Designing Analytics for Collaboration Literacy and Student Empowerment

Marcelo Worsley[1], Khalil Anderson[2], Natalie Melo[3], JooYoung Jang[4]

Abstract
Collaboration has garnered global attention as an important skill for the 21st century. While researchers have been doing work on collaboration for nearly a century, many of the questions that the field is investigating overlook the need for students to learn how to read and respond to different collaborative settings. Existing research focuses on chronicling the various factors that predict the effectiveness of a collaborative experience, or on changing user behaviour in the moment. These are worthwhile research endeavours for developing our theoretical understanding of collaboration. However, there is also a need to centre student perceptions and experiences with collaboration as an important area of inquiry. Based on a survey of 131 university students, we find that student collaboration-related concerns can be represented across seven different categories or dimensions: Climate, Compatibility, Communication, Conflict, Context, Contribution, and Constructive. These categories extend prior research on collaboration and can help the field ensure that future collaboration analytics tools are designed to support the ways that students think about and utilize collaboration. Finally, we describe our instantiation of many of these dimensions in our collaborative analytics tool, BLINC, and suggest that these seven dimensions can be instructive for re-orienting the Multimodal Learning Analytics (MMLA) and collaboration analytics communities.

Notes for Practice

- Multiple studies provide ways to analyze collaboration, but most approaches are post-hoc and fail to provide feedback that might promote collaboration literacy.

- This paper details different dimensions of collaborations, as viewed by university students, as well as how to design a system to support learning collaboration literacy.

- The dimensions of collaboration and the example platform could help steer the direction of future collaboration analytics tools for student and teachers.

Corresponding author [1]Email: marcelo.worsley@northwestern.edu Address: Department of Computer Science, Northwestern University, 2233 Tech Drive, Mudd 3104, Evanston, IL, USA. ORCID ID: https://orcid.org/0000-0002-2982-0040
[2]Email: khalil.anderson2023@u.northwestern.edu Address: Department of Computer Science, Northwestern University, 2233 Tech Drive, Mudd 3104, Evanston, IL, USA.
[3]Email: nataliemelo2019@u.northwestern.edu Address: Department of Computer Science, Northwestern University, 2233 Tech Drive, Mudd 3104, Evanston, IL, USA.
[4]Email: joojang2019@u.northwestern.edu Address: Department of Computer Science, Northwestern University, 2233 Tech Drive, Mudd 3104, Evanston, IL, USA.

1. Introduction

Analyzing collaboration has been a topic of interest among researchers for nearly a century (Bos, 1937; Mayo, 1939). Early research in analyzing collaboration challenged researchers to employ new techniques and conduct painstaking observations of their participants. Many of these early analyses were motivated by the goal of better understanding the utility of collaboration relative to individual work. This new orientation towards the utility of collaboration marked an important shift from analyses of individuals towards analyses of the dynamics that reside within collectives. This shift helped spur the development of many sociological and distributed learning theories currently treated as common knowledge. Fast-forward to the 21st century: collaboration has become a global topic of importance. It sits among the oft-cited 21st century skills and has been the centre of many industry, academic, and governmental projects (Binkley et al., 2012). This renewed attention coincides with novel technological and analytic tools. These recent tools for studying and supporting collaborative settings have exponentially increased from the early experiments by Bos and Mayo. In many ways, however, the overall goal of much of the research on

collaboration has not experienced a complementary exponential shift. Much of the early work aimed to chronicle the ways that individuals collaborate, and understand the conditions associated with achieving improved outcomes through collaboration. Contemporary research takes on many of the same types of questions, by interrogating what enables groups to succeed (Borge & Rosé, 2016; Terken & Sturm, 2010; Worsley & Blikstein, 2018; Spikol, 2017). Armed with a novel set of multimodal sensors, researchers are able to study this question with a much more focused, yet broader reaching lens than mid-20th century researchers. Some researchers have begun to think more critically about how to influence the course of a collaborative experience through innovative interfaces and other orchestration strategies (Schneider et al., 2018; Dillenbourg et al., 2011; Berland, Davis, & Smith, 2015; Praharaj, Scheffel, Drachsler, & Specht, 2018). However, we propose that achieving the type of exponential changes in collaboration research, and collaboration analytics that many researchers want to see, requires an additional dimension that re-considers how we situate work on collaboration. As a complement to research that chronicles the collaborative experience through multimodal sensors, we resonate with the idea of advancing the notion of collaboration literacy (Worsley & Ochoa, 2020). Collaboration literacy speaks to the importance of being able to both read and respond to a collaborative setting. Worsley & Ochoa (2020) define collaboration literacy as

> The ability to ascertain and respond to changes in the quality of a collaborative experience. From the student perspective this amounts to being conscious of one's own contribution to a group, as well as the awareness and ability to intervene in order to ensure a strong collaboration. From the teacher perspective this includes awareness of how different groups are progressing, being able to respond to those groups in a timely fashion, and developing prompts and activities that afford good collaboration.

