International Handbook of Computer-Supported Collaborative Learning

Preprint — **Wise, A., Knight, S., Buckingham Shum, S. (In Press) Collaborative Learning Analytics. In: Cress, U., Rosé C., Wise, A. & Oshima, J. (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (Springer)**

Collaborative Learning Analytics

Alyssa Friend Wise[1], Simon Knight[2], Simon Buckingham Shum[3]

[1] New York University, Steinhardt School of Culture, Education, and Human Development, New York, US, Alyssa.Wise@nyu.edu.au

[2] University of Technology Sydney, Faculty of Transdisciplinary Innovation, Sydney, Australia, Simon.Knight@uts.edu.au

[3] University of Technology Sydney, Connected Intelligence Centre, Sydney, Australia, Simon.BuckinghamShum@uts.edu.au

Abstract

The use of data from computer-based learning environments has been a longstanding feature of CSCL. Learning analytics can enrich this established work in CSCL. This chapter outlines synergies and tensions between the two fields. Drawing on examples, we discuss established work to use learning analytics as a research tool (analytics of collaborative learning – ACL). Beyond this potential though, we discuss the use of analytics as a mediational tool in CSCL – Collaborative Learning Analytics (CLA). This shift raises important challenges regarding the role of the computer – and analytics –in supporting and developing human agency and learning. LA offers a new tool for CSCL research. CSCL offers important contemporary perspectives on learning for a knowledge society, and as such is an important site of action for learning analytics research that both builds our understanding of collaborative learning, and support that learning.

Introduction, Definitions & Scope

There are two different histories that can be told about the relationship between Learning Analytics (LA) and Computer-Supported Collaborative Learning (CSCL). One is a story of continuity, in which the two fields have approached each other through a natural convergence of aligned interests. In this story, the promise of automated analyses to create dynamic support for collaboration has always been a part of the vision for CSCL, developing from early ideas about the relevance of machine learning for supporting collaboration made by Dillenbourg in 2005 (Rosé, 2018), to the creation of group awareness tools (Bodemer & Dehler, 2011), adaptive scripts (Vogel, Weinberger, & Fischer, 2020) and other intelligent support for groups (Kumar &

Kim, 2014). In parallel, the use of analytics to support groups of learners in their interactions has been a growing theme of work in LA since the inception of the field (Gašević, Dawson, Mirriahi, & Long, 2015), with attention paid both to carefully designed small group collaboration in formal learning environments (as commonly studied in CSCL) as well as large networks of people in informal settings (subject to less control by researchers/educators). This work has often been considered under the umbrella of Social Learning Analytics (Buckingham Shum & Ferguson, 2012), including attention to participant relationships (Bakharia & Dawson, 2011), discourse (De Liddo, Buckingham Shum, Quinto, Bachler & Cannavacciulo, 2011), sequential patterns of interaction (Suthers & Rosen, 2011) and communities (Haythornthwaite, 2011).

While this story of productive convergence is compelling, there is also a second, somewhat less harmonious, history that can be told. In this story, one of the critical characteristics of CSCL as a field is the study of collaborative processes as unique constructions achieved by (primarily small) groups, where units of discourse come to have meaning in an ever-evolving mediated context. The importance of developing theories to make sense of the local meaning-making that emerges through participation in specific situations has always stood in tension with more quantitative approaches to the study of CSCL (in which learning analytics is now included). Thus the rise of analytic approaches that attend *only* to quantitative representations of collaboration can be met with skepticism as a productive route to understanding. Similarly, there are concerns that quantification from the 'bottom up' (for example through structure discovery methods such as topic modelling) without attention to existing theory (or the generation of new theory) will not help to advance the collective knowledge base about collaborative learning[1]. How, then, can the CSCL and LA communities work together to develop collaborative learning analytics in ways that hold true to the collective values of the field? In the following sections we address this question by unpacking two distinct but complementary appeals that LA holds for the field of CSCL: first, as a set of *methods* useful to better *understand* collaborative learning; and second as a set of *tools* useful to better *facilitate* it. In doing so, the chapter aims to provide both an overview of the history of *analytics of collaborative learning* (to generate understanding of CSCL), and give signposts to recent and emerging work to create *collaborative learning analytics* (to create support for CSCL). Finally, we ask critical questions for both the CSCL and LA communities to consider about how the area of collaborative learning analytics (CLA) should develop.

History & Development: Analytics of Collaborative Learning

The core challenge in using learning analytics to better understand collaborative learning is to conceive (conceptually) and implement (technically) connections between (1) fine-grained trace data of the sort captured in software logs and (2) learning constructs (see Figure 1). How can this bridge from "clicks to constructs" (Knight, Buckigham Shum, & Littleton, 2014) be built? The

[1] Rosé (2018) discusses many of the same tensions existing between LA and the Learning Sciences more broadly: for example the need to consider the relative value of model accuracy versus interpretability, and top-down (theory-driven) versus bottom-up (data-driven) approaches. A key differentiator for CSCL in addressing these tensions is a longstanding history of considering the role of computers and computation in learning, which has been a central part of the fiber of the CSCL community from the beginning.

shaded triangles at the foot of Figure 1 reflect what we might think of as the traditional strengths of CSCL (anchored to the left) and LA (anchored to the right). CSCL is theoretically robust in its definition of learning constructs and investigation of them through careful manual qualitative and quantitative analysis of data. LA has arisen from data science and analytics, working with large amounts of machine-generated information. It is methodologically strong in using computational tools and techniques to mine insights from low-level trace data captured from online platforms, mobile computers and environmental / physiological sensors used 'in the wild' at scale. Bringing these strengths together offers the possibility to engineer complex higher order data features that richly represent meaningful learning constructs. Examination of such features can both offer direct insight into collaborative learning processes from a quantitative perspective and provide indicators of 'where to look' for in-depth qualitative examinations in large datasets. The potential for mutual enrichment across CSCL and LA is significant, therefore, if rigorous mappings can be designed and implemented.

