Exploring a Text-Mining Approach as Rapid Prototyping Tool for Formative Assessments in Inquiry-Based Online Learning

Alejandro Andrade, Chris Georgen, and Michael Stucker
laandrad@indiana.edu, cgeorgen@indiana.edu, mstucker@indiana.edu
Indiana University Bloomington

Abstract: We make a preliminary case for a computational method intended to facilitate real-time formative assessment in online inquiry-based learning environments. With a focus on talk and text as disciplinary, we aim to address how learning analytics, in this case, text mining, can provide learners and instructors with meaningful information in rapid and real-time to support learning and engagement. Our results show that in measuring the distance between the expert and learners' discourse from forum posts and verbal discussions, resulting similarity values can offer stakeholders evidence of student learning trajectories. Moreover, similarity values provide teachers with an automated measure of students' progress toward disciplinary discourse, and also reveal critical moments during the collaborative activity where more alignment with disciplinary ways of talk are being enacted.

Introduction

In recent years, there has been increasing recognition of the synergies between Computer-Supported Collaborative Learning (CSCL) and Learning Analytics (LA) (Jeong & Hmelo-Silver, 2016; Ludvigsen, 2016; Rodríguez-Triana, Martínez-Monés, Asensio-Pérez, & Dimitriadis, 2015). For example, LA has been used to analyze large amounts of student-generated data to explore and refine learning trajectories during collaboration (e.g., Xing, Guo, Petakovic, & Goggins, 2015). Another avenue of research uses LA to provide rapid or real-time feedback on student performance to instructors as they organize and orchestrate collaboration (e.g., Berland, Davis & Smith, 2015; Vatrapu, Teplovs, Fujita, & Bull, 2011). With this in mind, exploring the use of automated or semi-automated techniques is important in online, semi-structured learning environments where teachers are required to parse large amounts of information in order to make quick inferences about student learning. To make such inferences, teachers need support in developing both summative and formative assessments of the learning that takes place during collaboration. The purpose of this paper is to explore the affordances of a computational approach to facilitating real-time formative feedback in CSCL environments. With a focus on the discursive practices of collaboration, we aim to address how learning analytics–in particular text-mining–can capture whether a particular collaborative activity has an effect on the quality of learning. Moreover, we see text-mining and automation as capable of delivering rapid, iterative assessment prototypes that provide stakeholders with meaningful, albeit coarse-grained, data that may support collaborative learning.

Theoretical background

Our focus is on the effect of collaboration in fostering alignment between student and expert discourses as students progress through inquiry-based learning activities. Discourse alignment serves as an indicator of sociocultural learning, because it is assumed that the way in which learners engage in dialogue is evidence of how they engage with knowing and reasoning in a particular field or discipline (De Liddo, Shum, Quinto, Bachler, & Cannavacciuolo, 2011). From a sociocultural perspective of learning and knowing (Brown, Collins, & Duguid, 1989), it is expected that the more students learn in a particular field, the more they are to adopt the ways of expert talk—as in the process of enculturation (Lave & Wenger, 1991). Here, we focus on a circuitry course wherein students discuss current and voltage in a particularly disciplinary way. For example, initially, a student may underspecify disciplinary principles when referring to the elements in a circuit: "The switch is open which means no *electricity* is flowing." Later, as she advances through the class and develops her disciplinary discourse (Mercer, 2008), she might say: "Since the switch was open, this *series* circuit is *incomplete* and the *current* couldn't flow around the circuit any more." Compared to the former, the latter statement is more disciplined to circuitry. To understand the development of, and to explore automated ways for, assessing students' disciplinary discourse, we explored a computational approach to automatically analyze text-based data. Subsequently, we applied this analysis to transcripts of student inquiry in order to target particular moments of disciplinary talk in collaboration. This text-mining approach provides measures of learners' discursive alignment to expert benchmarks throughout online, inquiry-based learning environments.