From this definition, particularly the part about the teacher perspective, it is clear that the goal of collaboration literacy is not to discredit the existing research on collaboration. On the contrary, advancing the field's understanding of collaboration processes and creating novel interventions and interfaces to support students and teachers is important. Instead, as we see it, the goal of positioning collaboration as a literacy extends current work on collaboration and collaboration analytics to more explicitly consider student development of collaboration skills and awareness. To this point, the current paper aims to answer the following research questions:

1. In what ways do students conceptualize their needs related to collaboration?
2. How might we instantiate those ideas in the design of collaboration analytics technologies?

We address the first research question by presenting our findings from a survey of 131 university students about their perceptions and experiences with collaboration. We suggest that students' prior experiences have contributed to their growth and development of different collaboration skills, but believe that there are important ways that researchers and practitioners can intentionally design to promote collaboration literacy. Concretely, we find that student perceptions and descriptions of collaboration can reasonably be grouped into seven categories, described in more detail later. We suggest that these categories are a primary contribution that other researchers and designers can use as they consider ways that their proposed and existing platforms provide a means to empower students. We answer the second research question by providing examples of how to incorporate the different dimensions into the design of collaboration analytics tools. Specifically, we draw examples from the design of our collaboration analytics tool, BLINC (Building Literacy in In-Person Collaboration). While we will not describe all aspects of the platform design, we will touch on several of the features that closely map onto the needs and preferences that students described within the seven categories. The BLINC tool also exemplifies ways that an innovative technical architecture combined with Multimodal Learning Analytics (MMLA) can be used to support the goal of growing collaboration literacy.

While student perspectives are a primary orientation for this project, we are also influenced by the novel body of literature that learning analytics and collaboration scholars have contributed to the field. Hence, the next section will present some of the relevant prior work that informs our approach. We then describe our survey study, and the analytic approaches used to interrogate that data. Afterwards, we move into a discussion of the categories that emerged from student responses and provide both descriptions and examples of each category. The discussion of categories is followed by a description of the BLINC platform that emphasizes ways in which it aligns with the seven categories.

2. Literature Review

The work we describe in this paper is heavily influenced by a collection of prior research on MMLA, educational dashboards, and teacher analytics, specifically.

2.1. Multimodal Learning Analytics: Tools, Techniques, and Relevant Constructs

Conceptually, this paper is heavily motivated by both the affordances and shortcomings of current work in MMLA. At a high level, researchers doing work with MMLA aim to examine and triangulate among a variety of modalities (e.g., audio, video, gesture, heart rate, electro-dermal activation, location, and eye tracking) to investigate the different ways that users may be demonstrating interest, engagement, or learning, for example.

Since its introduction, MMLA has had a strong connection to studying collaboration and collaborative problem solving. Much of the early work in this sub-field emerged out of the Mathematics Data Corpus (Oviatt, Cohen, & Weibel, 2013). The dataset provided researchers with access to audio, video, and digital pen data from groups of high school students completing a collection of mathematics problems. Some of the papers that used this dataset were able to demonstrate the ease with which predictions of domain expertise and knowledge of the answer could be determined using different features. For instance, Zhou, Hang, Oviatt, Yu, and Chen (2014) highlight the ways that the combination of multimodal data (written and spoken) could be used to predict the dominant domain expert within a group 100% of the time. This multimodal prediction greatly exceeded what could be garnered from an individual modality. Ochoa and colleagues did complementary work by surfacing the utility of relatively easy to extract features (Ochoa et al., 2013). Some of these features included the amount of time they used calculators, the writing speed, and the frequency that mathematical terminology is utilized, which are all effective for discriminating between domain experts and domain novices. While this work represented a strong first step for MMLA in support of studying collaboration, the learning context was limited, and the development of strategies that could be used to support student collaboration in real-time was missing. Instead, this work highlights features that are primarily relevant for researchers who may want to develop theories or tools to assess students participating in a collaboration, as opposed to finding ways to support student collaboration literacy.

