https://doi.org/10.1007/978-3-030-33232-7_3
- Buckingham Shum, S., & Ferguson, R. (2012). Social learning analytics. *Journal of Educational Technology & Society*, 15(3), 3–26. <https://www.jstor.org/stable/jeductechsoci.15.3.3>
- Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019b). Human-centred learning analytics. *Journal of Learning Analytics*, 6(2), 1–9. <https://doi.org/10.18608/jla.2019.62.1>
- Buckingham Shum, S., & McKay, T. A. (2018). Architecting for learning analytics: Innovating for sustainable impact. *EDUCAUSE Review*, March/April 2018, 25–37. <https://er.educause.edu/articles/2018/3/architecting-for-learning-analytics-innovating-for-sustainable-impact>
- Bull, S., & Vatrappu, R. (2011). Supporting collaborative interaction with open learner models: Existing approaches and open questions. *Proceedings of the Ninth International Conference on Computer Supported Collaborative Learning 2011 (CSCL '11)*, 4–8 July 2011, Hong Kong, China (pp. 761–765). ISLS.
- Chandrasegaran, S., Bryan, C., Shidara, H., Chuang, T.-Y., & Ma, K.-L. (2019). TalkTraces: Real-time capture and visualization of verbal content in meetings. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '19)*, 4–9 May 2019, Glasgow, UK (pp. 577:571–577:514). New York: ACM. <https://doi.org/10.1145/3290605.3300807>
- Chen, J., Wang, M., Kirschner, P. A., & Tsai, C.-C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research*, 88(6), 799–843. <https://doi.org/10.3102/0034654318791584>
- Choudhury, T., & Pentland, A. (2002). The sociometer: A wearable device for understanding human networks. *Proceedings of the Workshop: Ad Hoc Communications and Collaboration in Ubiquitous Computing Environments at CSCW '02*, 16–20 November 2002, New Orleans, USA (pp. 1–6). MIT. <https://alumni.media.mit.edu/~tanzeem/shortcuts/workingpaper.pdf>
- Clow, D. (2012). The learning analytics cycle: closing the loop effectively. *Proceedings of the Second International Conference on Learning Analytics & Knowledge (LAK '12)*, 29 April–2 May 2012, Vancouver, Canada (pp. 134–138). New York: ACM. <https://doi.org/10.1145/2330601.2330636>
- Clow, D. (2014). Data wranglers: Human interpreters to help close the feedback loop. *Proceedings of the Fourth International Conference on Learning Analytics & Knowledge (LAK '14)*, 24–28 March 2014, Indianapolis, USA (pp. 49–53). New York: ACM. <https://doi.org/10.1145/2567574.2567603>
- Colvin, C., Rogers, T., Wade, A., Dawson, S., Gasevic, D., Buckingham Shum, S., . . . Kennedy, G. (2016). *Student Retention and Learning Analytics: A Snapshot of Australian Practices and a Framework for Advancement*. Australian Government Office for Learning and Teaching (Report), University of South Australia. Retrieved from <https://opus.lib.uts.edu.au/handle/10453/117173>

- Cooper, A. (2012). What is analytics? Definition and essential characteristics. *CETIS Analytics Series*, 1(5), 1–10. <https://publications.cetis.org.uk/2012/521>
- Corrin, L., & De Barba, P. (2015). How do students interpret feedback delivered via dashboards? *Proceedings of the Fifth International Conference on Learning Analytics & Knowledge (LAK '15)*, 16–20 March 2015, Poughkeepsie, New York, USA (pp. 430–431). New York: ACM. <https://doi.org/10.1145/2723576.2723662>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281. <https://doi.org/10.1037/h0040957>
- Davenport, T. H., & Harris, J. G. (2007). *Competing on Analytics: The New Science of Winning* (Vol. 84). Cambridge, MA, USA: Harvard Business School Press.
- Dawson, S., Mirriahi, N., & Gasevic, D. (2015). Importance of theory in learning analytics in formal and workplace settings. *Journal of Learning Analytics*, 2(2), 1–4. <https://doi.org/10.18608/jla.2015.22.1>
- Day, D. V., Gronn, P., & Salas, E. (2004). Leadership capacity in teams. *The Leadership Quarterly*, 15(6), 857–880. <https://doi.org/10.1016/j.leaqua.2004.09.001>
- De Liddo, A., Buckingham Shum, S., Quinto, I., Bachler, M., & Cannavacciuolo, L. (2011). Discourse-centric learning analytics. *Proceedings of the First International Conference on Learning Analytics & Knowledge (LAK '11)*, 27 February–1 March 2011, Banff, AB, Canada (pp. 23–33). New York: ACM. <https://doi.org/10.1145/2090116.2090120>
- Dillenbourg, P. (1998). What do you mean by “collaborative learning”? In P. Dillenbourg (Ed.), *Collaborative Learning: Cognitive and Computational Approaches. Advances in Learning and Instruction Series* (pp. 1–19). Oxford: Elsevier Science.
- Dillenbourg, P. (1999). *Collaborative Learning: Cognitive and Computational Approaches. Advances in Learning and Instruction Series*. New York: Elsevier.
- Dillenbourg, P., & Baker, M. (1996). Negotiation spaces in human-computer collaborative learning. *Proceedings of the International Conference on Design of Cooperative Systems (COOP '96)*, 12–14 June 1996, Juan-Les-Pins, France (pp. 187–206). EUSSET.