Methods

Participants and data sources

We collected textual and video data from undergraduate students (n=21) enrolled in a sound engineering course focused on the mathematical and engineering principles of analog electronics. The course is taught by two engineers, who serve as our expert benchmarks, each with considerable technical, professional, and teaching experience. In general, the course is structured around six modules that culminate in collaborative inquiry around unintuitive, yet foundational, concepts of audio engineering. This inquiry is conducted in Peer Investigation Groups (PIGs) composed of stable groups of 3-5 students. PIGs are structured by four sequential activities: 1) individual answers (text), 2) preliminary discussion and group answers (text), 3) online collaborative inquiry (video), 4) final group answers (text). In phases 1, 2, and 4, each expert individually scored student textual responses [0,5]. In phase 3, video from synchronous online inquiry was collected using screen capture tools embedded within the videoconference software (Figure 1).

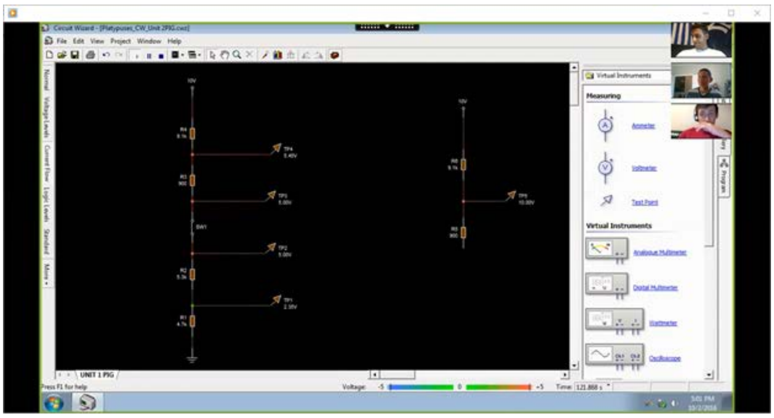

Figure 1. Online synchronous videoconference meeting with the circuitry simulator.

Our analysis will focus on student work in the second of six PIGs (PIG 2). PIG 2 targets interactions between voltage and current through the exploration of resistors and dividers. Students were given a simple circuit (Figure 2.a) and asked to calculate voltage at each test point. Next, the circuit was opened between two test points (Figure 2.b) and voltages were calculated again. The common misconception is based on a misinterpretation of Ohm's Law: you cannot have voltage without current. Since the circuit is open and current cannot flow through the entire circuit, students incorrectly predict the voltage as zero at each test point. Finally, students were asked to "describe why your answers for the open switch are what they are." The goal here was to require students to articulate the engineering principles behind their calculations.

a) Simple Circuit　　a) Open Circuit

Figure 2. PIG 2 circuitry problem.

Textual data

To examine the effect of collaboration we collected student individual, group, and final group answers to the question "describe why your answers for the open switch are what they are." We focused on this question due to prevalence of disciplinary discourse. Moreover, mathematical calculations of voltage are quite simple, whereas the majority of errors are based in misapplications of engineering principles. An example of these forum posts can be seen in Table 1.

Table 1: Examples of pre-, during-, and post-collaboration answers

Pre-Collaboration	During Collaboration	Post-Collaboration

"The switch is open which means no electricity is flowing. Therefore, no voltage is going through the circuit either."	"Since the switch was open, this series circuit is incomplete and the current couldn't flow around the circuit any more. Therefore, there would be no voltage at any point of the circuit except for the battery."	"Once the switch is open, TP3 and TP4 would connect to ground individually and it becomes a parallel circuit. In this parallel circuit, R4 and R3 would have the same voltage as the source, which is 10V."

Video data

To investigate the process of collaboration we recorded online collaborative inquiry. The video data was transcribed clean verbatim.

Expert benchmarks from course instructors

The course instructors provided their own answers to the questions and were used as the expert benchmarks. An example of a benchmark response is shown in Table 2.a.