Worsley and Blikstein (2015) expand the types of learning settings under consideration by studying pairs of students as they complete hands-on engineering design challenges. They also introduce specific learning interventions and examine how those treatments impact the overall process of the collaborative engineering design task. Like the previous work on MMLA, Worsley and Blikstein triangulate among several modalities: speech, gesture, and skin conductance. This multimodal data is used to construct prototypical multimodal engagement states that serve as the basis for interrogating each group's design process. This work, though, while having expanded to a hands-on learning context, once again focuses on post-hoc analysis and provides little in the way of direct feedback to the participants based on their group interactions.

Recent work by Cukurova, Luckin, Millan, and Mavrikis (2018) proposes non-verbal indexes for students' physical interactivity as a means to study collaborative problem-solving among high school and undergraduate students completing hands-on tasks. The paper specifically examines measures based on head direction and hand position to represent synchrony, accountability, equality, and intra-individual variability. Importantly missing from this analysis of collaboration practices, however, are the actual verbal cues, which the authors acknowledge can be central to studying group interactions. Beyond this, the study describes a platform that is utilized post-hoc, leaving the creation of a real-time system for future development.

The three papers mentioned here are a small subset of the many papers from the MMLA community that address collaboration-related constructs (Cukurova, Luckin, Mavrikis, & Millan, 2017; Cukurova et al., 2018; Schneider et al., 2018; Schneider & Pea, 2015; Worsley & Blikstein, 2018, 2015; Martinez-Maldonado et al., 2017; Martinez-Maldonado et al., 2019; Echeverria, Martinez-Maldonado, Shum, & Sydney, 2019; Vrzakova, Amon, Stewart, Duran, & Mello, 2020; Oviatt & Cohen, 2014). Collectively, prior work in MMLA and collaboration analytics highlights the many ways that MMLA can be used to identify and extract collaboration-relevant constructs. These constructs range from being individual and group-based, from unimodal to multimodal, and from literal to semantic (Worsley & Ochoa, 2020). However, one reality across most of this prior research is that the analyses are almost exclusively conducted after the completion of the activity (Worsley & Ochoa, 2020; Wise & Chiu, 2011). Hence, there is no closing the loop with the data being provided. While, there are instances of interactive table tops (Tse, Greenberg, Shen, & Forlines, 2007), or gesture-based systems that provide real-time information (Oviatt et al., 2013) about how much each student is participating, none of these platforms have really been developed with the goal of improving collaboration literacy. Instead, many existing tools were created to advance theories about collaborative problem solving, improve the quality of student work on the specific task, or provide insights for teachers. In contrast, the current paper primarily considers the need to help students develop collaboration literacy, which has not been extensively studied within the MMLA community.

2.2. Educational Dashboards and Human-Centred Learning Analytics

While not explicitly about supporting collaboration literacy, Reimers and Noevesky (2015) aim to centre student use of learning analytics dashboards as an important area of inquiry. Their specific study involved surveys of K–12 and university students about which features should be included within the student view of a learning management system (LMS). Student survey responses highlighted the importance of features like automatic assignment reminders, the ability to compare their progress with their peers, and integration with social media platforms. Given the specific context of an LMS and a fairly broad set of platform goals, Reimers and Noevesky (2015) do not provide much in the way of specific design considerations. However, their work is an early effort in a larger literature around human-centred learning analytics and the design of educational dashboards, and brings forth the importance of student voice.

More recent work has explicitly broader consideration of the different stakeholders within learning analytics (Alhadad, 2018; Teasley, 2015; Echeverria et al., 2018). Researchers have drawn attention to bridging between the theoretical and the

practical. Alhabad refers to this distinction as visualizing data as a means for communication versus visualizing data as a research methodology. Especially in the case of visualizing data as a means of communication, it is important to leverage prior research from cognitive psychology that describes potential challenges with human attention and general graphic design best practices for focusing user attention (Williams, 2015). Echevarria et al. (2018) utilize the language of explanatory and exploratory visual learning analytics. Namely, they position researchers as engaging in exploratory work, and students and practitioners utilizing learning analytics for explanatory purposes. In line with the emphasis to support explanatory learning analytics, they discuss the importance of task alignment, and the contextual visibility of information. When students and practitioners engage with a dashboard, they typically do so with specific tasks in mind, and may only need certain information in specific contexts. These ideas generally align with a best practice around minimalism and supporting the ability for users to tell a story with the data (Williams, 2015; Alhadad, 2018; Echeverria et al., 2018). This idea of storytelling also speaks to the importance of supporting multiple interpretations of the data as well as avoiding generalizations and equating causation with correlation (Alhadad, 2018). These are all important considerations when designing and implementing educational dashboards that inform this work.