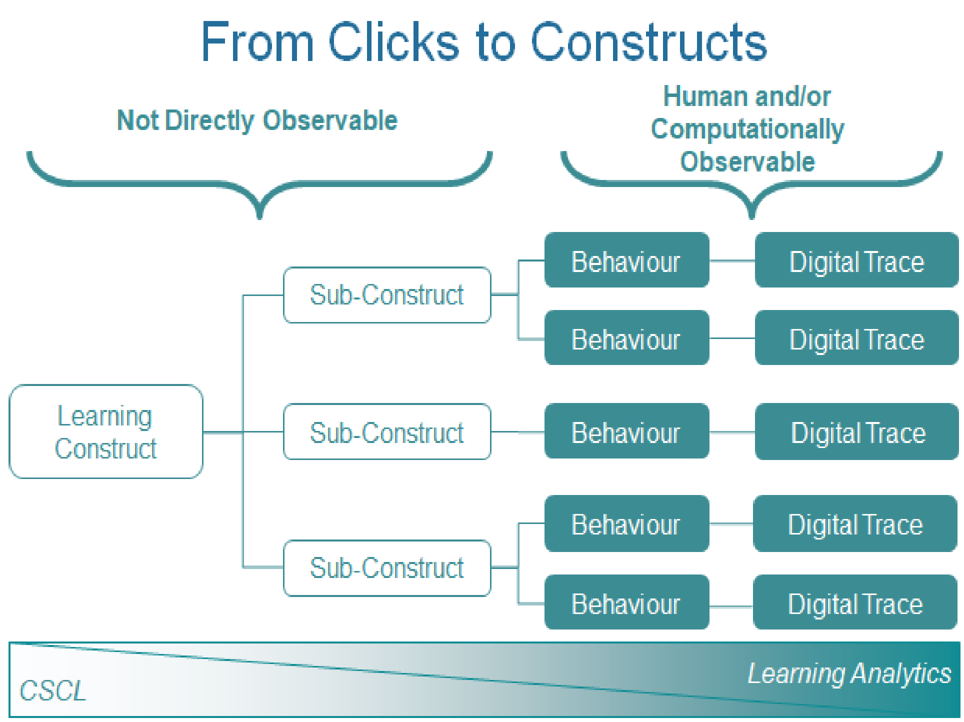

Fig. 1 A core challenge for analytics of collaborative learning is to map digital traces to learning constructs

In the case of studying argumentative knowledge construction (see e.g. Weinberger & Fischer, 2006) or in a collaborative tool designed for argumentation such as Argunaut (McLaren, Scheuer, & Mikšátko, 2010), theory might direct us to consider (1) *distribution of the discourse* (i.e., it is not argumentation if only one person engages); (2) *responsiveness* (i.e., that contributions made connect to one another); (3) the use of *formal reasoning* (e.g. evidence, warrants and qualifications); or (4) employing particular *argumentation strategies* (e.g. 'argument by analogy'). Each of these constructs can then be mapped onto *human observable behaviors*, such as turn-taking (as a window to discourse distribution), or the use of argumentative elements (as evidence of formal reasoning). For example, adding a node in an argument graph might be one indicator of taking a turn, while the presence of sources, reasons or epistemic modals in the node's text (detected using natural language processing) might be an indicator of the use of evidence in communication. Similarly, detecting when a student moves an argument node in a way that substantively affects the arguments could offer another indicator of taking a turn.

These kinds of analysis sit well in established CSCL work. For example, as Figure 2 indicates, in research to create analytics of relations among collaborative contributions, Suthers, Dwyer, Medina, and Vatrapu (2010) operationalized the construct of *uptake* between learners using evidence mapping and threaded discussion tools, based on the notion of contingency constructed from a combination of semantic relatedness between peers' contributions (involving many derived linguistic features), media dependency (taking action on a peer's object), temporal proximity (happening close together in time) as well as spatial organization and inscriptional similarity. LA amplifies such possibilities for analytically constructing complex features from digital traces. In the following sections we unpack in more detail further examples of work in this space, as summarised in Figure 2.

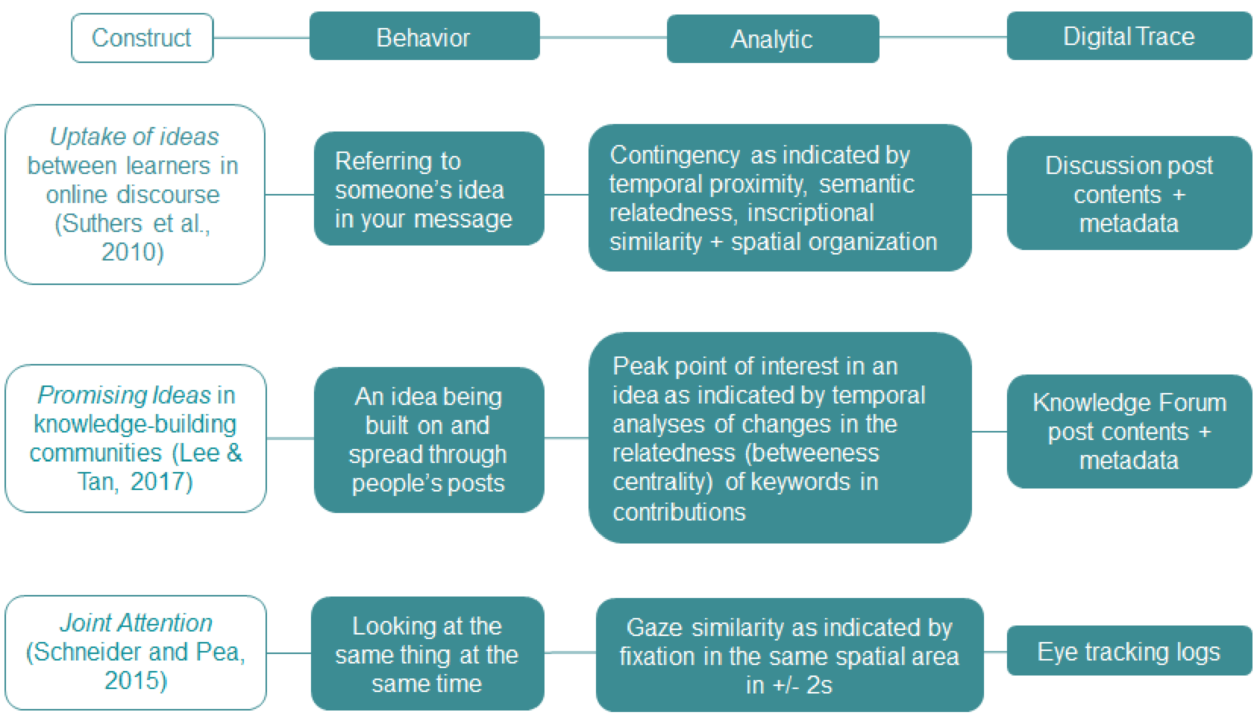

Fig. 2 Research on analytics of collaborative learning seeks to bridge "from clicks to constructs" by creating construct-aligned analytics, derived from digital trace

