International Society of
the Learning Sciences

Using Learning Analytics to Assess the Intended and Enacted Learning Design: An Epistemic Network Analysis Approach

Rogers Kaliisa, Department of Education, University of Oslo. rogers.kaliisa@iped.uio.no
Jennifer Scianna, Department of Educational Psychology, University of Wisconsin; jscianna@wisc.edu
Anders I. Mørch, Department of Education, University of Oslo; anders.morch@iped.uio.no
Anders Kluge, Department of Education, University of Oslo; anders.kluge@iped.uio.no

Abstract: This paper explores the potential of using learning analytics (LA) to provide context for the interpretation of the intended and enacted learning design. We use Epistemic Network Analysis (ENA) to analyze and visualize students' learning patterns enacted across seven asynchronous online discussions conducted in a semester-long undergraduate course and how they relate to the course's learning objectives. The findings showed connections between the design of each week's tasks and the epistemic and social connections made by the students. The implications of the findings are for designing activities that promote student engagement and productive knowledge-building discourse in computer-supported collaborative learning (CSCL) environments.

Introduction and Background

Most teachers rely on summative assessments (coarse-grained analysis) such as the end of term examinations, as a benchmark to measure students' learning and to retrospectively make decisions regarding how best to teach their subjects to the next cohort of students (Conole, 2012). However, such methods are prone to challenges such as personal bias and the failure to monitor students' online learning patterns (i.e., course logins, discussions attended, student-student, and student-course artifact interactions) during the run of the course (Rienties et al., 2017). One way to address this challenge is by using more objective and automated methods to evaluate students' online learning in real-time enabling teachers to make timely, informed (formative) learning design decisions. LA researchers have pointed to the rich potential for connecting LA and learning design (Lockyer & Dawson, 2011). They argue that effective alignment would provide a valuable context for the interpretation of metrics from LA algorithms (Rienties et al., 2017) and would offer teachers and researchers the evidence they need to assess the effectiveness of their teaching practices and innovations (Kaliisa et al., 2020). Consequently, several researchers have directed their efforts into exploring the connection between LA outputs and learning design.

This paper suggests LA as a possible approach to exploring students' online learning patterns for the purpose of design. Specifically, we are interested in exploring students' digital artefacts (i.e. discussion posts) to produce insights into students' participation and meaningful discourse patterns that could support teaching and learning design decisions. Adopting a sociocultural approach, learning is viewed as participation in, and mastery of, subject-specific discourses and practices (Säljö, 2009). We argue that the use of students' artefacts (e.g. discussion forum posts) with epistemic network analysis (ENA) and interpreting the results based on the course objectives could allow for a more nuanced description of student learning and knowledge building patterns (Oshima et al., 2012). From a sociocultural perspective, we regard the learning design as a cultural and material tool that is used to guide social encounters (e.g. online discussions) through the sequential ordering of actions and activities that students will be engaged in (Ludvigsen et al., 2010). As a methodological contribution, we use ENA, a quantitative ethnography approach (Shaffer, 2017) to analyze the discourse features of students' discussion posts in light of the intended course objectives. Our approach is based on the consensus among sociocultural theorists that discourse is an important unit of analysis for understanding learning through analyzing subject knowledge, argumentation skills, rhetorical moves, and student-student interactions (Säljö, 2009; Oshima et al., 2012). Towards this goal, we address the following research question: How and to what extent, can ENA capture students' progress with the course to provide context for the interpretation of the intended and enacted learning design?

Methods

The data used in this study were extracted from discussion forum contributions posted on Canvas, a learning management system, within a blended bachelor's course (involving face-to-face and online activities) in technology-enhanced learning. The course included 36 enrolled students and four teachers. The researchers received ethics approval from the national review board, and informed consent was gained from 30 students and four teachers whose data informed the analysis presented in this paper. There were six learning modules in the course, and weekly online discussions were created based on each module. The discussions ran for 7 weeks, from

January to April 2020. They were conducted asynchronously, and all subsequent messages in the thread were text-only. Participation in the discussion was compulsory, with each student expected to make two contributions and respond to at least one other student every week. Each week, the teachers created a new discussion thread based on the topic of the next lecture. Thereafter, students posted their contributions in response to either the main discussion question or posts by other students. The discussion instructions were provided by the teacher in charge of the respective module, and relevant instructions were provided to the students alongside the discussion questions. Some teachers actively participated in the discussions, while others did not.

Data Coding and Analysis

To code the discussion posts for use in ENA, we followed a deductive coding approach and developed codes based on the learning objectives of the course. Two raters coded the data manually using a spreadsheet. During the first phase, 5% of the posts from the different weeks were randomly selected, and the two coders worked separately to code the data. Later, the two coders compared the results of their respective codes. Next, to ensure validation of the codes in relation to the course objectives, we invited two of the four-course teachers to discuss the suggested codes as they pertained to the learning objectives. During this process, some codes were added (e.g., integration, design, and personal), while others were revised (e.g., critical thinking was revised to critical analysis). This procedure resulted in 7 codes (Theory, Technology, Critical Analysis, Group Engagement, Personal, Integration, and Design). Finally, the two raters coded the remainder of the dataset, settling any disagreements through social moderation. Some posts were assigned multiple codes.