Analysis of student-expert discursive alignment in pre- and post-collaboration responses

In order to get a similarity value from each student response to the expert's benchmark response, we constructed a document-term-frequency matrix. This approach is sometimes referred to as a "bag-of-words" approach (Bird, Klein, & Loper, 2009). The rows in this matrix are text-documents such as the students' and instructor answers to the PIG question; columns are terms (relevant words); and entries are term frequencies. Thus, each row is a vector of term frequencies for a particular document. Following Kopainsky, Pirnay-Dummer, and Alessi (2012), we extracted only nouns and names by using part-of-speech (POG) tags. We only took names and nouns as they might represent concepts (Kopansky et al., 2012), but we also could have also accounted for verbs and adverbs that represent relationships between those concepts. In future iterations of the project, we will address these methodological variables by systematically comparing the inclusion of various parts of speech to the text-mining algorithm.

We illustrate the procedure with the instructor answer shown in Table 2.a. We used Python's Natural Language Processing NLTK package to produce POS tags, as shown in Table 2.b. In keeping only those terms with 'NN', 'NNP', or 'NNS' tags, we stripped the document of everything but nouns and names. The following step was to get rid of capitalization and stem the words so that 'circuit' and 'circuits' are not counted as different terms, as shown in Table 2.c. Note that though the word 'resistance' was stemmed as 'resist,' 'resistor' was not. This list also contains some non-relevant terms such as 'word', 'end,' and 'matter,' but for the most part the automated word stemmer works relatively well. Then, we converted the list into a term-frequency vector where entries are term frequencies (see Table 2.d). Although the term-frequency vector (Table 2.d) is a high-level representation of the original text (Table 2.a), the claim we make here is that it contains all the relevant features to represent a viable approximation of enacted discourse.

Table 2: Example of Text-Mining Process and Analysis

(a) Expert benchmark example: "You can have Voltage without current, but not current without voltage. Since there is an open switch we know that current equals zero. To find the voltage drop of each resistor we insert that 0 amps of current into Ohm's law to get V=0*R. We know that V will always equal zero (no matter what the resistance is). If each resistor drops 0 volts, then that means there is 0 volts difference between the two ends of each resistor, in other words the voltage is the same at both ends of each resistor."
(b) Vector with Part-of-Speech tags for each term in expert answer: [('You', 'PRP'), ('can', 'MD'), ('have', 'VB'), ('Voltage', 'NNP'), ('without', 'IN'), ('current', 'JJ'), (',', ','), ('but', 'CC'), ('not', 'RB'),...('in', 'IN'), ('other', 'JJ'), ('words', 'NNS'), ('the', 'DT'), ('voltage', 'NN'), ('is', 'VBZ'), ('the', 'DT'), ('same', 'JJ'), ('at', 'IN'), ('both', 'DT'), ('ends', 'NNS'), ('of', 'IN'), ('each', 'DT'), ('resistor', 'NN'), ('.', '.')]
(c) Vector with only stemmed nouns and names: ['voltag', 'voltag', 'switch', 'equal', 'zero', 'voltag', 'drop', 'resistor', 'amp', 'ohm', 'law', 'v=0*r.', 'v', 'zero', 'matter', 'resist', 'resistor', 'volt', 'volt', 'differ', 'end', 'resistor', 'word', 'voltag', 'end', 'resistor']

(d) Term-Frequency Vector:
FreqDist({'voltag': 4, 'resistor': 4, 'end': 2, 'zero': 2, 'volt': 2, 'ohm': 1, 'differ': 1, 'law': 1, 'word': 1, ...})

We followed the same procedure with each student response, and by stacking these vectors, we produce a document-term matrix, with as many rows as there are students, and as many columns as there are terms contained in the forum posts. For instance, an excerpt of the document-term-frequency matrix for the pre-collaboration PIG looks like the one shown in Table 3.

Table 3: Example of a Document-Term-Frequency Matrix

Document/Term	+10v	account	act	amp	chang	…	circuit
Expert	4	0	0	1	0	…	1
Student 1	0	1	0	0	0	…	2
…	…	…	…	…	…	…	2
Student 21	0	0	1	0	1	…	2

The distance from each student document to the benchmark document was computed using the cosine similarity metric (Leydesdorff, 2005), which measures the cosine of the angle between two vectors in a multi-dimensional space. Cosine similarity values range between 0 and 1, where 0 represents two totally different and 1 represents two totally identical texts. We selected the cosine similarity metric because of its robustness in the presence of sparse vectors (vectors with many zeros, Leydesdorff, 2005), as is the case with our dataset. To study the effect of collaboration in the improvement of individual students' disciplinary discourse, group mean similarity values were computed between the expert benchmark and pre- and post-collaboration phases and compared using a dependent-measures t-test at 5% significance level.

