Intern. J. Comput.-Support. Collab. Learn (2023) 18:1–9
https://doi.org/10.1007/s11412-023-09389-x

Editorial: Nine elements for robust collaborative learning analytics: A constructive collaborative critique

Alyssa Friend Wise[1] · Carolyn Rosé[2] · Sanna Järvelä[3]

Published online: 13 March 2023

Introduction

This editorial represents a collaborative effort between the current co-editors-in-chief of the International Journal of Computer-Supported Collaborative Learning (ijCSCL) and the recent co-editor-in-chief of the Journal of Learning Analytics (JLA), Alyssa Wise, who is also a member of the ijC(LA) have made a presence in ijCSCL. This issue in particular comprises four full articles within this scope, in addition to a timely exposition on Collaborative Learning from an ethics perspective. Thus, it is high time to bring in a voice of leadership from the LA community together with those of the CSCL community to think together about the intersection of work across the two fields.

The four full articles of this March issue offer a view of the kind of work that the CSCL community is engaging in to a) *capture meaningful traces of learning*, b) *map them onto valued learning constructs*, and discover useful ways to c) *present them back to teachers, students and other educational stakeholders* (Chen & Teasley, 2022; Wise et al., 2021a, b). This work is of particular interest to the field of LA due to the skills and experience the CSCL community has developed, both conceptually and methodologically, for d) *dealing with the temporality of learning processes and e) addressing the interdependence in the activity of learners collaborating together*. These *five qualities of CSCL research* offer distinct value to the field of LA, where the majority of attention has focused on modeling individual learning, often through a series of snapshot views (in exception note the 2021 JLA Special Section on Collaboration Analytics and the 2017/2018 two-part Special Section on Temporal Analyses of Learning Data, which featured a large proportion of collaborative learning scenarios). Thus, in addition to thinking about how this particular set of

✉ Carolyn Rosé
cp3a@andrew.cmu.edu

Alyssa Friend Wise
alyssa.wise@nyu.edu

Sanna Järvelä
sanna.jarvela@oulu.fi

[1] NYU Steinhardt, New York, NY, USA

[2] Carnegie Mellon University Language Technologies Institute and HCI Institute, Pittsburgh, PA, USA

[3] University of Oulu, Oulu, Finland

papers contributes to our understanding of and ability to support CSCL, it is worthwhile to consider how they synergize with the LA literature. Doing so is a step towards encouraging productive cross talk across the disciplines that can identify both complementary contributions and productive gaps that offer opportunities to further expand and enhance the work.

This editorial is written as a constructive critique, meant to challenge the field of CSCL to find its unique place in the landscape while striving towards more vital collaboration with the LA community. We do so by evaluating the articles in light of an expansion of the five qualities into eight elements formulated from an LA perspective. We add one more element dealing with *ethical dimensions* discussed in the squib in this issue noting that an ethical focus was not featured prominently in any of the four full articles but serves as the primary theme of the squib that concludes this issue and is a critical aspect of responsible LA work.

Nine elements for robust collaborative learning analytics

In this section, we lay out the nine elements, which expand on the five qualities of CSCL and form the basis for our constructive critique. The first two papers most closely align with elements 2, 3, 4 and 5, which take the interplay between the group and individual levels as well as temporality into account in an online learning context, with the second paper notably featuring multi-channel data. These papers focus on analysis to reveal new scientific insights about collaborative processes rather than attempting to use these analytics as tools or scaffolding to support learning. The second two papers are more LA at their core, with reference to elements 1, 6, 7 & 8 in that they do close the loop by creating a tool for learners and then look at impact with real students. The tradeoff, however, is less attention to the details of temporality and group/individual levels in the analysis, as seen in the first two papers. Below we delve into the specifics of the nine elements, which we see as characterizing a robust lifecycle for collaborative learning analytics work. While all will not necessarily be focused on in a single article, an eye towards each, at least in peripheral vision, is important to keep the larger picture and purpose in view (Wise et al., 2021a, b).