- Dollinger, M., Liu, D., Arthars, N., & Lodge, J. (2019). Working together in learning analytics towards the co-creation of value. *Journal of Learning Analytics*, 6(2), 10–26. <https://doi.org/10.18608/jla.2019.62.2>
- Dönmez, P., Rosé, C., Stegmann, K., Weinberger, A., & Fischer, F. (2005). Supporting CSCL with automatic corpus analysis technology. *Proceedings of the International Conference on Computer-Support for Collaborative Learning (CSCL '05)*, 30 May–4 June 2005, Taipei, Taiwan (pp. 125–134). ISLS. <https://dl.acm.org/doi/10.5555/1149293.1149310>
- Dowell, N. M. M., Nixon, T. M., & Graesser, A. C. (2019). Group communication analysis: A computational linguistics approach for detecting sociocognitive roles in multiparty interactions. *Behavior Research Methods*, 51(3), 1007–1041. <https://doi.org/10.3758/s13428-018-1102-z>
- Drachsler, H., Hoel, T., Scheffel, M., Kismihók, G., Berg, A., Ferguson, R., . . . Manderveld, J. (2015). Ethical and privacy issues in the application of learning analytics. *Proceedings of the Fifth International Conference on Learning Analytics & Knowledge (LAK '15)*, 16–20 March 2015, Poughkeepsie, New York, USA (pp. 390–391). New York: ACM. <https://doi.org/10.1145/2723576.2723642>
- Echaluze, M. L. S., Fidalgo-Blanco, A., Esteban-Escano, J., Peñalvo, F. J. G., & González, M. Á. C. (2018). Using learning analytics to detect authentic leadership characteristics in engineering students. *The International Journal of Engineering Education*, 34(3), 851–864. <http://repositorio.grial.eu/handle/grial/1582>
- Echeverria, V. (2020). *Designing and Validating Automated Feed-back for Collocated Teams Using Multimodal Learning Analytics* (PhD in Learning Analytics). University of Technology Sydney (UTS), Sydney, Australia.
- Echeverria, V., Martinez-Maldonado, R., & Buckingham Shum, S. (2019). Towards collaboration translucence: Giving meaning to multimodal group data. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '19)*, 4–9 May 2019, Glasgow, UK (paper number 39, pp. 1–16). New York: ACM. <https://doi.org/10.1145/3290605.3300269>
- Echeverria, V., Martinez-Maldonado, R., Buckingham Shum, S., Chiluiza, K., Granda, R., & Conati, C. (2018a). Exploratory versus explanatory visual learning analytics: Driving teachers’ attention through educational data storytelling. *Journal of Learning Analytics*, 5(3), 72–97. <https://doi.org/10.18608/jla.2018.53.6>
- Echeverria, V., Martinez-Maldonado, R., Granda, R., Chiluiza, K., Conati, C., & Buckingham Shum, S. (2018b). Driving data storytelling from learning design. *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge (LAK '18)*, 5–9 March 2018, Sydney, Australia (pp. 131–140). New York: ACM. <https://doi.org/10.1145/3170358.3170380>

- Echeverria, V., Martinez-Maldonado, R., Power, T., Hayes, C., & Buckingham Shum, S. (2018c). Where is the nurse? Towards automatically visualising meaningful team movement in healthcare education. *Proceedings of the International Conference on Artificial Intelligence in Education (AIED '18)*, 27–30 June 2018, London, UK (pp. 74–78). Springer. https://doi.org/10.1007/978-3-319-93846-2_14
- Ellis, C. A., Gibbs, S. J., & Rein, G. (1991). Groupware: Some issues and experiences. *Communications of the ACM*, 34(1), 39–58. <https://doi.org/10.1145/99977.99987>
- Ferreira, M., Rolim, V., Mello, R. F., Lins, R. D., Chen, G., & Gašević, D. (2020). Towards automatic content analysis of social presence in transcripts of online discussions. *Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK '20)*, 23–27 March 2020, Frankfurt, Germany (pp. 141–150). New York: ACM. <https://doi.org/10.1145/3375462.3375495>
- Fiore, S. (2019). Integrating theorizing on embodied, enactive, extended, and embedded cognition to augment CSCL research. *Keynote at the International Conference of Computer-Supported Collaborative Learning (CSCL '19)*, 17–21 June 2019, Lyon, France. <https://www.cscl2019.com/en/keynotes-broadcast/51>.
- Fiore, S. M. (2008). Interdisciplinarity as teamwork: How the science of teams can inform team science. *Small Group Research*, 39(3), 251–277. <https://doi.org/10.1177/1046496408317797>
- Fitzpatrick, G., & Ellingsen, G. (2013). A review of 25 years of CSCW research in healthcare: Contributions, challenges and future agendas. *Computer Supported Cooperative Work*, 22(4-6), 609–665. <https://doi.org/10.1007/s10606-012-9168-0>
- 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>
- Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education*, 10(3), 157–172. <https://doi.org/10.1016/j.iheduc.2007.04.001>
- Gašević, D., Adesope, O., Joksimović, S., & Kovanović, V. (2015a). Externally-facilitated regulation scaffolding and role assignment to develop cognitive presence in asynchronous online discussions. *The Internet and Higher Education*, 24, 53–65. <https://doi.org/10.1016/j.iheduc.2014.09.006>
- Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, 28, 68–84. <https://doi.org/10.1016/j.iheduc.2015.10.002>
- Gašević, D., Dawson, S., & Siemens, G. (2015b). 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. (2018). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior*, 92(March 2019), 562–577. <https://doi.org/10.1016/j.chb.2018.07.003>
- Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the learning analytics puzzle: A consolidated model of a field of research and practice. *Learning: Research and Practice*, 3(1), 63–78. <https://doi.org/10.1080/23735082.2017.1286142>
- Gibson, A., & Martinez-Maldonado, R. (2017). That dashboard looks nice, but what does it mean? Towards making meaning explicit in learning analytics design. *Proceedings of the Australian Conference on Computer-Human Interaction (OzCHI '17)*, 28 November–1 December 2017, Brisbane, Australia (pp. 528–532). New York: ACM. <https://doi.org/10.1145/3152771.3156171>
- Goggins, S., Jahnke, I., & Wulf, V. (2012). CSCL@work revisited — Beyond CSCL and CSCW? Are there key design principles for computer supported collaborative learning at the workplace? *Proceedings of the International Conference on Supporting Group Work (GROUP '12)*, 27–31 October 2012, Sanibel Island, FL, USA (pp. 323–326). New York: ACM. <https://doi.org/10.1145/2389176.2389239>
- Gomez, K., Kyza, E. A., & Mancevice, N. (2018). Participatory design and the learning sciences. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (pp. 401–409). New York: Taylor and Francis. <https://doi.org/10.4324/9781315617572-39>
- Goodyear, P. M., Banks, S., Hodgson, V., & McConnell, D. (2006). *Advances in Research on Networked Learning* (Vol. 4). Dordrecht, Netherlands: Springer Science & Business Media. <https://doi.org/10.1007/1-4020-7909-5>
- Graesser, A., Dowell, N., Hampton, A. J., Lippert, A. M., Li, H., & Shaffer, D. W. (2018). Building intelligent conversational tutors and mentors for team collaborative problem solving: Guidance from the 2015 Program for International Student Assessment. In *Building Intelligent Tutoring Systems for Teams* (Vol. 19, pp. 173–211). Emerald Publishing Limited. <https://doi.org/10.1108/S1534-085620180000019012>
- Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society*, 15(3), 42–57. <https://www.jstor.org/stable/10.2307/jeductechsoci.15.3.42>

- Grudin, J. (1991). CSCW. *Communications of the ACM*, 34(12), 30–34. <https://doi.org/10.1145/125319.125320>
- 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>
- Harrer, A. (2013). Analytics of collaborative planning in Metafora: Architecture, data, and analytic methods. *Proceedings of the Third International Conference on Learning Analytics & Knowledge (LAK '13)*, 8–12 April 2013, Leuven, Belgium (pp. 255–259). New York: ACM. <https://doi.org/10.1145/2460296.2460348>
- Herder, T., Swiecki, Z., Fougat, S. S., Tamborg, A. L., Allsopp, B. B., Shaffer, D. W., & Misfeldt, M. (2018). Supporting teachers' intervention in students' virtual collaboration using a network based model. *Proceedings of the Eighth International Conference on Learning Analytics & Knowledge (LAK '18)*, 5–9 March 2018, Sydney, Australia (pp. 21–25). New York: ACM. <https://doi.org/10.1145/3170358.3170394>
- Hernandez-Leo, D., Martinez-Maldonado, R., Pardo, A., Muñoz-Cristobal, J., & Rodriguez-Triana, M. (2016). Analytics for learning design: A layered framework and tools. *British Journal of Educational Technology*, 50(1), 139–152. <https://doi.org/10.1111/bjet.12645>
- Hernandez-Leo, D., Villasclaras-Fernandez, E. D., Asensio-Perez, J. I., Dimitriadis, Y. A., & Retalis, S. (2006). CSCL scripting patterns: Hierarchical relationships and applicability. *Proceedings of the International Conference on Advanced Learning Technologies (ICALT '06)*, 5–7 July 2006, Kerkrade, Netherlands (pp. 388–392). IEEE. <https://doi.org/10.1109/ICALT.2006.1652452>
- Hernández Leo, D., Asensio-Pérez, J. I., Dimitriadis, Y., & Villasclaras-Fernández, E. D. (2010). Generating CSCL scripts: From a conceptual model of pattern languages to the design of real scripts. In P. Goodyear & S. Retalis (Eds.), *Technology-Enhanced Learning: Design Patterns and Pattern Languages* (pp. 49–64). Rotterdam: Sense Publishers. https://doi.org/10.1163/9789460910623_004
- Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills: Methods and Approach* (pp. 37–56). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-017-9395-7_2
- Hey, A., Tansley, S., & Tolle, K. (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Redmond, WA, USA: Microsoft Research.
- Isotani, S. (2011). Guest editorial: Special issue on intelligent and innovative support systems for CSCL. *IEEE Transactions on Learning Technologies*, 4(1), 1–4. <https://doi.org/10.1109/TLT.2011.7>
- Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2019). What multimodal data can tell us about the students' regulation of their learning process. *Learning and Instruction*, 4, in press. <https://doi.org/10.1016/j.learninstruc.2019.04.004>
- Jeong, H., & Hmelo-Silver, C. E. (2010). Technology use in CSCL: A content meta-analysis. *Proceedings of the Hawaii International Conference on System Sciences*, 5–8 January 2010, Honolulu, HI, USA (pp. 1–10). IEEE. <https://doi.org/10.1109/HICSS.2010.364>
- Jeong, H., Hmelo-Silver, C. E., & Jo, K. (2019). Ten years of computer-supported collaborative learning: A meta-analysis of CSCL in STEM education during 2005–2014. *Educational Research Review*, 28, 100284. <https://doi.org/10.1016/j.edurev.2019.100284>
- Jermann, P., & Dillenbourg, P. (2008). Group mirrors to support interaction regulation in collaborative problem solving. *Computers & Education*, 51(1), 279–296. <https://doi.org/10.1016/j.compedu.2007.05.012>
- Johnson, S. D., Suriya, C., Yoon, S. W., Berrett, J. V., & La Fleur, J. (2002). Team development and group processes of virtual learning teams. *Computers & Education*, 39, 379–393. [https://doi.org/10.1016/S0360-1315\(02\)00074-X](https://doi.org/10.1016/S0360-1315(02)00074-X)
- Joksimovic, S., Gasevic, D., Kovanovic, V., Adesope, O., & Hatala, M. (2014). Psychological characteristics in cognitive presence of communities of inquiry: A linguistic analysis of online discussions. *The Internet and Higher Education*, 22, 1–10. <https://doi.org/10.1016/j.iheduc.2014.03.001>
- Jørnø, R. L., & Gynther, K. (2018). What constitutes an “actionable insight” in learning analytics? *Journal of Learning Analytics*, 5(3), 198–221. <https://doi.org/10.18608/jla.2018.53.13>
- Kasempalu, R., Santos, L. P. P., & Ley, T. T. (2019). Providing teachers with individual and group-level collaboration analytics: A paper prototype. *Proceedings of the International Workshop on Collaboration Analytics: Making Learning Visible in Collaborative Settings* (pp. 1–4). Held at the International Conference on Computer-Supported Collaborative Learning (CSCL '19), 17–21 June 2019, Lyon, France. <https://collaborationanalytics.files.wordpress.com/2019/06/submission-2-reet.pdf>

- Kay, J., Maisonneuve, N., Yacef, K., & Zaïane, O. (2006). Mining patterns of events in students' teamwork data. *Proceedings of the Workshop on Educational Data Mining* (pp. 45–52). Held at the Eighth International Conference on Intelligent Tutoring Systems (ITS '06), 26–30 June 2006, Jhongli, Taiwan. IEDMS. https://www.educationaldatamining.org/ITS2006EDM/Kay_Yacef.pdf
- Khan, S. M. (2017). Multimodal behavioral analytics in intelligent learning and assessment systems. In A. A. von Davier, M. Zhu, & P. C. Kyllonen (Eds.), *Innovative Assessment of Collaboration* (pp. 173–184). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-33261-1_11
- Kim, J., & Kumar, R. (Eds.). (2012). *Proceedings of the Full-Day Workshop on Intelligent Support for Learning In Groups*. Held at the 11th International Conference on Intelligent Tutoring Systems (ITS 2012), 15 June 2012, Chania, Greece. <https://sites.google.com/site/islg2012/>
- Kim, J. W., Sottolare, R. A., Brawner, K., & Flowers, T. (2018). Integrating sensors and exploiting sensor data with GIFT for improved learning analytics. *Proceedings of the Annual GIFT Users Symposium* (GIFTSym6), 9–11 May 2018, Orlando, FL, USA (pp. 299–312). https://gifttutoring.org/attachments/download/2695/13_GIFTSym6_Exp_Analysis_Eval_paper_15.pdf
- Kitto, K., Buckingham Shum, S., & Gibson, A. (2018). Embracing imperfection in learning analytics. *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (LAK '18), 5–9 March 2018, Sydney, Australia (pp. 451–460). New York: ACM. <https://doi.org/10.1145/3170358.3170413>
- Knight, S., & Buckingham Shum, S. (2017). Theory and learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of Learning Analytics* (pp. 17–22). Edmonton, AB, Canada: SOLAR. <https://doi.org/10.18608/hla17.001>
- Knipfer, K., Prilla, M., Cress, U., & Herrmann, T. (2011). Computer support for collaborative reflection on captured teamwork data. *Proceedings of the International Conference on Computer Supported Collaborative Learning* (CSCL '11), 4–8 July 2011, Hong Kong, China (pp. 938–939). ISLS.
- Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts — A conceptual analysis. *Educational Psychology Review*, 18(2), 159–185. <https://doi.org/10.1007/s10648-006-9007-2>
- 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. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (LAK '16), 25–29 April 2016, Edinburgh, Scotland (pp. 15–24). New York: ACM. <https://doi.org/10.1145/2883851.2883950>
- Kreijns, K., Kirschner, P. A., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environments. *Journal of Educational Technology & Society*, 5(1), 8–22. <https://www.jstor.org/stable/jeductechsoci.5.1.8>
- Kumar, R., & Kim, J. (2014). Special issue on intelligent support for learning in groups. *International Journal of Artificial Intelligence in Education*, 24(1), 1–7. <https://doi.org/10.1007/s40593-013-0013-5>
- Kumar, R., Rosé, C. P., Wang, Y.-C., Joshi, M., & Robinson, A. (2007). Tutorial dialogue as adaptive collaborative learning support. *Proceedings of the International Conference on Artificial Intelligence in Education* (AIED '07), 9–13 July 2007, Los Angeles, CA, USA (pp. 383–390). IOS Press. <https://dl.acm.org/doi/10.5555/1563601.1563663>
- Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2013). *Handbook of Latent Semantic Analysis*. New York: Psychology Press.
- 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. <https://doi.org/10.18608/jla.2017.43.5>
- Lim, L., Dawson, S., Joksimovic, S., & Gašević, D. (2019). Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference? *Proceedings of the Ninth International Conference on Learning Analytics & Knowledge* (LAK '19), 4–8 March 2019, Tempe, Arizona, USA (pp. 250–259). New York: ACM. <https://doi.org/10.1145/3303772.3303804>
- Liu, A. L., & Nesbit, J. C. (2019). Dashboards for computer-supported collaborative learning. In M. Virvou, E. Alepis, G. A. Tsihrantzis, & L. C. Jain (Eds.), *Machine Learning Paradigms: Advances in Learning Analytics* (pp. 157–182). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-13743-4_9
- Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist*, 57(10), 1439–1459. <https://doi.org/10.1177/0002764213479367>
- Ludvigsen, S., Cress, U., Law, N., Stahl, G., & Rosé, C. P. (2017). Future direction for the CSCL field: Methodologies and eight controversies. *International Journal of Computer-Supported Collaborative Learning*, 12(4), 337–341. <https://doi.org/10.1007/s11412-017-9268-4>

- Macfadyen, L. P., & Dawson, S. (2012). Numbers are not enough. Why e-learning analytics failed to inform an institutional strategic plan. *Journal of Educational Technology & Society*, 15(3), 149–163. <https://www.jstor.org/stable/10.2307/jeductechsoci.15.3.149>
- 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*. <https://doi.org/10.1016/j.chb.2018.06.030>
- Maltese, A. V., Harsh, J. A., & Svetina, D. (2015). Data visualization literacy: Investigating data interpretation along the novice–expert continuum. *Journal of College Science Teaching*, 45(1), 84–90. <https://www.jstor.org/stable/43631889>
- Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, 42(1), 30–37. <https://doi.org/10.3102/0013189x12459803>
- Mangaroska, K., & Giannakos, M. N. (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, 12(4), 516–534. <https://doi.org/10.1109/TLT.2018.2868673>
- Martinez-Maldonado, R. (2019). A handheld classroom dashboard: Teachers' perspectives on the use of real-time collaborative learning analytics. *International Journal of Computer-Supported Collaborative Learning*, 14(3), 383–411. <https://doi.org/10.1007/s11412-019-09308-z>
- Martinez-Maldonado, R., Elliot, D., Axisa, C., Power, T., & Buckingham Shum, S. (2019a). Making the design of CSCL analytics interfaces a co-design process: The case of multimodal teamwork in healthcare. *Proceedings of the International Conference on Computer-Supported Collaborative Work (CSCL '19)*, 17–21 June 2019, Lyon, France (pp. 859–860). ISLS. <https://repository.isls.org/handle/1/1698>
- Martinez-Maldonado, R., Fernandez, G., Echeverria, V., & Buckingham Shum, S. (2020). From data to insights: A layered storytelling approach for multimodal learning analytics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '20)*, 25–30 April 2020, Honolulu, HI, USA (pp. 1–15). New York: ACM. <https://doi.org/10.1145/3313831.3376148>
- Martinez-Maldonado, R., Goodyear, P., Kay, J., Thompson, K., & Carvalho, L. (2016). An actionable approach to understand group experience in complex, multi-surface spaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '16)*, 7–12 May 2016, San Jose, California, USA (pp. 2062–2074). New York: ACM. <https://doi.org/10.1145/2858036.2858213>
- Martinez-Maldonado, R., Kay, J., Buckingham Shum, S., & Yacef, K. (2019b). Collocated collaboration analytics: Principles and dilemmas for mining multimodal interaction data. *Human–Computer Interaction*, 34(1), 1–50. <https://doi.org/10.1080/07370024.2017.1338956>
- Martinez-Maldonado, R., Kay, J., Wallace, J., & Yacef, K. (2011). Modelling symmetry of activity as an indicator of collocated group collaboration. *Proceedings of the International Conference on User Modeling, Adaptation and Personalization (UMAP '11)*, 11–15 July 2011, Girona, Spain (pp. 196–204). Berlin: Springer. https://doi.org/10.1007/978-3-642-22362-4_18
- Martinez-Maldonado, R., Buckingham Shum, S., Schneider, B., Charleer, S., Klerkx, J., & Duval, E. (2017). Learning analytics for natural user interfaces: A framework, case studies and a maturity analysis. *Journal of Learning Analytics*, 4(1), 24–57. <https://doi.org/10.18608/jla.2017.41.4>
- Martinez-Maldonado, R., Worsley, M., Schneider, B., & Kharrufa, A. (2019c). International Workshop on Collaboration Analytics: Making learning visible in collaborative settings. *Proceedings of the International Conference on Computer-Supported Collaborative Learning (CSCL '19)*, 17–21 June 2019, Lyon, France (pp. 1044–1049). ISLS.