Examples of How Analytics of Collaborative Learning Further Understanding of CSCL

Analytics of Collaborative Knowledge Building

In the second row of Figure 2 we introduced the example of operationalizing the construct of *promising ideas* in knowledge-building discourse. This is a good example of how research, and the student/educator experience, has evolved in an established CSCL software tool and research platform (Knowledge Forum). Knowledge Building as a strand of CSCL research has focused on the use of graphical networks of nodes and links (enabled with hypertext software) as mediating representations for student discourse. These notations structure discourse by providing a vocabulary of contributions and relationships that focus learners' attention on the conceptual structure of the conversation: explicitly created links signal discourse relationships, and node types signal further the intended nature of contribution. The advantage over conventional contributions to a flat or threaded discussion forum is that such representations both provide distinct visual and classification affordances for reasoning by humans and are more tractable computationally. Other well-known CSCL tools for structuring thinking and argument include Belvedere (Suthers, Weiner, Connelly, & Paolucci, 1995) and Argunaut (McLaren, et al. 2010); for an overview of these and diverse other forms of tools to support structured discourse and deliberation, see Kirschner, Buckingham Shum, and Carr (2003) and Okada, Buckingham Shum, and Sherborne (2014).

Early research analyses of Knowledge Forum data used conventional techniques such as statistical summaries of how networks grew, and human, qualitative coding of contributions (Scardamalia, 2003). More recent analyses have seen the introduction of text analytics approaches to provide new proxies for constructs such as *similarity of contributions* (using Latent Semantic Analysis: Teplovs & Fujita, 2013), and *productive threads* and *improvable threads* (Chen, Resendes, Chai, & Hong, 2017). Threads coded as *productive* or *improvable* by human analysts were interrogated using novel (to CSCL) analytic techniques, such as Lag-Sequential Analysis, "to determine whether there is cross-dependence between a specified behavior and another behavior that occurs earlier in time" (p. 228). This revealed which metrics were powerful enough to differentiate the two thread types (*productive* versus *improvable*), yielding new insights into the temporal dimensions of each kind of discourse, for example, fine grained claims such as: "...productive threads of inquiry involved significantly more transitions among *questioning, theorizing, obtaining information*, and *working with information*, while improvable inquiry threads showed more transitions involving *giving opinions*" (p. 229).

In recent work Lee and Tan (2017) demonstrated how the construct of a *promising idea* (one that helps to move a knowledge-building community forward: Chen, 2017) can be operationalized in the Knowledge Forum software tool through the combination of text mining (to reveal relationships between keywords in students' contributions) with network analysis (using the metrics of betweenness, centrality and degree centrality on the text mining results to identify strongly connected ideas). These metrics were designed to "reveal the associations between keywords, discourse units, and participants within a discourse" (p.81). Coupled with visualizations, Lee and Tan argue that these are a form of visual analytic for the quality of *"promisingness"* that is a hallmark of a strong discourse contribution.

Analytics of Joint Attention

Analytics of collaborative knowledge building evolved through innovating the analysis applied to existing data sources from an established CSCL software tool. Opportunities are also emerging to generate analytics based on new sources of data. To illustrate, this example details how eye-tracking has facilitated the examination of *joint attention*, a central concept in CSCL (Tomasello, 1995). Joint attention is theorized as an important mechanism by which shared focus on a common reference helps collaborators coordinate with one another to ground communication (Clark & Brennan, 1991) but has traditionally been difficult to empirically assess. Considering (computationally detectable) gaze as an indicator of visual attention (which can be taken as a reasonable proxy for attention more generally), researchers have used dual eye-tracking systems to construct measures of joint visual attention, operationalized as the composite features of gaze similarity or cross-recurrence (Jermann, Mullins, Nüssli, & Dillenbourg,, 2011; Schneider & Pea, 2013; Sharma, Caballero, Verma, Jermann, & Dillenbourg, 2015). These measures have been useful in investigating the role of joint attention in CSCL in ways not previously possible.

For example, Schneider and Pea (2013) found that an intervention designed to support joint visual attention in dyads collaborating at a distance (by making participants' gaze visible to each other) achieved its goal, and that greater joint visual attention correlated with higher quality collaborative processes and improved learning outcomes. Going beyond the simple notion that more joint visual attention is better for collaboration, Schneider and Pea (2015) sought to examine its relationship with the related concept of *transactive discourse* (Stahl, 2013). They examined the association of moments of relatively high and low joint visual attention in the previous study, with patterns in the coherence of talk (a derived feature constructed from the computationally detectable words used via a sliding time-window analytic of cosine similarity). This allowed them to identify different potential modes of collaboration. For example an exchange when both joint visual attention and verbal coherence was high was spatially focused, with referents used to anchor the dialogue (e.g. "this one right there"), while an exchange with high coherence but low joint visual attention illustrated an attempt to integrate information across the task.

In a separate study of collocated collaboration using a tangible tabletop, Schneider et al. (2015) found that using three-dimensional representations of shelves in a warehouse optimization task led to greater joint visual attention than two-dimensional representations, and that the greater overall levels of joint visual attention were associated with higher task performance, and in some cases, learning outcomes and quality of collaborative processes. Deeper examination of the eye-tracking data revealed a critical aspect of group dynamics: groups in which learners equally shared responsibility for initiating joint visual attention showed higher learning gains than those in which joint visual attention was primarily initiated by only one of the learning partners (Schneider et al., 2016). This second example illustrates how the technical capabilities of analytics (ability to digitally capture gaze by microsecond) can also support the generation of new conceptual categories (construction of the derived feature of joint attention initiator opened the possibility for the construct of a *leading learner*). Additionally, in both this set of studies and the ones described above, there was evidence that the more successful dyads moved fluidly between different (spatial) parts of the problem to be solved, rather than focusing on them serially (Schneider & Pea, 2013; Schneider et al., 2016). Together, these studies advance our understanding of joint attention by identifying different patterns of partner gaze that can occur, examining how they do (or do not) contribute to learning, and testing designs through which they can be supported.