Element 1. Overall orientation to mobilize data traces to inform learning Learning analytics is much more than just a set of educational data science methods for analyzing learning. The guiding vision is one of data as a tool for improving learning and the field is decidedly interdisciplinary in nature, including: researchers and practitioners focused on questions of data policy, infrastructure, and ethics; designers exploring how best to create analytic tools for and with educators and students; educational researchers studying how people work with data-based tools to make decisions; as well as the very important expertise brought by those focused on innovating ways to generate useful data and insightfully applying computational techniques for analyzing them. Thus, aligned with CSCL's orientation to not only understand but improve collaborative learning, key opportunities for collaboration analytics include not only building better theories and models of collaboration but also supporting students' collaborative interactions and their acquisition of better collaborative skills (Schneider et al., 2021). Having this in view from the start of a project is important as more than just a laudable aspiration. It also raises important ethical questions related to visibility (who will data be collected on, who will this data be available to, and given this, what data is appropriate to collect) and critical questions about how such visibility will intersect with existing structures of power and patterns of (in)equity within an institution that provide an important framing for the work.

Element 2. Careful clicks-to-constructs mappings that attend to the learning task Another key element, discussed at length in the Collaborative Learning Analytics chapter of the International Handbook of CSCL (Wise et al., 2021a, b), is the central importance of the clicks-to-constructs mapping (Buckingham Shum & Crick, 2016). This is the chain of inference that allows us to take some automated processing of raw data as evidence for the presence of some conceptually interesting entity or process. For example, Lee and Tan (2017) identify promising ideas in a knowledge-building community through a temporal peak in the betweenness centrality of relevant keywords extracted from student contributions. A strong clicks-to-constructs mapping necessarily relies both on thoughtful data collection and manipulation, but also rich theorization of the collaborative learning interactions of interest (Wise, 2023), which in turn relates to the design of the CSCL task (Martinez-Maldonado et al., 2021).

Elements 3 & 4. Theorization about group and/or individual level. Theorization and modeling of learning as a temporal process In the context of collaborative learning, theorization of clicks-to-constructs should be purposeful with respect to operation at the group and/or individual levels as well as relationships between them and often necessarily the dynamics of evolution over time. Particularly, there is a need to expand from the dominance of basic social network analysis (SNA) as an analytical technique (Kaliisa et al., 2022) to unpack the detailed nature of unfolding interactions and the dynamics of how these themselves evolve. This can be analyzed both in terms of sequences of events as well as flow of activity over time (Molenaar & Wise, 2022). Again such thinking can be coordinated to thinking ahead to whom (individual students, groups of students, teacher, the classroom community) might usefully be shown such information and how they could be expected to act upon it.

Element 5. Multi-channel and/or physical space data Looking at data collection itself more closely, collaborative learning naturally lends itself to multi-channel data (e.g. in the asynchronous context: language plus clicks as well as a potentially a shared product and its evolution). And while much of CSCL has developed based on data collected in virtual spaces, greater availability of low-cost sensing devices (e.g. eye-trackers, motion detectors) as well as advances in the ability of computer vision to extract useful traces from video data open exciting new doors for the study and support of collaborative learning in the physical spaces (e.g. Nguyen et al., 2022) where much of the collaborative learning we care about happens.

Element 6. Careful attention to what information to provide to whom and how Taking to heart the goal of both CSCL and LA to not only understand but support learning, there remains great potential for thoughtful attention to the ways that analytics can be used to inform collaborative interactions and help students improve their collaborative skills. Whether seeing the potential for responsive feedback on collaborative processes as a more agentic alternative to upfront scripting (Wise & Schwarz, 2017) or the use of analytics as a way to make scripting more adaptive (Vogel et al., 2021), the question of what kinds of information to make available to what actors in what ways and at what points in time is an area of work ripe for exploration. Here there are also several potential ways for CSCL to make a contribution back to the field of LA more generally. First, while in LA dashboards have been the primary form taken by analytic technologies, CSCL has a long tradition exploring various mediums for generating group awareness (Buder et al., 2021) that can happen integrated into the flow of collaboration, rather than necessitating "stepping out" to take a reflective moment (Ackermann, 1996). Second, while LA has focused more on the creation of tools for teachers, tapping into some of the original ideas related to student agency and autonomy motivating the pursuit of CSCL offer potential to explore the largely uncharted space of student-facing analytics.