- Matcha, W., Gasevic, D., & Pardo, A. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, 13(2), 226–245. <https://doi.org/10.1109/TLT.2019.2916802>
- Mazzocchi, F. (2015). Could Big Data be the end of theory in science? A few remarks on the epistemology of data-driven science. *EMBO Reports*, 16(10), 1250–1255. <https://doi.org/10.15252/embr.201541001>
- McNamara, D. S., Louwse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes*, 47(4), 292–330. <https://doi.org/10.1080/01638530902959943>
- Milligan, S. K. (2020). Standards for developing assessments of learning using process data. In M. Bearman, D. Boud, P. Dawson, J. Tai, & R. Ajjawi (Eds.), *Re-imagining University Assessment in a Digital World* (pp. 179–192). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-3-030-41956-1_13
- Milligan, S. K., & Griffin, P. (2016). Understanding learning and learning design in MOOCs: A measurement-based interpretation. *Journal of Learning Analytics*, 3(2), 88–115. <https://doi.org/10.18608/jla.2016.32.5>
- Mislevy, R. J. (2019). Advances in measurement and cognition. *The Annals of the American Academy of Political and Social Science*, 683(1), 164–182. <https://doi.org/10.1177/0022716219843816>

- Mor, Y., Ferguson, R., & Wasson, B. (2015). Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action. *British Journal of Educational Technology*, 46(2), 221–229. <https://doi.org/10.1111/bjet.12273>
- Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2019). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior*, 100(Nov 2019), 298–304. <https://doi.org/10.1016/j.chb.2018.12.019>
- Olsen, J. (2017). *Orchestrating Combined Collaborative and Individual Learning in the Classroom* (doctoral dissertation). Carnegie Mellon University, Pittsburgh, PA, USA.
- Pardo, A., Mirriahi, N., Martinez-Maldonado, R., Jovanovic, J., Dawson, S., & Gašević, D. (2016). Generating actionable predictive models of academic performance. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (LAK '16)*, 25–29 April 2016, Edinburgh, Scotland (pp. 474–478). New York: ACM. <https://doi.org/10.1145/2883851.2883870>
- Pauleen, D. (2004). *Virtual Teams: Projects, Protocols and Processes*. Hershey, PA, USA: IGI Global.
- Perera, D., Kay, J., Koprinska, I., Yacef, K., & Zaiane, O. R. (2008). Clustering and sequential pattern mining of online collaborative learning data. *IEEE Transactions on Knowledge and Data Engineering*, 21(6), 759–772. <https://doi.org/10.1109/TKDE.2008.138>
- Persico, D., & Pozzi, F. (2015). Informing learning design with learning analytics to improve teacher inquiry. *British Journal of Educational Technology*, 46(2), 230–248. <https://doi.org/10.1111/bjet.12207>
- Pijera-Díaz, H. J., Drachsler, H., Järvelä, S., & Kirschner, P. A. (2016). Investigating collaborative learning success with physiological coupling indices based on electrodermal activity. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (LAK '16)*, 25–29 April 2016, Edinburgh, Scotland (pp. 64–73). New York: ACM. <https://doi.org/10.1145/2883851.2883897>
- Porter, G., & Beyerlein, M. (2000). Historic roots of team theory and practice. In M. M. Beyerlein (Ed.), *Work Teams: Past, Present and Future* (pp. 3–24). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-015-9492-9_1
- Praharaj, S., Scheffel, M., Drachsler, H., & Specht, M. (2018). Multimodal analytics for real-time feedback in co-located collaboration. *Proceedings of the European Conference on Technology Enhanced Learning (ECTEL '18)*, 3–6 September 2018, Leeds, UK (pp. 187–201). Springer. https://doi.org/10.1007/978-3-319-98572-5_15
- Prieto-Alvarez, C., Martinez-Maldonado, R., & Buckingham Shum, S. (2018). Mapping learner-data journeys: Evolution of a visual co-design tool. *Proceedings of the ACM Australian Computer-Human Interaction Conference (OzCHI '18)*, 5–7 December 2018, Melbourne, Australia (pp. 205–214). New York: ACM. <https://doi.org/10.1145/3292147.3292168>
- Prieto-Alvarez, C., Martinez-Maldonado, R., & Buckingham Shum, S. (2020). LA-DECK: A card-based learning analytics co-design tool. *Proceedings of the 10th International Learning Analytics and Knowledge Conference (LAK '20)*, 23–27 March 2020, Frankfurt, Germany (pp. 63–72). New York: ACM. <https://doi.org/10.1145/3375462.3375476>
- Prieto, L. P., Rodríguez Triana, M. J., Martínez Maldonado, R., Dimitriadis, Y. A., & Gašević, D. (2018). Orchestrating learning analytics (OrLA): Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level. *Australasian Journal of Educational Technology*, 35(4), 14–33. <https://doi.org/10.14742/ajet.4314>
- Prinsloo, P., & Slade, S. (2016). Student vulnerability, agency and learning analytics: An exploration. *Journal of Learning Analytics*, 3(1), 159–182. <https://doi.org/10.18608/jla.2016.31.10>
- Puntambekar, S., Erkens, G., & Hmelo-Silver, C. (2011). *Analyzing Interactions in CSCL: Methods, Approaches and Issues* (Vol. 12). Boston, MA, USA: Springer Science & Business Media. <https://doi.org/10.1007/978-1-4419-7710-6>
- Reimann, P. (2016). Connecting learning analytics with learning research: The role of design-based research. *Learning: Research and Practice*, 2(2), 130–142. <https://doi.org/10.1080/23735082.2016.1210198>