From Understanding to Action

The above examples showed how researchers have worked to connect clicks to constructs in order to develop meaningful insight into collaborative learning processes. Once valid features and metrics have been built to serve as proxies for pedagogically significant constructs, it may then be possible for an analytics infrastructure to partially or fully automate that analytical workflow. This makes it possible to go beyond tools that support researchers to create tools that support students and educators by providing meaningful representations of collaborative activity patterns in a timely manner. Such *Collaborative Learning Analytics* create a feedback loop by generating information than can trigger computer-initiated adaptations to the conditions of collaboration or be provided to students and educators to provoke reflection, and potentially, changes that improve collaborative learning.

State-of-the-Art: Collaborative Learning Analytics

Recent years have seen a shift in LA, from a tool for researchers to understand the processes and learning impacts of CSCL (*analytics of collaborative learning: ACL*), to a way to support CSCL (*collaborative learning analytics: CLA*). This new emphasis is marked by a focus on designing learner- and instructor-facing analytics that provide timely feedback on collaborative learning process. In this way, CLA involves treating the outputs of analyzing collaborative learning (ACL) as inputs to improve collaboration quality. Consequently, a key challenge in translating “ACL” to “CLA” are human-computer interaction and human factors considerations; for instance how to usefully present analytic information as feedback on collaborative learning processes and how to support people in interacting with, comprehending and taking action on that feedback. In this way the developing area of CLA has significant potential to become a core mediating tool to support collaborative learning.

This broad idea is not a new one in CSCL. For example, in 2007 the notion of group mirrors was introduced (Dillenbourg & Fischer, 2007), while a 2011 special issue of Computers in Human Behavior (Bodemar & Dehler, 2011) built on earlier efforts in the related field of computer supported collaborative work (CSCW) to address the question of group awareness in CSCL environments. In this work, devices such as interactive tabletops and other large displays were used to display back information to groups about their interactions (in the simplest example, the number of times each participant speaks) with the goal of encouraging reflection in real time (Jermann & Dillenbourg, 2008). LA extends and expands this tradition of showing collaborating learners information about their process by capturing and analysing data in richer, more complex ways.

CLA provides potential for the CSCL community to design new kinds of support for collaborative learning which are able to impose less structure a priori by providing ongoing assistance through real-time, theory-grounded, scalable feedback. This addresses a concern within the CSCL community that some of the original emancipatory spirit of the field is lost when the implementation of effective CSCL requires strong scripts that limit the range of interaction possibilities (Wise & Schwartz, 2017). The emergence of LA offers the opportunity to address this issue by developing learning environments in which the processes of interaction with computer support are less tightly predefined, with the system instead acting responsively to the learners and their interactions.

Such responsiveness also has the potential to extend the sophistication of CSCL scripts from relatively fixed templates for learning interactions to dynamic models in which conditions and patterns of collaboration are adjusted and calibrated responsively before and during a learning episode. For example, this might include the customisation of scripts for particular groups of learners or group-interaction-dynamics. It might also involve the use of scripts that – to use Fischer, Kollar, Stegmann and Wecker's (2013) term – fade, or adapt over time, adjusting the nature of the structures put in place to support learning. In doing so, there is also the potential to advance our knowledge of collaboration and the systems we develop to support it from being relatively domain-general (e.g. *"assigning roles to learners is helpful for collaboration"*) to one in which the support (and related knowledge claims) is more tightly specified to the people and learning tasks involved (e.g. *"assigning roles that take distinct perspectives may be helpful for this task since the goal is broad idea generation and your group tends to arrive at quick consensus..."*). In summary CLA offer the potential for a new chapter of CSCL research; however it must navigate several key tensions to do so.

Changing the Shape of Support for Collaborative Learning through CLA

Collaborative Learning Analytics (CLA) is an emerging area of research and practice, which builds on prior CSCL work to make a distinctive contribution to the nature of computer support for collaborative learning. The final piece of the technical puzzle is the growing development of systems and infrastructure to "close the loop" in real-time; that is, to automate the capture, analysis, and feeding-back of information about collaborative learning processes to the collaboration while it is still in progress. Increasingly, marking the shift to CLA, we thus see a focus on new questions that arise in this final phase with respect to the design features of how learning analytics information can be integrated and implemented in practical learning contexts to inform learning. In so doing, CLA must navigate three core tensions which we discuss in detail below:

1. What will CLA do? The relative balance of technology and human agency
2. Who will CLA attend to? Support for activity at different levels (group, individual, collective)
3. How will CLA operate? Iterations of refining collaborative learning efforts

What will CLA Do? The Relative Balance of Technology and Human Agency

A key consideration for CLA is the mode of action for the computer and the agency of the actors in a particular learning context. CSCL research has a long history of attention to the agency of human (particularly student) actors; as stated by Scardamalia et al. (1989), "the computer environment should not be providing the knowledge and intelligence to guide learning, it should be providing the facilitating structure and tools that enable students to make maximum use of their own intelligence and knowledge" (p 54). Attention to human agency has also been an area of focus in LA. Drawing on precedents in a range of fields that seek to keep the human 'in the loop' when working with intelligent machines, Kitto, Buckingham Shum & Gibson (2018) have argued that knowing when to *disagree* with analytics (and being empowered to do so) is both an important competence to build, and an effective pedagogic strategy. Deriving from this, in deploying CLA, attention must be paid to the ways in which, through flexible implementation, LA can be used to *support* rather than *supplant* the agency of learners. Consideration must also be given to the extent to which analytics are seen as a *temporary* scaffold for collaborative learning whose role will eventually be taken over and internalized by learners, as compared to a performance support system which will *continually* provide data to inform collaboration on an ongoing basis. To consider these questions, we explore two different routes to building CLA: (1) via *Adaptive CSCL systems*, in which changes to collaboration based on analytics are algorithmically initiated; and (2) *Adaptable CSCL systems*, in which changes to collaboration based on analytics are initiated by the users (Wise, 2019). *Adaptive CSCL* assigns a great degree of agency to the computer to adjust learning tasks and content for specific tasks, potentially with input (or a veto from) learners or educators. In contrast, *Adaptable CSCL* is designed to promote reflection on the part of learners and instructors and support action based on that reflection.