Element 7. Human-centered approach to LA design Recent directions in LA that could prove useful to CSCL researchers working to "close-the-loop" include the development of analytics that are explainable and configurable. The focus here is on allowing for two-way interaction between human and tool (as opposed to just the one-way flow of information), as well as promoting actionability (i.e., supporting movement from an understanding of the current learning situation to what can be changed to productively improve it, Ochoa & Wise, 2021). These are both inherently tied to a growing emphasis on human-centered learning analytics (Buckingham Shum et al., 2019), which centrally draws on the human–computer interaction techniques of participatory and co-design to include the intended users of learning analytics systems in the process of their design (Sarmiento & Wise, 2022). Such approaches attempt to find a productive intersection among (a) the needs of teachers and students within a given CSCL situation, (b) theoretically valued constructs of learning, and (c) available data (Martinez-Maldonado et al., 2021) and are also helping the field move towards greater transparency in the design of LA tools. Here again, CSCL and the larger Learning Sciences have a valuable contribution to make to the field of LA by drawing on their long traditions of design-based research for documenting iterative processes of making and testing design decisions (Hoadley & Campos, 2022).

Element 8. Examination of how LA are used in the world There is increased emphasis in the field of LA on understanding how analytics tools are actually used by teachers, students and others to inform and improve their (collaborative) learning activity. Certainly we want to show that the tools we build have an impact, but simple evaluations showing that learning with an analytic solution (which contains many component decisions with respect to data, analysis, presentation and integration) is more effective than without do not offer specific generalization knowledge contributions to inform other future efforts. In contrast, studies that include detailed mixed-methods component that help us understand how and/or why teachers and students work with analytics tools in specific ways and how this informs their understanding of and engagement in collaborative activity have much greater power to both improve models of intentional collaborative learning and design better CSCL tools in the future. Here "better" is defined as an ability to positively impact collaborative learning activity and the development of collaborative learning and group regulatory skills.

Element 9. Attention to systems level and ethical concerns There is a systems level view of LA that has hopefully been apparent throughout this discussion where the ultimate lens is larger than just the interaction of a number of individuals with an analytic learning tool, but one which takes into account the existing institutions into which such tools will be introduced. This includes elements such as their power structures, expectations of privacy, and underlying technical infrastructure and, most explicitly, an anticipatory consideration of how LA will interface with these to cement and/or change the situation. Considering this larger context, the ways organizations do or don't integrate LA tools into their practices, and, most importantly, the ways in which they do so incorporates an important, socio-technical perspective to the work. This is necessary for us to understand how our efforts ultimately make a difference in the world through both anticipated and unintended effects.

Analytics to support theory development

First in this issue, Mohamed Saqr and Sonsoles López-Pernas offer an article entitled, "The temporal dynamics of online problem-based learning: Why and when sequence matters". This interesting paper combines qualitative coding with sequence and process mining to unpack how problem-based learning (PBL) unfolds online. The paper offers useful insights into group processes of PBL and is thoughtful in the creation of the analytic metrics used, with important attention to both the temporality of learning and the interplay between group and individual levels. For example, in terms of construct mappings that attend to a learning task, the coding featured in this work was well-described theoretically and based on existing schemes that made sense in connection with theories of learning in the well-characterized PBL task. In terms of theorization about group and/or individual level constructs, this article is particularly strong as it features both individual and group level analyses as well as multi-level modeling taking group dependencies into account in modeling individuals. The temporal aspects of collaboration are a central part of the analysis and contribution as the work builds on prior theorization in CSCL and methods from LA conceptualizing temporality in terms of the sequential nature of events.