- Reimann, P., Yacef, K., & Kay, J. (2011). Analyzing collaborative interactions with data mining methods for the benefit of learning. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing Interactions in CSCL: Methods, Approaches and Issues* (pp. 161–185). Boston, MA, USA: Springer. https://doi.org/10.1007/978-1-4419-7710-6_8
- Rodríguez Triana, M. J., Martínez Monés, A., Asensio Pérez, J. I., & Dimitriadis, Y. (2014). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *British Journal of Educational Technology*, 46(2), 330–343. <https://doi.org/10.1111/bjet.12198>
- Rodríguez-Triana, M. J., Prieto, L. P., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2018a). Monitoring Collaborative Learning Activities: Exploring the Differential Value of Collaborative Flow Patterns for Learning Analytics. *Proceedings of the 18th International Conference on Advanced Learning Technologies (ICALT '18)*, 9–13 July 2018, Mumbai, India (pp. 155–159). IEEE. <https://doi.org/10.1109/ICALT.2018.00044>

- Rodriguez Triana, M. J., Prieto, L. P., Martínez-Monés, A., Asensio Pérez, J. I., & Dimitriadis, Y. (2018b). The teacher in the loop: Customizing MMLA in blended CSCL scenarios. *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge (LAK '18)*, 5–9 March 2018, Sydney, Australia (pp. 417–426). New York: ACM. <https://doi.org/10.1145/3170358.3170364>
- Rogers, T., Gašević, D., & Dawson, S. (2016). Learning analytics and the imperative for theory driven research. In C. Haythornthwaite, R. Andrews, J. Fransman, & E. M. Meyers (Eds.), *The SAGE Handbook of E-Learning Research* (pp. 232–250). Thousand Oaks, CA, USA: SAGE Publications Ltd. <http://doi.org/10.4135/9781473955011.n12>
- Rolim, V., Ferreira, R., Lins, R. D., & Gašević, D. (2019). A network-based 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>
- Rosé, C. (2017). Discourse analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of Learning Analytics* (pp. 105–114). Edmonton, AB, Canada: SOLAR. <https://doi.org/10.18608/hla17.009>
- Rosé, C., Wang, Y.-C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., & Fischer, F. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, 3(3), 237–271. <https://doi.org/10.1007/s11412-007-9034-0>
- Rosé, C. P. (2018). Learning analytics in the learning sciences. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (pp. 511–519). New York: Routledge.
- Rummel, N., Walker, E., & Alevin, V. (2016). Different futures of adaptive collaborative learning support. *International Journal of Artificial Intelligence in Education*, 26(2), 784–795. <https://doi.org/10.1007/s40593-016-0102-3>
- Rummel, N., Weinberger, A., Wecker, C., Fischer, F., Meier, A., Voyiatzaki, E., . . . Joshi, M. (2008). New challenges in CSCL: Towards adaptive script support. *Proceedings of the International Conference for the Learning Sciences (ICLS '08)*, 23–28 June 2008, Utrecht, Netherlands (pp. 338–345). ISLS. <https://dl.acm.org/doi/10.5555/1599936.1600036>
- Rupp, A. A., Nugent, R., & Nelson, B. (Eds.). (2012). Special issue on diagnostic measurement in complex learning environments using evidence-centered design. *Journal of Educational Data Mining*, 4, 1–230. <https://jedm.educationaldatamining.org/index.php/JEDM/issue/view/10>
- Salas, E., Sims, D. E., & Burke, S. (2005). Is there a “Big Five” in teamwork? *Small Group Research*, 36(5), 1–46. <https://doi.org/10.1177/1046496405277134>
- Sanders, E. B.-N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *Co-design*, 4(1), 5–18. <https://doi.org/10.1080/15710880701875068>
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 97–115). New York: Cambridge University Press. <https://psycnet.apa.org/record/2006-07157-007>
- Schneider, B. (2020). A methodology for capturing joint visual attention using mobile eye-trackers. *Journal of Visualized Experiments* (155), e60670. <https://doi.org/10.3791/60670>
- Schneider, B., Dich, Y., & Radu, I. (2020). Unpacking the relationship between existing and new measures of physiological synchrony and collaborative learning: A mixed methods study. *International Journal of Computer-Supported Collaborative Learning*, 15, 89–113. <https://doi.org/10.1007/s11412-020-09318-2>
- Schneider, B., & Pea, R. (2017). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *Educational Media and Technology Yearbook*, 40, 99–125. <https://doi.org/10.1007/s11412-013-9181-4>
- Schneider, B., Worsley, M., & Martinez-Maldonado, R. (2021). Gesture and gaze: Multimodal data in dyadic interactions. In U. Cress, C. Rosé, A. Wise, & J. Oshima (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (in press). London, UK: Springer.
- Sergis, S., & Sampson, D. G. (2017). Teaching and learning analytics to support teacher inquiry: A systematic literature review. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundamentals, Applications, and Trends* (pp. 25–63). Basel, Switzerland: Springer. https://doi.org/10.1007/978-3-319-52977-6_2
- Shaffer, D. W. (2017). *Quantitative ethnography*. Madison, WI, USA: Cathcart Press.
- Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, 3(3), 9–45. <https://doi.org/10.18608/jla.2016.33.3>
- Shute, V. J., & Ventura, M. (2013). *Stealth Assessment: Measuring and Supporting Learning in Video Games*. Cambridge, MA, USA: MIT Press. <https://mitpress.mit.edu/books/stealth-assessment>
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review*, 46(5), 30–32. <https://doi.org/10.17471/2499-4324/195>

- Soller, A., Martinez, 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. <https://dl.acm.org/doi/10.5555/1434935.1434937>
- Spikol, D., Ruffaldi, E., & Cukurova, M. (2017a). Using multimodal learning analytics to identify aspects of collaboration in project-based learning. *Proceedings of the International Conference on Computer Supported Learning (CSCL '17)*, 18–22 June 2017, Philadelphia, PA, USA (pp. 263–270). ISLS.
- Spikol, D., Ruffaldi, E., Landolfi, L., & Cukurova, M. (2017b). Estimation of success in collaborative learning based on multimodal learning analytics features. *Proceedings of the International Conference on Advanced Learning Technologies (ICALT '17)*, 3–7 July 2017, Timisoara, Romania (pp. 269–273). <https://doi.org/10.1109/ICALT.2017.122>
- Stadler, M., Herborn, K., Mustafić, M., & Greiff, S. (2020). The assessment of collaborative problem solving in PISA 2015: An investigation of the validity of the PISA 2015 CPS tasks. *Computers & Education*, 157, 103964. <https://doi.org/10.1016/j.compedu.2020.103964>
- Stahl, G. (2004). Building collaborative knowing: Elements of a social theory of CSCL. In J.-W. Strijbos, P. A. Kirschner, & R. L. Martens (Eds.), *What We Know about CSCL: And Implementing It in Higher Education* (pp. 53–85). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/1-4020-7921-4_3
- Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge, MA, USA: MIT Press. <https://mitpress.mit.edu/books/group-cognition>
- Stahl, G. (2013). Theories of collaborative cognition: Foundations for CSCL and CSCW together. In S. P. Goggins, I. Jahnke, & V. Wulf (Eds.), *Computer-Supported Collaborative Learning at the Workplace: CSCL@Work* (pp. 43–63). Boston, MA, USA: Springer. https://doi.org/10.1007/978-1-4614-1740-8_3
- Stahl, G. (2017). *Global Introduction to CSCL*. Philadelphia, PA, USA: Lulu.
- Stahl, G., & Hakkarainen, K. (2021). Theories of CSCL. In U. Cress, C. Rosé, A. Wise, & J. Oshima (Eds.), *International Handbook of Computer Supported Collaborative Learning* (in press). London, UK: Springer.
- Steier, R., Shapiro, B., Christidou, D., Pierroux, P., Davidsen, J., & Hall, R. (2019). Tools and methods for “4E analysis”: New lenses for analyzing interaction in CSCL. *Proceedings of the International Conference on Computer Supported Collaborative Learning (CSCL '19)*, 17–21 June 2019, Lyon, France (pp. 759–766). ISLS.
- Stevens, R., Galloway, T., & Willemsen-Dunlap, A. (2020). Approaches for inserting neurodynamics into the training of healthcare teams. In C. S. Nam (Ed.), *Neuroergonomics: Principles and Practice* (pp. 251–269). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-34784-0_13
- Suthers, D. D. (2017). Multilevel analysis of activity and actors in heterogeneous networked learning environments. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of Learning Analytics* (pp. 189–197). Edmonton, AB, Canada: SOLAR. <https://doi.org/10.18608/hla17.016>
- Suthers, D. D., Dwyer, N., Medina, R., & Vatrappu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. *International Journal of Computer-Supported Collaborative Learning*, 5(1), 5–42. <https://doi.org/10.1007/s11412-009-9081-9>
- Swiecki, Z., Ruis, A., Farrell, C., & Shaffer, D. W. (2020). Assessing individual contributions to collaborative problem solving: A network analysis approach. *Computers in Human Behavior*, 104(March 2020), 105876. <https://doi.org/10.1016/j.chb.2019.01.009>
- Szewkis, E., Nussbaum, M., Rosen, T., Abalos, J., Denardin, F., Caballero, D., . . . Alcoholado, C. (2011). Collaboration within large groups in the classroom. *International Journal of Computer-Supported Collaborative Learning*, 6(4), 561–575. <https://doi.org/10.1007/s11412-011-9123-y>
- Tan, J. P.-L., Yang, S., Koh, E., & Jonathan, C. (2016). Fostering 21st century literacies through a collaborative critical reading and learning analytics environment: User-perceived benefits and problematics. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (LAK '16)*, 25–29 April 2016, Edinburgh, Scotland (pp. 430–434). New York: ACM. <https://doi.org/10.1145/2883851.2883965>
- Tarmazdi, H., Vivian, R., Szabo, C., Falkner, K., & Falkner, N. (2015). Using learning analytics to visualise computer science teamwork. *Proceedings of the Conference on Innovation and Technology in Computer Science Education (ITiCSE '15)*, 6–8 July 2015, Vilnius, Lithuania (pp. 165–170). New York: ACM. <https://doi.org/10.1145/2729094.2742613>
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54. <https://doi.org/10.1177%2F0261927X09351676>
- Tchounikine, P. (2019). Learners’ agency and CSCL technologies: Towards an emancipatory perspective. *International Journal of Computer-Supported Collaborative Learning*, 14(2), 237–250. <https://doi.org/10.1007/s11412-019-09302-5>

- Teasley, S. D. (1997). Talking about reasoning: How important is the peer in peer collaboration? In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools and reasoning: Essays on situated cognition*. Heidelberg, Germany: Springer-Verlag. https://doi.org/10.1007/978-3-662-03362-3_16
- Tempelaar, D. T., Rienties, B., & Nguyen, Q. (2017). Towards actionable learning analytics using dispositions. *IEEE Transactions on Learning Technologies*, 10(1), 6–16. <https://doi.org/10.1109/TLT.2017.2662679>
- Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin*, 63(6), 384–399. <https://psycnet.apa.org/doi/10.1037/h0022100>
- van Leeuwen, A. (2019). Teachers' perceptions of the usability of learning analytics reports in a flipped university course: When and how does information become actionable knowledge? *Educational Technology Research and Development*, 67(5), 1043–1064. <https://doi.org/10.1007/s11423-018-09639-y>