Adaptive CSCL

In Adaptive CSCL systems, tools are designed to algorithmically alter learning environments in response to data during a learning episode. More specifically, the aim of adaptive CSCL is to use "*intelligent technologies to improve student collaboration and learning by assessing the current state of the interaction and providing a tailored pedagogical intervention (Soller et al. 2005)"* Rummel, Walker, and Aleven (2016, p.785), pointing to the potential of ongoing work at the intersection of learning analytics and artificial intelligence in education. This support may target the ways in which groups are configured and formed (e.g., the characteristics of members and their roles), the nature of group interaction, or the nature of the group's understanding (Magnisalis, Demetriadis, & Karakostas, 2011). For example Howley, et al. (2012) investigated support for group composition and interaction, by looking at how an intelligent dialogue system influenced interaction in groups with varying self-efficacy compositions. Similarly, Walker, Rummel, and Koedinger (2011) explored the impact of adaptive support on helping behaviors in peer tutoring. Indeed, there is an emerging body of work on *Intelligent Support for Learning in Groups,* exemplified by a special issue of the International Journal of Artificial Intelligence in Education (Kumar & Kim, 2014) as well as a series of workshops at the AIED and ITS conferences by the same name, and recent CSCL group formation symposium with a similar theme (Tsovaltzi et al., 2019).

While our default assumption might be that the primary agent in such systems is the computer itself, alternative adaptive models have been developed in which humans retain more control. Rummel et al. (2016) directly address concerns about a "dystopian" future in which artificial intelligence support for collaboration is reactive, rigid, and robs learners (and teachers) of agency, by describing a vision for a more "utopian" vision in which support is provided in a responsive, nuanced and flexible way. In this vision, the adaptive agent supports development grounded in the learning sciences, with agency to both educators and students, through the use of explainable models. Thus, we can imagine a continuum of systems from those in which the adaptive system drives adaptation in the learning context, to ones in which agents retain ability to adjust, 'speak back to', or ignore adaptive features, blurring into the kinds of adaptable systems we now turn to.

Adaptable CSCL

In Adaptable CSCL, collaborative learning processes are made visible for reflection by educators and learners so that they can adjust their learning interactions or the learning environment itself. Adaptable CSCL systems commonly deploy analytics (that were previously used in research contexts) to display information to students and educators about their collaborative interactions for reflection (see Liu & Nesbit, 2019 for a recent review of CSCL dashboards). In the future, adaptable CSCL may expand to to coaching systems that also offer support in interpreting and actioning on these representations of interaction processes (see e.g. Soller, Jermann, Mühlenbrock, & Martinez, 2005).

Adaptable CSCL systems can include analytics *embedded* as part of the interface used for collaboration to support reflection-in-action, or *extracted* from it to support reflection-on-action (Wise, Zhao, & Hausknecht, 2014). For example, the Starburst discussion forum tool was built to embed analytics by visualizing the online conversation as a hyperbolic tree in which the size and color of the nodes communicates information to students about the structure of the discussion, the location of their comments within it, whether contributions are receiving replies and if threads are being abandoned or ignored (Wise et al., 2014). Separately, log-file trace data about students' speaking and listening activity was extracted from the system and provided to students in a separate (digital) space, where they were asked to "step back" from the action and reflect on the group's collaborative process and their role in it. Other examples of embedded CLA include various visual representations of interaction designed to maintain group awareness during collaboration (e.g. Bodemer & Dehler, 2011), while the classic example of extracted CLA is an analytics dashboard whose use is separated from the collaboration itself in time and/or space (e.g. van Leeuwen, 2015; Tan, Koh, Jonathan, & Yang, 2017).

Recent work in multimodal analytics has also sought to combine real-time temporally and spatially embedded analytics, with extracted analytics in the form of post-hoc visualisation for the purposes of reflection. For example, as described in *Gesture and Gaze: Multimodal Data in CSCL* (Schneider, Worsely, & Martinez-Maldonado, 2020), visualisations can be used to support educator agency in real time to indicate how they move through collaborative learning spaces, and make suggestions for groups that might require more attention. After sessions, then, the same technology can be used to support both individual students, and collaborative groups, through their exploration of extracted analytics that can be used to show key events for individuals in a teamwork context, against an idealised process (see Echeverria, Martinez-Maldonado, & Buckingham Shum, 2019). Indeed, tools are similarly emerging that shift from analytics of CSCL in the context of the knowledge building construct described earlier, to CLA for that construct. For example Tan et al. (2017) and Tan, Koh, Jonathan, and Tay (2018), discuss the use of learning analytics dashboards to support pedagogic adaptation by a teacher with their students in the context of complex literacy practices. In summary, Adaptable CSCL are an active area of research and innovation in which a variety of systems are being built and tested.

Who will CLA Attend To? Support for activity at different levels

The level(s) of analysis for *studying* collaborative learning is a long-standing area of consideration in CSCL, with different work taking individual learners, small groups and large collectives as the unit-of-analysis (as well as tackling the challenging question of how to make claims that bridge across these different levels, Stahl et al., 2013). The introduction of CLA brings an additional layer of complexity as the goal is no longer simply understanding collaborative learning from a group, individual or collective perspective, but also acting on it. This is an important issue as learning analytics more broadly have largely stayed focused on the individual as the “target” for analytic insight and resultant action, but for CLA action at multiple levels may be needed. For example, we can imagine a CSCL system in which core concerns might include feedback that prompts reflections from “how is *my* contribution to the collaboration and what might *I* do to improve it?” to “how is *our* collaboration going and what can *we* do to improve it?” Drawing on work from the field of self-regulation, these can helpfully be distinguished by considering whether the goal is to support self or socially-shared regulation (Wise et al., 2015). Where computer-action supports collaborative learning the question arises of how the intervention or changes to the conditions of collaboration are intended to impact learners as individuals and/or at the level of the unified group.

How will CLA operate? Iterations of refining collaborative learning efforts

The final core question for CLA relates to their theory of action: How are CLA expected to impact collaboration and how can this process can be designed for? CLA emphasizes *improvement to collaborative processes and learning* as the outcome measure; thus, our attention shifts from accuracy of representations of theoretical constructs to considering the audience(s) for action and means by which such action will occur (Wise, Knight, & Ochoa, 2018). From this perspective, while analytics may be imperfect, they can still provide useful insights to learners and instructors with appropriate design (Kitto et al., 2018).