Validity and reliability of similarity values

To analyze the validity of the similarity values for capturing the forum post accuracy/correctness and student discursive alignment, quantitative scores for each PIG text were obtained from the instructor and correlated with the similarity values. To analyze the reliability of the similarity values, a correlation analysis was conducted between the similarity values computed from each expert benchmark.

Analysis of student-expert discursive alignment during synchronous collaboration

To study the development of each PIG's collaboration quality, similarity values were computed between the expert benchmark and a 100-word sliding-window of the transcripts. A sliding-window is a segment of certain number of words (100 in this case) in which the segment, or window, moves forward 1 word at a time. For instance, suppose you want to create a 3-word sliding-window for the text: "There can be voltage without a current", thus, the first three windows would be: ["There", "can", "be"], ["can", "be", "voltage"], ["be", "voltage", "without"]. A 100-word window includes enough information as to the dialogue that is taking place during that window of collaboration, and it is short enough as to produce a smooth measure of the ongoing flow of the conversation over time. The values computed per sliding window are then plotted on a line-over-time chart, which shows changes in student-expert discursive alignment over the span of the inquiry. The goal is to produce a visualization to show critical moments during the collaboration where students are engaging in disciplinary discourse. After identifying these moments, one can compare the prevalence of such moments across groups, or go back to the video data and conduct more in-depth qualitative analyses of these identified collaboration moments.

We also computed similarity values between each sliding window and each student's final texts. Although this is not possible to do in real-time while collaboration develops, because students are yet to produce the final text, it can be done as a post facto analysis. The goal is to provide a vantage into relevant moments during the collaboration that helped students orient their writing for their final texts. We hypothesize that sometimes expert's and final texts' similarity values would align, indicating that students were able to capture these productive conversations in their final texts. But we also anticipate that sometimes students can have some productive conversations and yet fail to capture these reflections in their final texts. When this happens, a lack of alignment would be observed between the two similarity values trajectories.

Results

Effect of collaboration on individual learning
Results show that, according to our discursive alignment measure, collaboration had a significant effect in helping students align their discourse toward more disciplinary ways. In comparing the mean groups between the pre- (M_{pre} = 0.295, SD_{pre} = 0.176) and post-collaboration (M_{post} = 0.434, SD_{post} = 0.133) texts, there is a statistically significant increase of .139 points in the similarity values, $t(17) = 2.59$, $p = .019$, $d = 0.61$. This means that the forum posts became more similar to the way an expert would answer the question after students had an opportunity to interact in the collaborative forum and the videoconference activity.

Validity and reliability of similarity values
Two pieces of information provide validity evidence that our measure of disciplinary discourse, represented by the similarity values, actually attest an improvement in students' conceptual understanding. First, there is a positive moderate-to-large association between the similarity values and the teacher-assigned forum post scores, $r = .58$, $p = .011$, 95% CI [0.157, 0.824]. Second, the teacher-assigned scores also show a significant increase of 1.66 points from pre- to post collaboration (M_{pre} = 1.167, M_{post} = 2.833), $t(17) = 3.58$, $p = .002$, $d = 0.84$. This means that both teacher-assigned scores and similarity values covariate in the same direction and equivalent magnitudes.

We also found that this similarity value approach seems to be reliable at finding students' discursive development. Again, there are two pieces of reliability evidence. First, there is a positive moderate-to-large association between the two ratings provided by the two expert benchmarks, $r = .54$, $p = .016$, 95% CI [0.117, 0.799]. Second, the second expert benchmark also shows a significant increase in students' similarity values, $t(17) = 8.04$, $p = .078$, $d = 5.68$. These pieces of evidence indicate that results are very similar regardless of which expert benchmark is used, and also imply that similarity values can serve as proxies for conceptual understanding because of their association with teacher-assigned scores.