We then used the ENA Web Tool (version 1.7.0) (Marquart et al., 2018) to examine the content of the weekly online discussions. The unit of analysis consisted of each post by a given participant grouped by the week of participation. Code connections in our ENA model were defined as occurring only between initial posts to the main forum and the replies to them. We ensured this by assigning metadata to each post that denoted the branching structure of the discussion and defining the conversations via this structure. The resulting networks were aggregated for posts for each student in the model. The ENA model normalized the networks for all the students, accounting for varying levels of participation between students, prior to determining the visualization via singular value dimensional reduction (Marquart et al., 2018). This yielded coordinate points in a two-dimensional space that describe the data. The means of the groups were represented by squares surrounded by a confidence interval (see Figure 1). For example, plotting the epistemic networks for Weeks 1 through 7 yielded a plot that showed 19% of the variance in the data along the y-axis and 22.9% along the x-axis. The ENA had co-registration correlations of 0.97 (Pearson) and 0.98 (Spearman) for the first dimension, and 0.96 (Pearson) and 0.96 (Spearman) for the second. These measures indicated strong goodness of fit between the visualization and the original model, which provides justification for our ability to interpret the position of the codes, units, and groups in a meaningful way.

Results

Visualization of the ENA results is shown in Figures 1 & 2. In Figure 1, the y-axis can be characterized by a progression of course topics from theory-driven and personal to design. This coincides with the structure of the course syllabus, implying the early weeks lie along the negative y-axis and the late weeks in the positive direction. The x-axis is characterized by the type of source (integration) students used to justify their positions within their posts. The positive direction in the x-axis includes discourse that has more connections to *Personal* and *Group Engagement,* whereas discourse that lies along the negative x-axis includes connections to external source codes, such as *Integration* and *Analysis.* As the course is centred on educational technology, it is unsurprising to see strong connections between *Technology* and various other codes, thus positioning *Technology* near the origin of the model's plot.

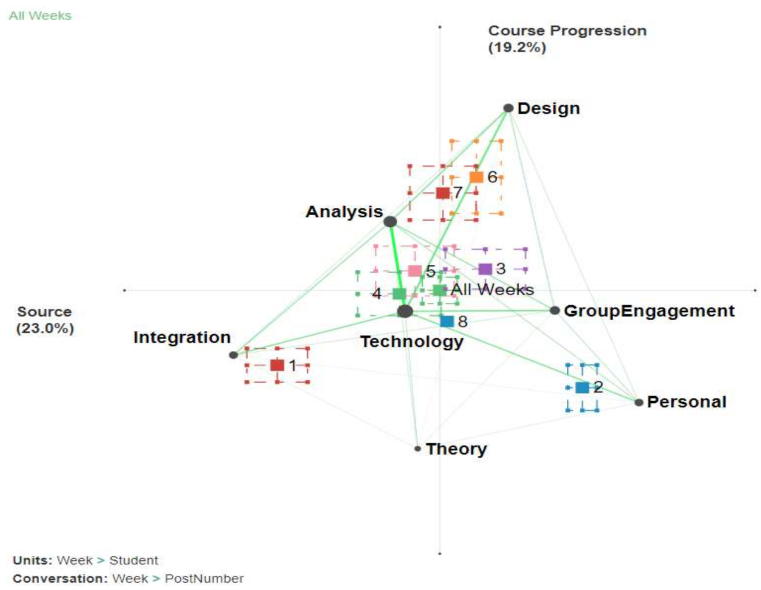

Figure 1. Overall ENA graph representing connections of codes across the seven weeks of discussion. The x-axis is based on svd1 and characterized by connections to source codes such as *Personal* or *Integration*. The y-axis is based on svd2 and is primarily focused on connections to *Design*.

The content of any given week can be extrapolated based on its position in the ENA graph. For example, Week 1 lies closer to *Integration* and is, therefore, more likely to include posts that connect other codes to *Integration,* whereas Week 2 lies closer to *Personal*, so posts in that week are more likely to connect to that code. This can be visualized using a subtraction plot (Figure 2) to demonstrate which connections between codes are more prevalent between two weeks. When the position of the individual units and their relationship to the codes are able to indicate the connections students were making across the semester, ENA can serve as a valuable tool for content analysis. For example, the plots can help us to identify if students reveal understanding with respect to certain parts of the course and how they relate to the intended course objectives.