To support learning, CLA must be embedded into practical learning contexts in ways that support thoughtful interpretation of, and action on, the collaborative learning interactions that occur within them. This embedding relates to guidance surrounding the implementation of learning analytics generally, such as for their use to be incorporated as an integral part of the learning experience and for reference points to be provided such that users have a ready way to evaluate the meaning of the information and representations provided (Wise & Vytasek, 2017). In addition, a key concern in implementing CLA is that for LA to be effective learners need both to recognise when they are experiencing a learning problem (perhaps prompted by the analytic), and to know how to address the problem - a translation requiring a conceptual leap from the analytic information (Wise et al., 2014).

There are also additional concerns specific to CLA. For example, van Leeuwen and colleagues conducted a program of research examining specific ways in which CLA are useful to instructors in monitoring and supporting collaborating groups. They found evidence to suggest that while CLA may or may not improve teacher's ability to *notice* problems in collaboration, they did increase their *specificity* of diagnosis and their *confidence* in interpretation, leading them to offer more support to the groups (van Leeuwen, Janssen, Erkens, & Brekelmans, 2014; van Leeuwen 2015; van Leeuwen, Janssen, Erkens, & Brekelmans, 2015). While this work examines how instructors work with CLA to improve collaborative learning, other models of action address the ways in which students work with CLA (Wise, Vytasek, Hausknecht & Zhao, 2016) or how students and teachers can come together to make sense of and act on CLA (Tan et al., 2017; 2018). Further documentation of the mechanisms by which CLA can support and impact collaborative learning is an active area of research that can also contribute to the effective design of such systems.

Conclusions & Future Directions

This chapter has identified some of the key tensions at the intersection of CSCL and Learning Analytics, and introduced exemplars that demonstrate how these have been — and looking to the future could be — resolved in productive ways. Spanning a rich variety of learning contexts, the potential of log-file data mining, natural language processing and multimodal analytics to support online and collocated CSCL is clear. In this chapter, we have foregrounded the challenge of grounding analytics in CSCL constructs in a principled way, and identified the distribution of *agency* between learners, educators and computers as a key design consideration. We have argued that Learning Analytics can provide a powerful new capability in the CSCL toolbox, firstly, by yielding new insights when deployed as *Analytics of Collaborative Learning*, and secondly, deployed as *Collaborative Learning Analytics,* directly supporting learners and educators as they engage in CSCL.

The potential of digital trace data to inform our understanding of collaborative learning has been of longstanding interest to the CSCL community. The history of questions about how these new techniques can and should be applied and used is similarly extensive. Discussing the role of 'Computer' in Computer-Supported Collaborative Learning over 25 years ago, Bannon (1994) lists a number of ways we might understand the potential of computers for learning:

1. As a tool for researchers, to gather data for analysis: "the computer makes the task of the researcher easier but does not really affect the collaborative learning process per se" (p.4).
2. As a platform or 'rich microworld' in which students can interact (p.4)
3. As an automated tutoring tool with which the student interacts or collaborates (pp. 4-5)

4. As a resource to support collaborative learning (the viewpoint he argues for): “The computer can help students to communicate and collaborate on joint activities, providing assistance in the coordination process. This mediational role of the technology emphasizes the possibilities of using the computer not simply as an individual tool but as a medium through which individuals and groups can collaborate with others. In such studies the computer acts as a support and resource for the collaborating students.” (p.5)

Substituting *Analytics* for *Computer*, at a basic level, learning analytics is a tool that can augment research through the collection and analysis of data; the first of these possibilities. There is well established work that has deployed learning analytic techniques as a research tool to understand the processes of learning in CSCL contexts, often making use of technological affordances that support student interaction (on and offline; the second of the possibilities). Analytics of Collaborative Learning (ACL) raise new kinds of challenges, including:

1. How to conduct research in this interdisciplinary space, requiring bringing experts in data mining, learning analytics, education, CSCL, and more together in productive dialogue
2. The relationship between theory and data (and its analysis) in CLA
3. The specific object of the analytics, for example the group or individual; or the episode or idea
4. How to deal with new kinds of data in this context (multimodal, textual, and so on), and particularly the practical challenges of interoperability of such data across CLA systems and contexts.

Moreover, in the second part of this chapter we have pointed to the potential of learning analytics to support the fourth potential raised by Bannon: of an emerging approach to what we have called Collaborative Learning Analytics. In this view analytics can act as a computer support for collaborative learning. Such a potential allows us as researchers to use analytics as a tool-to-think with, to instantiate and test theoretical notions about what matters for collaboration by creating analytics that are a part of the collaboration rather than the environment in which it occurs. This novel potential raises important challenges for CLA:

1. How can CLA support and develop human agency?
2. Who is the audience for CLA? (Groups, individuals and/or collectives; Students and/or teachers)
3. What is the intended impact of CLA (and how do we evaluate that impact)?
4. How do we design for impact, respecting the needs to integrate and implement CLA in practical learning contexts?
5. How can CSCL have impact in a context where, often commercial, vendors are rapidly developing products?

This chapter introduces a set of exemplifications of research being pursued at the intersection of CSCL and learning analytics; there are, of course, many more that could not be discussed. The potential of log-file data mining, natural language processing and sensors that support multimodal learning analytics for CSCL is clear, across a range of learning contexts. We have suggested that learning analytics provides a new tool in the CSCL toolbox. Moreover, that CSCL, which offers contemporary perspectives on learning for a knowledge society, is a specific and important site of action for learning analytics research, to create CLA, that both build our understanding of collaborative learning, and support that learning.

References

Bakharia, A., & Dawson, S. (2011, February). SNAPP: a bird's-eye view of temporal participant interaction. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 168-173). ACM.

Bannon, L. (1994). Issues in computer supported collaborative learning. In C. O'Malley (Ed.), Computer supported collaborative learning. *Proceedings of NATO advanced research workshop on computer supported collaborative learning*, Aquafredda di Maratea, Italy , Sept. 24 - 28, 1989. NATO ASI Series, SERS F Vol.128. Berlin: Springer - Verlag. ISBN 3 - 540 - 57740 - 8.