2.3. Teacher Analytics

Echeverria et al. (2018) also pushes into the domain of teacher analytics, a growing sub-area of learning analytics (LA) that aims to enable teachers to better understand and support their learners across a variety of learning configurations (Teasley, 2015). This work is dually motivated by the challenges posed by teaching large classes, and the availability of increasing amounts of data within learning management systems (Carbone, 1999; Dillenbourg et al., 2013; Gibbs, Jenkins, & Alan, 1992). While the current paper focuses on responding to student needs and preferences, the larger goal of collaboration literacy reasonably involves both students and teachers. To this point, teachers can play an important role in helping students develop collaboration skills. Hence, work to support teacher interactions with collaboration data and visualizations is also relevant to the current work. Additionally, many of the insights gathered from research on teacher analytics can be informative for thinking about the student experience and how student data is represented. However, as prior researchers have suggested (Van Leeuwen, Janssen, Erkens, & Brekelmans, 2015), the actual creation and study of these types of tools, for teachers or students, has scarcely been discussed within the LA community, even less so in the context of studying or supporting collaboration.

Van Leeuwen provides an early attempt to address this gap by conducting an experimental study with 40 pre-service teachers who were tasked with responding to different collaboration scenarios. More specifically, the paper describes two tools that were implemented to give the pre-service teachers a window into group progress in the form of the Concept Trail and Progress Statistics. Concept Trail tracks the specific topics discussed within a given group, and puts that data on a timeline. Progress Statistics looks at how many words students have entered into the text editor so that the teacher can ascertain how much progress students have made on the assignment. In their study, the authors found that pre-service teachers using these two tools were generally more responsive to students.

However, they did not see noticeable differences in pre-service teachers' abilities to identify problematic group interactions. Moreover, they found that since the Concept Trail was specifically providing information about the topics being discussed, pre-service teachers in the learning analytics condition (i.e., those who used the Concept Trail) tended to focus on conceptual problems while overlooking certain social indicators of collaboration challenges. Some of these factors may be influenced by the inexperience of pre-service teachers in the field. Even so, this analysis provides a number of relevant insights that inform our work.

First, the information provided certainly has an impact on what teachers and students will pay attention to. In order to ensure a smooth interaction across the different levels of collaboration, we have to be cognizant of what and how to present that information. Second, this study reiterates the importance of collaboration literacy in that providing data does not necessarily result in greater awareness of inequitable or otherwise unproductive collaboration practices. Third, the Van Leeuwen et al. (2015) study was conducted using fairly simple measures of collaboration quality. This is both promising and an opportunity for improvement. It is promising from the perspective that even the simplest measures can change the nature of participant behaviour during a collaborative experience. However, consistent with Worsley and Ochoa's (2020) taxonomy, Van Leeuwen et al. (2015) primarily focuses on very literal, individual level measures. While these can likely be used in beneficial ways, the need is to consider more complex patterns of collaboration through multimodal data. Moreover, as we will see from our analysis of student perceptions in the following section, many important dimensions of the collaboration experience are being overlooked within these simple representations. Hence, we position the current work as addressing an important gap in the kind of information being provided to both students and teachers about the collaboration process.

3. Methods

3.1. Participants

Participants in this study were 131 university students from a private university in the United States. Most participants (124) were undergraduate students, two were masters students, and five were PhD students. First- and second-year students represented the majority of the pool, with 40 participants in each year. Twenty-two of the students were in third year and 26 were in fourth year. Two of the graduate students were in their fifth year. We received completed surveys from 87 of the 131 participants; however, partial responses are also included within this analysis. Individuals received $5 compensation for finishing the survey and were also entered into a raffle for five $50 gift cards.