- van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014). Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers and Education*, 79, 28–39. <https://doi.org/10.1016/j.compedu.2014.07.007>
- van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2015). Teacher regulation of multiple computer-supported collaborating groups. *Computers in Human Behavior*, 52, 233–242. <http://doi.org/10.1016/j.chb.2015.05.058>
- van Leeuwen, A., & Rummel, N. (2019). Orchestration tools to support the teacher during student collaboration: A review. *Unterrichtswissenschaft*, 47(2), 143–158. <https://doi.org/10.1007/s42010-019-00052-9>
- van Leeuwen, A., Rummel, N., & van Gog, T. (2019). What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations? *International Journal of Computer-Supported Collaborative Learning*, 14, 261–289. <https://doi.org/10.1007/s11412-019-09299-x>
- van Schuppen, J. H. (2015). What is team theory? In J. H. van Schuppen & T. Villa (Eds.), *Coordination Control of Distributed Systems* (pp. 147–154). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-10407-2_18
- Vrzakova, H., Amon, M. J., Stewart, A., Duran, N. D., & D'Mello, S. K. (2020). Focused or stuck together: multimodal patterns reveal triads' performance in collaborative problem solving. *Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK '20)*, 23–27 March 2020, Frankfurt, Germany (pp. 295–304). New York: ACM. <https://doi.org/10.1145/3375462.3375467>
- Vujovic, M., Tassani, S., & Hernández-Leo, D. (2019). Motion capture as an instrument in multimodal collaborative learning analytics. *Proceedings of the European Conference on Technology Enhanced Learning (ECTEL '19)*, 16–19 September 2019, Delft, Netherlands (pp. 604–608). Springer. https://doi.org/10.1007/978-3-030-29736-7_49
- Wake, J., Wasson, B., Bjørgen, E., & Heimsæter, F. (2019). Supporting firefighter training through visualising indoor positioning, motion and time use data. *Proceedings of the WiPe Paper — Workshop on Advance Technologies for First Responders at International Conference on Information Systems for Crisis Response and Management (ISCRAM '19)*, 19–22 May 2019, Valencia, Spain (pp. 87–90). ISCRAM. <https://hdl.handle.net/11250/2650379>
- Wang, Q., Jin, H., & Liu, Y. (2010). Collaboration analytics: Mining work patterns from collaboration activities. *Proceedings of the International Conference on Information and Knowledge Management (CIKM '10)*, 26–30 October 2010, Toronto, ON, Canada (pp. 1861–1864). New York: ACM. <https://doi.org/10.1145/1871437.1871748>
- West, D., Luzecky, A., Toohey, D., Vanderlelie, J., & Searle, B. (2020). Do academics and university administrators really know better? The ethics of positioning student perspectives in learning analytics. *Australasian Journal of Educational Technology*, 36(2), 60–70. <https://doi.org/10.14742/ajet.4653>
- Wilder, C. R., & Ozgur, C. O. (2015). Business analytics curriculum for undergraduate majors. *INFORMS Transactions on Education*, 15(2), 180–187. <https://doi.org/10.1287/ited.2014.0134>
- Wiley, K. J., Dimitriadis, Y., Bradford, A., & Linn, M. C. (2020). From theory to action: Developing and evaluating learning analytics for learning design. *Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK '20)*, 23–27 March 2020, Frankfurt, Germany (pp. 569–578). New York: ACM. <https://doi.org/10.1145/3375462.3375540>
- Winne, P. H. (2017). Leveraging big data to help each learner and accelerate learning science. *Teachers College Record*, 119(3), 1–24.
- Wise, A., Azevedo, R., Stegmann, K., Malmberg, J., Rose, C. P., Mudrick, N., . . . Fischer, F. (2015). CSCL and learning analytics: Opportunities to support social interaction, self-regulation and socially shared regulation. *Proceedings of the International Conference on Computer Supported Learning (CSCL '15)*, 7–11 June 2015, Gothenburg, Sweden (pp. 607–614). ISLS. <https://repository.isls.org/handle/1/446>
- Wise, A., Knight, S., & Buckingham Shum, S. (2021). Collaborative learning analytics. In U. Cress, C. Rosé, A. Wise, & J. Oshima (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (in press). London, UK: Springer.

- Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. *Proceedings of the Fourth International Conference on Learning Analytics & Knowledge (LAK '14)*, 24–28 March 2014, Indianapolis, USA (pp. 203–211). New York: ACM. <https://doi.org/10.1145/2567574.2567588>
- Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, 12(4), 423–467. <https://doi.org/10.1007/s11412-017-9267-5>
- Wise, A. F., & 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>
- Worsley, M. (2019). Computationally augmented ethnography: Emotion tracking and learning in museum games. *Proceedings of the Advances in Quantitative Ethnography: First International Conference (ICQE '19)*, 19–22 October 2019, Madison, WI, USA (pp. 141–153). Springer Nature. https://doi.org/10.1007/978-3-030-33232-7_12
- Worsley, M., Abrahamson, D., Blikstein, P., Grover, S., Schneider, B., & Tissenbaum, M. (2016). Situating multimodal learning analytics. *Proceedings of the International Conference of the Learning Sciences (ICLS '16)*, 20–24 June 2016, Singapore (pp. 1346–1349). ISLS.
- Worsley, M., & Blikstein, P. (2018). A multimodal analysis of making. *International Journal of Artificial Intelligence in Education*, 28(3), 385–419. <https://doi.org/10.1007/s40593-017-0160-1>
- Wu, L., Waber, B., Aral, S., Brynjolfsson, E., & Pentland, S. (2008). Mining face-to-face interaction networks using sociometric badges: Predicting productivity in an IT configuration task. *Proceedings of the International Conference on Information Systems (ICIS '08)*, 19–20 July 2008, Paris, France (pp. 1–19). AIS/ICIS. <https://aisel.aisnet.org/icis2008/127>
- Xing, W., Wadholm, B., & Goggins, S. (2014). Learning analytics in CSCL with a focus on assessment: An exploratory study of activity theory-informed cluster analysis. *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge (LAK '14)*, 24–28 March 2014, Indianapolis, USA (pp. 59–67). New York: ACM. <https://doi.org/10.1145/2567574.2567587>