Bodemer, D., & Dehler, J. (2011). Group awareness in CSCL environments. *Computers in Human Behavior*, *27*(3), 1043-1045. https://doi.org/10.1016/j.chb.2010.07.014

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

Chen, B. (2017). Fostering scientific understanding and epistemic beliefs through judgments of promisingness. *Educational Technology Research and Development, 65*(2), 255–277. http://dx.doi.org/10.1007/s11423-016-9467-0

Chen, B., Resendes, M., Chai, C. S., & Hong, H.-Y. (2017). Two tales of time: Uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interactive Learning Environments*, *25*(2), 162–175. https://doi.org/10.1080/10494820.2016.1276081

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

De Liddo, A., Shum, S. B., Quinto, I., Bachler, M., & Cannavacciuolo, L. (2011). Discourse-centric learning analytics. *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 23-33). ACM.

Dillenbourg, P., & Fischer, F. (2007). Computer-supported collaborative learning: The basics. *Zeitschrift für Berufs-und Wirtschaftspädagogik*, *21*, 111-130.

Echeverria, V., Martinez-Maldonado, R., & Buckingham Shum, S. (2019). Towards collaboration translucence: Giving meaning to multimodal group data. *Proceedings of ACM CHI conference (CHI'19)*, Glasgow, UK. (ACM: New York). https://doi.org/10.1145/3290605.3300269

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, *48*(1), 56-66.

Gasevic, D., Dawson, S., Mirriahi, N., & Long, P. D. (2015). Learning analytics–A growing field and community engagement (editorial). *Journal of Learning Analytics*, *2*(1), 1-6.

Haythornthwaite, C. (2011). Learning networks, crowds and communities. *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 18-22). ACM.

Howley, I., Adamson, D., Dyke, G., Mayfield, E., Beuth, J., & Rosé C. P. (2012). Group composition and intelligent dialogue tutors for impacting students' academic self-efficacy. In: S. A. Cerri, W. J. Clancey, G. Papadourakis, K. Panourgia (Eds.) *Intelligent tutoring systems*: *Lecture notes in computer science*, Vol 7315. Springer, Berlin, Heidelberg

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

Jermann, P., Mullins, D., Nüssli, M.-A., & Dillenbourg, P. (2011). Collaborative gaze footprints: Correlates of interaction quality. *CSCL2011 Conference Proceedings., Volume I - Long Papers*, 184–191.

Kitto, K., Buckingham Shum, S. & Gibson, A. (2018). Embracing imperfection in learning analytics. *Proceedings of LAK18: International conference on learning analytics & knowledge*, March 5–9, 2018, Sydney, AUS, pp.451-460. (ACM, New York, NY, USA). https://doi.org/10.1145/3170358.3170413

Kirschner, P.A., Buckingham Shum, S.J. & Carr, C.S. (Eds.) (2003). *Visualizing argumentation: Software tools for collaborative and educational sense-making.* London, UK: Springer.

Knight, S., Buckingham Shum, S., & Littleton, K. (2014). Epistemology, Pedagogy, Assessment: Where Learning Meets Analytics in the Middle Space. *Journal of Learning Analytics*, *1*(2), 23–47. https://doi.org/10.18608/jla.2014.12.3

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

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

Liu, A. L., & Nesbit, J. C. (2020). Dashboards for computer-supported collaborative learning. In M. Virvou, E. Alepis, G. A. Tsihrintzis, & L. C. Jain (Eds.), *Machine learning paradigms: Advances in learning analytics* (pp. 157–182). https://doi.org/10.1007/978-3-030-13743-4_9

Magnisalis, I., Demetriadis, S., & Karakostas, A. (2011). Adaptive and intelligent systems for collaborative learning support: A review of the field. *IEEE transactions on Learning Technologies*, *4*(1), 5-20.

McLaren B.M., Scheuer O., & Mikšátko, J. (2010) Supporting collaborative learning and e-discussions using artificial intelligence techniques. *International Journal of Artificial Intelligence in Education, 20*(1), 1-46. http://doi.org/10.3233/JAI-2010-0001

Okada, A., Buckingham Shum, S. & Sherborne, T. (Eds.). (2014). *Knowledge cartography: Software tools and mapping techniques* (2nd ed). London, UK: Springer.

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., & Aleven, V. (2016). Different futures of adaptive collaborative learning support. *International Journal of Artificial Intelligence in Education*, *26*(2), 784-795.

Scardamalia, M. (2003). Knowledge building environments: Extending the limits of the possible in education and knowledge work. In A. DiStefano, K. E. Rudestam, & R. Silverman (Eds.), *Encyclopedia of distributed learning*, (pp. 269–272). Thousand Oaks, CA: Sage Publications.

Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J., & Woodruff, E. (1989). Computer-supported intentional learning environments. *Journal of Educational Computing Research, 5*(1), 51–68.https://doi.org/10/c82wjr

Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning, 8*(4), 375-397.

Schneider, B., & Pea, R. (2015). Does seeing one another's gaze affect group dialogue? A computational approach. *Journal of Learning Analytics, 2*(2), 107-133.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. D. (2015). 3D tangibles facilitate joint visual attention in dyads. *Proceedings of CSCL 2015* (pp. 158-165). Gothenburg, Sweden: International Society of the Learning Sciences.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2016). Detecting collaborative dynamics using mobile eye-trackers. *Proceedings of ICLS 2016*.(pp. 522-529). Singapore: International Society of the Learning Sciences. Singapore: International Society of the Learning Sciences.

Schneider, B., Worsely, M., & Martinez-Maldonado, R. (2020). 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 (pp.). [Publisher's info]

Sharma, K., Caballero, D., Verma, H., Jermann, P., & Dillenbourg, P. (2015). Looking AT versus looking THROUGH: A dual eye-tracking study in MOOC context. *Proceedings of CSCL 2015* (pp. 260-267). Gottenburg, Sweden: International Society of the Learning Sciences.

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

Stahl, G. (2013). Transactive discourse in CSCL. *International Journal of Computer-Supported Collaborative Learning, 8*(2), 145–147.

Stahl, G., Jeong, H., Ludvigsen, S., Sawyer, R. K., & Suthers, D. D. (2013). *Workshop: Across levels of learning: A workshop on resources connecting levels of analysis*. Presented at the International conference of computer-supported collaborative learning (CSCL 2013), Madison, WI.