Time-sensitive analysis of collaboration quality
In this section, we explore the affordances of the discursive-alignment-over-time visualization chart. Figure 3 shows the chart for the group called "Koalas," which is a high-achieving group, according to the teacher-assigned scores (M = 4) and similarity values (M = 0.63). The chart displays two visible lines over time, though there are four lines in total. The solid bar represents the similarity distance to the expert benchmark, whereas the dotted line represents the distance to the final text, at every minute of the conversation. It is apparent that all students posted the same final text, because all the student lines overlay. The graph shows that this was a relatively long conversation of approximately 24 minutes (compared to other groups' videoconferences). From a quick read at the ebbs and flows of the lines, it is apparent that the discussion was really on-target around minutes 4, 6, 13 and 17, and that there was a drop in on-target talk between minutes 7 and 10. The slow decline at the end of the chart shows that productivity slowly went down after 17 minutes into the activity. The similarity values against both expert and final post benchmarks overlap for the most part, implying that students' final texts' ideas came from moments aligned with the expert's discourse.

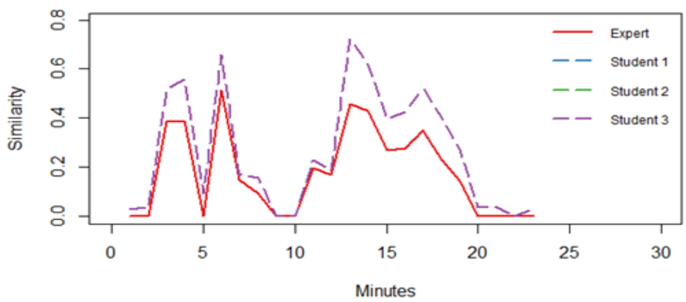

Figure 3. Koalas group collaborative discursive alignment with expert and final text over time.

As a validity check, we took a look at various points in time of the conversation to see if high similarity values represent good disciplinary discourse and low similarity values far off disciplinary discourse. For instance, during the first two minutes of the activity the similarity values were very low. As Excerpt 1 shows (see Table 4), students were settling in and thus their talk was not about circuitry, instead, their conversation revolved around technical issues with the videoconferencing software. On the other hand, one of the most productive moments in the conversation seemed to have occurred between minutes 11 and 14, where there is a noticeable spike in the similarity values. As Excerpt 2 shows (see Table 5), at this point in the conversation there is an interesting exchange of ideas around the question of whether there can be voltage without current and how switches affect current flow. These two excerpts provide good evidence that similarity values extracted from the group's dialogue can provide a valid approach for measuring how discursive alignment develops throughout a collaborative activity. Then again, this measure only captures discursive alignment with an expert benchmark and nothing more, this is why other important aspects of collaborative learning do not seem reflected by the chart of similarity values over time. For instance, Excerpt 3 (see Table 6) shows a very important moment between minutes 8 and 10. This point in time shows a sharp drop in discursive alignment, and yet the dialogue reflects an interesting collaborative exchange where students try to find common ground around their interpretations of what the question is asking. We believe that, however, in the future, we might explore ways to capture other relevant aspects of the collaborative activity by systematically examining and developing distinct collaboration benchmarks.

Table 4: Excerpt 1: No disciplinary talk

[00:00:04]	**Student 1:**	So, I'm just going to connect the rest of these real quick.
[00:00:18]	**Student 2:**	Wait, this share screen's kind of weird.
[00:00:21]	**Student 1:**	Is it?
[00:00:25]	**Student 2:**	I can't go to like my actual...
[00:00:29]	**Student 3:**	Go up to options and say exit full screen.
[00:00:34]	**Student 2:**	Okay, thanks.