The second article, by Fan Ouyang, Weiqi Xu, and Mutlu Cukurova entitled, "An Artificial Intelligence-driven Learning Analytics Method to Examine the Collaborative Problem-solving Process from the Complex Adaptive Systems Perspective" also builds on qualitative coding with sequence matching, cluster identification and temporal characterization to describe the different ways groups go about collaborative problem solving (CPS) in terms of cognitive, regulative, behavioral, and socio-emotional aspects. With respect to a careful clicks-to-constructs mapping that attends to the learning task, the article features a detailed description of the task and exploration of the "interactive, cognitive, regulative, behavioral, and socio-emotional aspects of the learning process". Similar to the first article, the coding scheme was theoretically grounded but applied by hand, which limits the ability to "close the loop". From the perspective of LA, the explicit attention to the importance of time and relationship between individual and group levels is quite valuable, as is the skillful use of multiple channels of data. In consideration of group and individual levels, the article takes the perspective of complex adaptive systems theory and emergent phenomena. This offers good justification for coding at individual level but then looking for patterns at the group level. As for temporality, the analysis supports the view that sequence matters. An optimal matching algorithm is used to compare sequences and epistemic network analysis (ENA) and hidden markov models (HMMs) are applied in order to characterize temporal patterns. A notable strength is the incorporation of multi-channel and/or physical space data–the work is in an online context, but uses audio, screen capture, text chat and the final products created–with extra value provided by the explicit exposition of the value of the multi-channel approach in the paper.

Both papers offer important insights into the different ways collaborative learning proceeds over time; however the stance taken is firmly that of a researcher with an emphasis on elements 2, 3, 4, & 5. This offers an opportunity for expansion with respect to the other elements: namely, Overall orientation to mobilize data traces to inform learning, Careful attention to what information to provide to whom and how, Human-centered approach to LA design*,* Examination of how LA are used in the world and Attention to systems level and ethical concerns. Thinking from the perspective of "closing-the-loop" LA and mobilizing these traces for instructors and/or learners, it is not difficult to imagine building a model that could automate either of the qualitative coding process. But this is just one of several important steps for developing useful LA. Perhaps the more interesting question is how the metrics themselves might be developed differently if we were instead to think of them from the perspective of informing collaborative activity as it is in progress. Specific questions might focus on *who* (e.g., groups, individual, the instructor) would most benefit from *what* information and *how* they might take action on it; the questions *they* have about their learning (perhaps as part of self- co- or socially shared-regulation) that this data could help inform; what questions of ethics, transparency and privacy might need to be considered; and how the introduction of this LA might perturb existing routine practices, social relations and power structures. These are things to be considered from top to bottom of the LA development process.

Analytic-enabled support of collaborative learning processes

The second two articles offer complementary strengths to the first two in that LA interventions are proposed and evaluated. Further probing into the underlying logic model of the experimental designs would enable further theorizing, lending additional strength and more nuanced validation of design principles from the successful intervention studies presented in these two contributions to the journal.

In the first of these, Leonardo Silva, António Mendes, Anabela Gomes and Gabriel Fortes present "Fostering Regulation of Learning Processes among Programming Students using Computational Scaffolding" in which the authors test the impact of a set of regulatory scaffolds, which include learning analytics among some other supports, to help students engage better in both individual and collaborative activities in the context of programming education. In contrast to the first two papers, here a core component of the analytics work is mobilizing the data to be shown back to people (in this case students). The specific analytics themselves (i.e., progress metrics, types of errors) are strong in that they are conceived of in an integrated way with the overall approach to collaborative support (here focused on regulation). Looking at these in more detail there is tracking of basic temporality related to regulation, specifically focused on how much a student has done so far and types of errors committed over time. In terms of careful clicks to constructs mapping, however, in this case it is challenging to draw inferences due to the low level nature of the signal and with respect to theorization at multiple levels, the analysis focuses mainly at the individual level. A notable aspect of the work is its explanation of what data to show to whom (students), which again fits well with their overall project focus on regulation. The transferable value of these analytics would be even greater if there was more transparency about their process of creation and a broader consideration of a human-centered approach to LA tool design, which offers an opportunity for growth going forward. Another strength here is that the authors present convincing results of a successful intervention. Many questions for future work are opened up by this article, in particular, by probing further than evaluation of overall effectiveness to delve into the separate effects of different aspects of the complex intervention and engaging in deep sense making about the patterns found.