3.2. Survey

Each student was asked to complete a survey that posed questions about their perceptions and prior experiences with collaboration. Students answered two Likert-scale (questions 1 and 2) and six text-response questions about their collaboration experiences (questions 3–8):

1. How often do you collaborate on classwork in school settings (inside of classrooms, for class)?
2. How often do you collaborate on classwork outside of school settings (work done outside of the classroom, etc.)?
3. How would you describe a good collaboration?
4. What are some characteristics of a bad collaboration?
5. How do you deal with poor collaboration experiences?
6. What are some things you've been taught about collaboration and collaborating?
7. What advice would you give a high school student about how to collaborate most effectively in college?
8. In your ideal world, how would you want to learn how to collaborate?

Students also had the option of responding to a ninth question: If you have any other comments on collaboration, please provide them here. Thirty-one students responded to this optional question. Aggregate statistics from the two Likert scale questions will be briefly presented, complemented by an analysis of the free-response questions. The six free response questions required students to enter at least 75 characters each; they form the bulk of our analysis.

3.3. Likert Scale Question Analysis

To the two Likert scale questions included in the survey — "How often do you collaborate on classwork in school settings (inside of classrooms, for class)?" and "How often do you collaborate on classwork outside of school settings (work done outside of the classroom, etc.)?" — students could respond Never, Sometimes, Regularly, or Everyday. To compare results from these two questions, responses were coded into ordinal values from 1–4 (Never = 1, Sometimes = 2, Regularly = 3, Everyday = 4). This order matches that of the response options presented to participants. A two-tailed Wilcoxon signed-ranks test was used to compare the mean response value, specifically to determine if students reported collaborating at different rates in class versus outside of class.

3.4. Free Response Question Analysis

As learning analytics enthusiasts, we began our text analysis by leveraging different computational techniques for mining student responses. For example, we developed word clouds of term frequencies for the different questions. We also used topic modelling to examine any underlying and hard to observe models that could help inform our analysis. However, we found that these two approaches, as implemented, were insufficient for the level of specificity we wanted to include in the analysis. Instead, we borrowed techniques from grounded theory (Glaser & Strauss, 1967) in analyzing student responses to the free response questions. We supplemented this human coding with a quick text analysis of common words associated with each category.

The data was coded by four team members. The first step was for individual team members to highlight student responses that surfaced one or more key ideas. From these initial exemplars, we identified a set of overarching themes. These initial themes included broad topics such as Emotion, Communication, and Equal Work. After some discussion, a preliminary set of categories was proposed. This preliminary set included Content, Context, Climate, Communication, and Contribution. Group members met to collectively code a selection of 10 responses taken from different questions. The team reached consensus on those responses but refined the categories to be the following: Communication, Compatibility, Conflict Resolution, Context, Contribution, and Constructive. The Communication category combined Communication and Climate, Constructive replaced Content, and Compatibility was added. The four research team members each coded 300 of the 567 total responses. After individually coding 300 responses, the research team met to reassess the list of categories and proposed a few modifications. Namely, the Climate category was re-added, and Conflict Resolution was changed to simply Conflict. Next, the research team completed an additional round of coding based on updated categories and their definitions. Finally, each team member reviewed all of the codes assigned to each response and added and removed codes until there was a consensus among the

research team for all student responses. Hence, the data that we report in this paper comprise the codes that the entire research team agreed were applicable to each student response.

4. Findings

4.1. Likert Scale Question Analysis

The two Likert scale questions about the contexts where students practise collaboration provides the starting point for this analysis. Students were independently asked about how frequently they collaborate in classrooms, and outside of the classroom. Students could respond Never, Sometimes, Regularly, or Everyday. On average, students noted that they collaborate more frequently outside of class than they do in class ($Z=-3.67$, $p<0.001$). This, as we discuss later, has important implications for the design of collaboration analytics technologies and where the tools are deployed.

4.2. Free Response Question Analysis

Moving now to the free-response questions, our iterative coding process surfaced seven overarching topics. We suggest that these topics represent the ways that students perceive and experience collaboration. These topics include Climate, Communication, Compatibility, Conflict, Context, Contribution, and Constructive. The research team reached a consensus that these seven topics cover the broad categories that students mentioned in response to our survey questions. As we describe each category or dimension, it is important to note that they are not meant to be orthogonal, or mutually exclusive. On the contrary, as the reader will see, many of the categories reflect, or are influenced by, other categories. Even so, each category does appear to have several characteristic sub-dimensions. We do not consider these sub-dimensions to be predictors of collaboration quality; instead, they reflect the ways that students may see and interpret a given collaborative experience. Our presentation of each dimension will follow the same format. First, we present a short description of the category followed by a set of terms that most frequently occur within responses coded with this dimension. Next, we include an exposition of the different sub-dimensions of that category as evidenced through one or more quotations. Finally, we discuss the ways that the given dimension is discussed, or absent from, prior research on collaboration and/or learning analytics.