Table 5: Excerpt 2: High disciplinary talk

[00:11:59]	**Student 2:**	But did we ever say specifically what the voltage would be without current?
[00:12:02]	**Student 1:**	Exactly, we had never been in a situation to apply that until right now.
[00:12:16]	**Student 2:**	I think we have to kind of figure out when there's no current and when you're also dealing with a switch, how does that affect the voltage. Rather than whether there is voltage or not, it was more how it was affected.
[00:12:35]	**Student 3:**	I thought that this question was like asking like we had to find total current first before we could find any of the voltage for the test points. I thought that was the question.
[00:12:58]	**Student 2:**	There is no current, because it's open. I'm pretty sure they're all talking about the open one.

Table 6: Excerpt 3: No disciplinary talk, yet good collaborative talk not captured by our measure

[00:08:24]	**Student 1:**	I don't really know what else I'd put for question four. I mean, our first answer was right.
[00:08:36]	**Student 3:**	It was saying if you have to go back and find a new answer, how would you do that?

[00:08:47]	**Student 1:**	Yes, but are you saying that's what we did?
[00:08:51]	**Student 3:**	No, it's saying like even though we did prove it on our first group answer, [...] So, it's saying if we were wrong, what would we do in that situation.
[00:09:19]	**Student 1:**	I'm actually interpreting that a little differently. I'm reading it as if you did have to go back, then you need to answer the second part of the question.
[00:09:30]	**Student 3:**	I see what you're saying. It's saying, or did you have to go back and do this?

Finally, it seems important to show that our measures of discursive alignment can show interesting individual dynamics within collaboration groups. For instance, Figure 4 shows the chart over time for the group called Otters, which is also a high-achieving group (Teacher-assigned score = 4.67, similarity values = 0.45). This was a relatively short conversation, 8 minutes approximately, where the similarity values against expert benchmark show that the alignment was higher during the second half of the collaborative activity than the first half. However, students' final forum posts reflect different trajectories; student's 1 similarity values reflect a higher similarity during minute 3, whereas students' 2 and 3 trajectories show higher values during the second half of the activity. This can be interpreted in the following way: student's 1 final forum post seemed to have come from the ideas discussed during minute 3, whereas for students 2 and 3, their ideas came from the second half of the discussion. Although all three students got a high grade (5, 4, and 5, respectively) in the teacher assigned score, similarity value for student 1 (0.37) is lower than for students 2 and 3 (0.46 and 0.51, respectively). We think that student 1 deserves further analysis in order to understand why she is getting good grades but her discourse is not yet aligned with that of the expert.

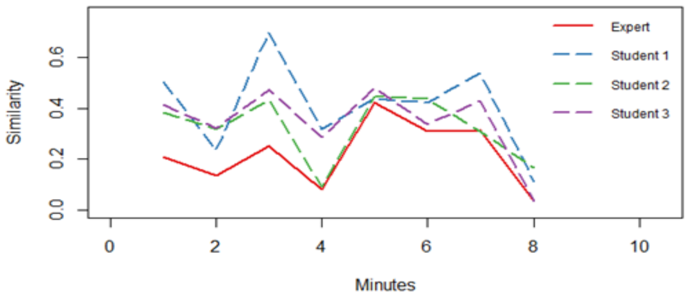

Figure 4. Otters group collaborative discursive alignment with expert and final post over time.

Conclusions and implications

Over the course of PIG2, students made significantly closer alignments to expert discourses when articulating the practical application of engineering principles. Although similarity values serve only as an approximation of full disciplinary discourses, they appear to be a useful measure of the positive effect of collaboration on learning. Moreover, from this analysis it is clear that similarity values can be used to provide rapid or real-time feedback to students or groups of students as they work toward expert discourse. Over the course of the videoconferencing collaborative activity, students' dialogue exhibited several shifts in its alignment to expert discourse. Similarity values prove to be a valid approach to display a simple visualization of the changes in the content of students' conversations. These shifts reveal some critical points where the discussions were more or less aligned with disciplinary discourse instructors would expect to see in these sorts of activities. Furthermore, we showed how similarity values can evidence whether students are able to document, within their final texts, the productive ideas that emerge during their discussion.