In the final full paper of the issue, Lanqin Zheng, Miaolang Long, Jiayu Niu, and Lu Zhong present a paper entitled, "An automated group learning engagement analysis and feedback approach to promoting collaborative knowledge building, group performance, and socially shared regulation in CSCL" in which the focus is on groups' collaborative engagement, including ways to automatically detect it and display back to learners while collaborating. The underlying theoretical framing pays primary attention to the group as the unit of analysis, except the socio-gram, which was developed to offer more insights about how group members interacted with specific others. Further investigation of the interplay between the individual and the group is left for future work as is an exploration of temporality. The data features text-based interactions and click-stream data and thus future work might offer more nuanced insights by incorporating more multi-modality in the setup. With respect to careful clicks to constructs mapping, the scheme for hand coding training data seems well-grounded, but an opportunity to probe deeper might relate to validation of the coding constructs via automated coding. Usage of deep learning in automated analysis of collaborative learning has increased over the past five years, and will increase still more with the current attention to Large Language Models in the education sphere. The specific focus of the analytic work in this article comprises cognitive engagement, metacognitive engagement, behavioral engagement, and emotional engagement to classify each segment of the transcripts. A BERT-based model is then used to automatically classify each segment into cognitive engagement, remembering, understanding, applying, and evaluating as well as off-topic. Further scrutiny of the use of such models would be valuable at this crossroads.

A strength of this work is that the LA provided suggestive feedback based on the group learning engagement analysis results. They paid attention to what information to show back to whom (again learners) embedded within principles for their high-level design. Further explanation of the decisions made in the process of designing the tool, for example in relation to the presentation of bar graphs vs. pie charts and sociograms would lend more insights into collaborative learning analytics creation that future researchers would find useful. A valuable next step in the analysis might be an investigation into the mechanisms of how / why aspects of the interventions were appropriated by students, and how these patterns of appropriation might be improved. It is notable that the authors examined effects of the designed tool on group learning engagement, group performance, collaborative knowledge-building level, socially shared regulation, and cognitive load. The rich data collected offer opportunities in future work to probe further into how often and when in the learning process students accessed the dashboard, whether and how it was discussed by the collaborative groups, and what parts of the tool and the information it provided were useful in what ways. This could serve as a powerful impetus for design-based research cycles of human-centered iteration that improve both our theories of how data can be useful for informing collaborative learning and our tools for doing so. Finally, examination of systems level and ethical concerns represent an important are for future work—for example unpacking students' experience of being monitored and the ways in which this and the information generated affect social dynamics and power relations within the group and with respect to the larger learning environment.

Looking forward

The March issue rounds out with a squib by Etan Cohen, Dani Ben-Zvi, Yotam Hod entitled,, "Visions of the Good in Computer-Supported Collaborative Learning: Unpacking the Ethical Dimensions of Design-Based Research". This article provides a much needed guide for addressing the human-centered and ethical perspectives on CSCL, though not focused specifically on LA per se. As these issues are not yet addressed in detail in the four full articles of this issue, this squib is well positioned in conversation with the articles ordered before it, calling for reflection on these issues as future work is planned.

This editorial has introduced the articles of this issue of the journal, couched in a Learning Analytics perspective, challenging the LA work of CSCL to benefit from the values and standards the community represents. Moving forward, in the spirit of cross-fertilization between sister research communities, we invite further such interactions as part of the discourse of this journal.

References

Ackermann, E. (1996). Perspective-taking and object construction. In Y. Kafai & M. Resnick (Eds.), *Constructionism in practice: Designing, thinking, and learning in a digital world* (pp. 25–37). Routledge.