Suthers, D., Dwyer, N, Medina, R. & Vatrapu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction*. International Journal of Computer-Supported Collaborative Learning, 5*(1), 5-42. 10.1007/s11412-009-9081-9.

Suthers, D., & Rosen, D. (2011). A unified framework for multi-level analysis of distributed learning. *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 64-74). ACM.

Suthers, D., A. Weiner, J. Connelly, & M. Paolucci. (1995). Groupware for developing critical discussion skills. In John L. Schnase and Edward L. Cunnius (Eds.), *Proceedings of CSCL95: First international conference on computer support for collaborative learning* (pp. 341-348). Bloomington, IN: Lawrence Erlbaum Associates.

Tan, J. P-L., Koh, E., Jonathan, C. R., & Tay, S. H. (2018, April). Visible teaching in action: Using the WiREAD learning analytics dashboard for pedagogical adaptivity. Paper presented at *the 2018 Annual meeting of the American educational research association*. Retrieved October 31, from the AERA Online Paper Repository.

Tan, J. P. L., Koh, E., Jonathan, C. R., & Yang, S. (2017). Learner dashboards a double-edged sword? Students' sense-making of a collaborative critical reading and learning analytics environment for fostering 21st century literacies. *Journal of Learning Analytics*, *4*(1), 117-140.

Teplovs, C., & Fujita, N. (2013). Socio-dynamic latent semantic learner models. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive multivocality in the analysis of group interactions* (pp. 383–396). Springer US. https://doi.org/10.1007/978-1-4614-8960-3_21

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc

Tsovaltzi, D., Weinberger, A., Schmitt, L., Bellhäuser, H., Müller, A., Konert, J., … Van Aalst, J. (2019). Group Formation in the Digital Age: Relevant Characteristics, Their Diagnosis, and Combination for Productive Collaboration. In K. Lund, G. P. Niccolai, É. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings 13 th International Conference on Computer Supported Collaborative Learning* (Vol. 2, pp. 719–726). Retrieved from https://ris.utwente.nl/ws/portalfiles/portal/129957345/CSCL_2019_Volume_2.pdf#page=205

van Leeuwen, A. (2015). Learning analytics to support teachers during synchronous CSCL: Balancing between overview and overload. *Journal of Learning Analytics, 2*(2), 138-162.

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

van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014). Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers & Education, 79*, 28-39.

Vogel, F., Weinberger, A., & Fischer, F. (2020). Collaboration Scripts: Guiding, Internalizing, and Adapting. In U. Cress, C. Rosé, A. Wise, & J. Oshima (Eds.), International handbook of computer-supported collaborative learning (pp.). [Publisher's info]

Walker, E., Rummel, N., & Koedinger, K. R. (2011). Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity. *International Journal of Computer-Supported Collaborative Learning, 6*(2), 279–306.

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

Wise, A. F. (2019). Learning analytics: Using data-informed decision-making to improve teaching and learning. In O. Adesope & A.G. Rudd (Eds.) *Contemporary technologies in education: maximizing student engagement, motivation, and learning* (pp. 119-143). New York: Palgrave Macmillan.

Wise, A. F., Azevedo, R., Stegmann, K., Malmberg J., Rosé C. P. et al. (2015). CSCL and learning analytics: Opportunities to support social interaction, self-regulation and socially shared regulation. *In Proceedings of the 11th International Conference on Computer Supported Learning* (pp. 607-614). Gothenburg, Sweden: ICLS.

Wise, A. F., Knight, S., & Ochoa, X. (2018). When are learning analytics ready and what are they ready for. *Journal of Learning Analytics*, *5*(3), 1-4.

Wise, A. F. & Vytasek, J. M. (2017). Learning analytics implementation design. In C. Lang, G. Siemens, A. Wise & D. Gašević (Eds.) *Handbook of learning analytics* (1st ed) (pp. 151-160). Edmonton, AB: SoLAR.

Wise, A. F., Vytasek, J. M., Hausknecht, S., & Zhao, Y. (2016). Developing learning analytics design knowledge in the "Middle Space": The student tuning model and align design framework for learning analytics use. *Online Learning, 20*(2), 155-182.

Wise, A. F. & Schwarz, B. S. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning, 12*(4), 423-467.

Wise, A., Zhao, Y., & Hausknecht, S. (2014). Learning analytics for online discussions: Embedded and extracted approaches. *Journal of Learning Analytics*, *1*(2), 48-71.

Further readings

For an overview of learning analytics, readers may refer to the journal (learning-analytics.info - including a forthcoming special section on collaborative learning analytics), conference, and handbook in that space. There are also examples of learning analytics work in CSCL, including via the following excellent NAPLES resources:

1. An example learning analytics approach grounded in the learning sciences, which demonstrates moving through analytic lenses.
   Williamson Shaffer, D. (n.d.). David Williamson Shaffer: Tools of Quantitative Ethnography: Epistemic Network Analysis and nCoder. Retrieved September 16, 2019, from ISLS NAPLES Network website: http://isls-naples.psy.lmu.de/intro/all-webinars/shaffer_video/index.html

2. A specific example of how physiological measures can give insight into constructs of interest to the CSCL community.
   Jermann, P. (n.d.). Patrick Jermann: Physiological measures in Learning Sciences Research. Retrieved September 16, 2019, from ISLS NAPLES Network website: http://isls-naples.psy.lmu.de/intro/all-webinars/jermann/index.html

3. An example of how analyses of argumentation can be automated for insight.
   Erkens, G. (n.d.). Gijsbert Erkens: Automated argumentation analyses. Retrieved September 16, 2019, from ISLS NAPLES Network website: http://isls-naples.psy.lmu.de/intro/all-webinars/erkens/index.html

4. Potential for CLA in applying learning analytics and educational data mining to learning discourses.
   Rosé, C. P. (2014). Carolyn Rosé: Learning analytics and educational data mining in learning discourses. Retrieved September 16, 2019, from ISLS NAPLES Network website: http://isls-naples.psy.lmu.de/intro/all-webinars/rose_all/index.html

5. An overview of a classic CSCL area with parallels in emerging learning analytics dashboard work.
   Janssen, J. (2013). Jeroen Janssen: Group awareness tools. Retrieved September 16, 2019, from ISLS NAPLES Network website: http://isls-naples.psy.lmu.de/intro/all-webinars/janssen_video/index.html