This initial implementation of LA in a CSCL environment serves to demonstrate two potential approaches for delivering rapid assessment prototypes to students and instructors. By exploring student-expert similarity values in both textual and video data, we focus on how students negotiate and construct discourses pre-, during-, and post collaboration, align to expert answers, and map to learning trajectories. Moreover, these promising results indicate where we can continue to refine our learning analytical approach to deliver finer-grained feedback. For example, the inclusion of other parts of speech such as adverbs in document-term-frequency matrices deserves a systematic exploration. Additionally, an exploration of other kinds of possible benchmarks with the purpose of capturing other valuable aspects within collaborative learning warrants further investigation. Finally, we are currently exploring what we call misconception benchmarks, which would help provide distance measures to possible misconceptions students may be falling prey to.

All in all, we believe that this preliminary case serves as a fruitful proof-of-concept that text-mining in online, inquiry-based learning can be used to provide rapid feedback to students and instructors during and after collaboration. This has the potential to foster learning and engagement in complex or semi-structured learning environments where students construct, negotiate, and implement disciplinary concepts. While powerful learning spaces, making timely sense of broad fields of disciplinary discourse may be inherently difficult in online CSCL environments. Text mining and the production of similarity values may provide a snapshot of student-expert alignment pre-, during-, and post-collaboration. However, an important question remains: How might stakeholders differently use similarity values as rapid prototypes of feedback?

Here, similarity values were used to assess student learning and collaboration. A vital next step in research is to investigate how similarity values are taken up and used by students, instructors, and researchers. For example, how do students reorient their discourse when provided with similarity values or qualitative representations of similarity values (e.g., word clouds)? How can instructors make use of similarity values to reorient delivery of instruction and understand the performance of collaborative student groups? Based on this work, similarity values can provide significant utility to instructors in cases where multiple groups of students are collaborating synchronously online (or in-person) or when making sense of large amounts of textual information. Finally, researchers can use the similarity values as a lens to investigate the collaboration, pinpointing moments of interaction worthy of deeper investigation. Practically, this may help organize trends in large amounts of data, providing the resources and convergent support for more nuanced analyses of discourse (e.g., conversation analysis or discursive psychology) in CSCL studies.

References

Berland, M., Davis, D., & Smith, C. P. (2015). AMOEBA: Designing for collaboration in computer science classrooms through live learning analytics. *International Journal of Computer-Supported Collaborative Learning, 10*(4), 425-447.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. Sebastopol, CA, USA: O'Reilly Media, Inc.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42. doi:10.3102/0013189x018001032

De Liddo, A., Shum, S. B., Quinto, I., Bachler, M., & Cannavacciuolo, L. (2011). *Discourse-centric learning analytics.* Paper presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge.

Jeong, H., & Hmelo-Silver, C. E. (2016). Seven Affordances of Computer-Supported Collaborative Learning: How to Support Collaborative Learning? How Can Technologies Help?. *Educational Psychologist, 51*(2), 247-265.

Kopainsky, B., Pirnay-Dummer, P., & Alessi, S. M. (2012). Automated assessment of learners' understanding in complex dynamic systems. *System Dynamics Review, 28*(2), 131-156.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.

Leydesdorff, L. (2005). Similarity Measures, Author Cocitation Analysis, and Information Theory. *JASIST, 56*(7), 769-772.

Ludvigsen, S. (2016). CSCL towards the future: The second decade of ijCSCL. *International Journal of Computer-Supported Collaborative Learning, 1*(11), 1-7.

Mercer, N. (2008). Talk and the Development of Reasoning and Understanding. *Human Development, 51*(1), 90-100. doi:10.1159/000113158

Rodríguez-Triana, M. J., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2015). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *British Journal of Educational Technology, 46*(2), 330-343.

Vatrapu, R., Teplovs, C., Fujita, N., & Bull, S. (2011). *Towards visual analytics for teachers' dynamic diagnostic pedagogical decision-making.* Paper presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge.

Xing, W., Guo, R., Petakovic, E., & Goggins, S. (2015). Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory. *Computers in Human Behavior, 47*, 168-181.