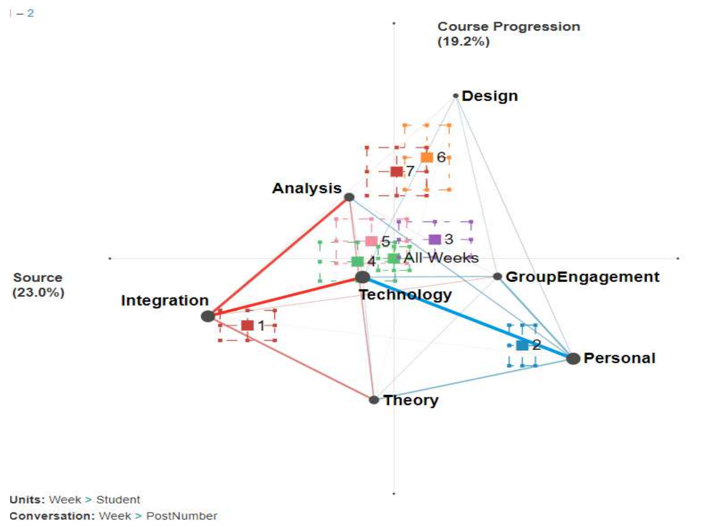

Figure 2. A comparison plot of Week 1 (red) and Week 2 (blue). Week 1 posts more often connect to *Integration* whereas Week 2 posts more often connect to *Personal*. Both weeks have a strong connection to *Technology*

Discussion and conclusion

The purpose of this study was to explore the potential of using LA to provide context for the interpretation of the intended and enacted learning design. To achieve this, we used ENA to explore and understand students' learning trajectories/behaviors on a week-by-week basis as informed by the learning design. By applying ENA, it allowed us to gain a methodologically sound view of students' learning patterns, and how these were related to the course objectives. ENA models illustrated the shift in connections between the different course objectives across the 7 weeks of the discussion. For example, the models showed that students made few connections between Design and Technology in Week 2, but by Week 7, there are strong connections between *Design* and both *Analysis* and *Technology*. Additionally, in Week 1, students focused more on *Theory*, *Integration*, and *Technology* because the discussion prompt and learning goals were ultimately related to these topics. At the same time, in Week 2, where the discussion question invited students to use their personal experiences to reflect on the role of technology, we saw the focus of the topics being more on *Personal*, *Technology*, and *Group Engagement*. The fact that we were able to model student progress through the course by operationalizing a coding scheme grounded in the course objectives suggests that the course design—particularly the tasks set by the respective teachers—could have influenced these patterns, as noted by Rienties et al. (2017). These findings align with Damşa et al.'s (2019) ecological perspective of learning environments, which states that what emerges is facilitated by designed activities, technologies, and spaces. In practice, the identification of such patterns could support teachers in self-monitoring their own learning designs and interventions to determine how effectively their intended designs are when utilized by students, and whether revising the course may be beneficial during the course's run or in future iterations.

Limitations to this study. First, the study was based on data from 30 students and 4 teachers within a blended learning context. This implies that the results and the processes followed may not be directly replicable to studies that use larger datasets. In particular, we took a deductive approach to the analysis of discussion forum data by coding the posts manually, guided by the course objectives. While we regard this approach as reliable when analyzing discussion forums, it might be laborious for researchers working with very large datasets. Moreover, this method required many methodological decisions, such as deciding on the labels for the different codes based on the course objectives. It is possible that if a different coding scheme were used, the results could have been different. More practically, we used this method over a complete dataset; it is unclear whether the same results would be obtained by plotting the datasets weekly, over time, or if the application of this method relies on the ability to code and model one semester's dataset and apply the model to future iterations of the same course. In this way, a delayed approach may not be as beneficial when teachers need the most critical feedback about a course design, which occurs during major iterations.

References

Conole, G. (2012). *Designing for learning in an open world* (Vol. 4). Springer Science & Business Media.

Damşa, C., Nerland, M., & Andreadakis, Z. E. (2019). An ecological perspective on learner-constructed learning spaces. *British Journal of Educational Technology*, *50*(5), 2075-2089.

Kaliisa, R., Kluge, A., & Mørch, A. I. (2020). Combining Checkpoint and Process Learning Analytics to Support Learning Design Decisions in Blended Learning Environments. *Journal of Learning Analytics*, *7*(3), 33-47.

Lockyer, L., & Dawson, S. (2011, February). Learning designs and learning analytics. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 153-156).

Ludvigsen, S., Rasmussen, I., & Krange, I. (2010). Intersecting trajectories of participation: Temporality and learning. In *Learning across sites* (pp. 117-133). Routledge.

Marquart, C., Hinojosa, C., Swiecki, Z., Eagan, B., & Williamson Shaffer, D. W. (2019). *Epistemic Network Analysis [Online software]* (1.5.2) [Computer software]. Epistemic Analytics. http://app.epistemicnetwork.org

Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: a social network analysis application for knowledge-building discourse. *Educational technology research and development*, *60*(5), 903-921

Rienties, B., Nguyen, Q., Holmes, W., & Reedy, K. (2017). A review of ten years of implementation and research in aligning learning design with learning analytics at the Open University UK. Interaction Design and Architecture (s), 33, 134-154.

Shaffer, D. W. (2017). *Quantitative ethnography*. Lulu. com.

Säljö, Roger (2009). Learning, Theories of Learning and Units of Analysis in Research. Educational psychologist, 44:3, 202-208