Buckingham Shum, S., & Crick, R. D. (2016). Learning analytics for 21st century competencies. *Journal of Learning Analytics, 3*(2), 6–21.

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics, 6*(2), 1–9.

Buder, J., Bodemer, D., & Ogata, H. (2021). Group awareness. In U. Cress, C. P. Rosé, A. F. Wise, & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning* (pp. 295–313). Springer.

Chen, B. & Teasley, S. D. (2022). Learning Analytics for Understanding and Supporting Collaboration. C. Lang, G. Siemens, A. F. Wise, D. Gašević & A. Merceron (Eds.) *Handbook of learning analytics (2nd ed)* (pp. 86–95). SoLAR.

Hassan, J., Leong, J., & Schneider, B. (2021). Multimodal data collection made easy: The EZ-MMLA Toolkit. *11th International Conference on Learning Analytics and Knowledge* (pp. 579–585). ACM.

Hoadley, C., & Campos, F. C. (2022). Design-based research: What it is and why it matters to studying online learning. *Educational Psychologist, 57*(3), 207–220.

Kaliisa, R., Rienties, B., Mørch, A. I., & Kluge, A. (2022). Social learning analytics in computer-supported collaborative learning environments: A systematic review of empirical studies. *Computers and Education Open, 3*, 100073. https://doi.org/10.1016/j.caeo.2022.100073

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.

Martinez-Maldonado, R., Gaševic, D., Echeverria, V., Fernandez Nieto, G., Swiecki, Z., & Buckingham Shum, S. (2021). What Do You Mean by Collaboration Analytics? A Conceptual Model. *Journal of Learning Analytics, 8*(1), 126–153.

Molenaar, I. & Wise, A. F. (2022). Temporal aspects of learning analytics: Grounding analyses in concepts of time. C. Lang, G. Siemens, A. F. Wise, D. Gašević & A. Merceron (Eds.) *Handbook of learning analytics (2nd ed)* (pp. 66–76). SoLAR.

Nguyen, A., Järvelä, S., Wang, Y., & Róse, C. (2022). Exploring socially shared regulation with an AI deep learning approach using multimodal data. In C. Chinn, E. Tan, C. Chan & Y. Kali (Eds.), *Proceedings of International Conferences of Learning Sciences (ICLS)* (pp. 527–534).

Ochoa, X., & Wise, A. F. (2021). Supporting the shift to digital with student-centered learning analytics. *Educational Technology Research and Development, 69*, 357–361.

Sarmiento, J. P., & Wise, A. F. (2022). Participatory and co-design of learning analytics: An initial review of the literature. *12th International Learning Analytics and Knowledge Conference* (pp. 535–541). ACM.

Schneider, B., Dowell, N., & Thompson, K. (2021). Collaboration analytics—current state and potential futures. *Journal of Learning Analytics, 8*(1), 1–12.

Vogel, F., Weinberger, A., & Fischer, F. (2021). Collaboration scripts: Guiding, internalizing, and adapting. *International Handbook of Computer-Supported Collaborative Learning* (pp. 335–352). Cham: Springer.

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*, 423–467.

Wise, A. F., Knight, S., & Ochoa, X. (2021a). What makes learning analytics research matter. *Journal of Learning Analytics, 8*(3), 1–9.

Wise, A. F., Knight, S., & Buckingham Shum, S. (2021b). Collaborative learning analytics. In U. Cress, C. P. Rosé, A. F. Wise, & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning* (pp. 425–443). Springer.

Wise, A. F. (2023). Collaborative learning theory and analytics. K. Bartimote, S. Howard & D. Gasevic (Eds.) *Theory informing and arising from learning analytics*. Springer.

Worsley, M., Anderson, K., Melo,N., & Jang, J. Y.(2021). Designing analytics for collaboration literacy and student empowerment. *Journal of Learning Analytics*, 8(1), 30–48. https://doi.org/10.18608/jla.2021.7242