4.2.1. Climate

The Climate dimension refers to emotional or affective considerations of the collaboration. When we look at the most frequently occurring words from responses that received this code, a few exemplify the idea of Climate: ideas, open, help, try and respectful. Notably, students are thinking about the extent to which they are able to openly share ideas, in a respectful and welcoming environment. For example, one student wrote that bad collaboration can be characterized as an "overall toxic atmosphere." Another student described it as "Feeling like I want to withdraw. A place that I don't feel welcome." In contrast, they described good collaboration as a space where "everyone is given a fair chance to voice their opinion and does not feel scared to give that opinion on the basis that others will make fun of them."

Consistent throughout this category is the realization that feelings, emotions, and affective states are important parts of a collaborative experience. Our category of Climate bears some similarity to psychological safety climate (Edmondson & Lei, 2014). Edmondson and Lei (2014) refer to psychological safety, or a psychological safety climate, as representing how comfortable people feel taking interpersonal risks within a given context. Psychology researchers have demonstrated ways that the psychological safety climate is linked to several different measures of team performance and organizational change. Outside of the psychology community, several LA researchers have talked about the inclusion of emotion, or affect related measures, when studying groups of collaborators. For example, Vrzakova et al. (2020), Worsley and Blikstein (2018), and Furuichi and Worsley (2018) all utilize electro-dermal activation wristbands that can proxy for feelings of stress or arousal. These measures, however, can be challenging to disentangle from correlated measures like cognitive load. Others have utilized text mining, albeit post-hoc, to surface positive and negative sentiment in spoken utterances (Worsley & Blikstein, 2011). A primary challenge, then, is how the field might look to include a more complex set of affective states without being subject to the bias present in many perceived affective state recognition tools (Marsh, Elfenbein, & Ambady, 2003). Additionally, there is the challenge of conflating the different multimodal signals with felt emotions.

4.2.2. Communication

The Communication dimension refers to the quantity and quality of information shared among group members participating within a given collaboration. Terms frequently used in responses labelled as Communication included the following: communication, everyone, talk, share, open, and visualization. The words "communication" and "talk" highlight ways that this category generally includes the simple act of talking to one another when problems arise, or when people are uncertain about next steps, for example. In a more concrete sense, responses coded as Communication often included people establishing and being aware of expectations. For instance, one student wrote, "Constant communication is important to make sure everyone is on the same page about goals and expectations." Another student writes, "Everyone must be open to compromise and people should try to have open communication." Open communication also touches on another key idea: listening. Connecting these

two ideas, one student wrote that "Collaboration requires active listening and understanding from both sides." Finally, several responses also referred to visual representations of data as a way to communicate the information to students.

Communication is the area where prior research has been most prolific. Many projects have endeavored to quantify talk using speech signals (Dillenbourg & Bachour, 2010; Bassiou et al., 2016) or chat-based messages (Vrzakova et al., 2020). These systems can make approximations of speaking turns and use body language to proxy for participant attention. For instance, Vrzakova et al. (2020), Worsley and Blikstein (2018), and Cukurova, Zhou, Spikol, and Landolfi (2020) use head-pose to identify moments when participant attention seems to be situated away from their collaborator. These existing platforms have done this as part of post-hoc analysis. Other research explicitly measures visual attention using mobile eye trackers (Schneider et al., 2018). Albeit prohibitively expensive, mobile eye trackers can monitor and focus participant attention.

4.2.3. Compatibility

Compatibility refers to how well group members use complementary working and interaction strategies. Common compatibility words included the following: working, together, team, help, and friends. As these terms suggest, students think about how well people in the group work together and help one another. Other responses noted ways that students may have "compatible work styles." Importantly, though, students did not simply see compatibility as static. Instead some indicated an awareness that "team building and bonding activities give people a sense of closeness and community." The term "friend" arose from student responses that highlighted the advantages of working with and without friends. In giving advice to younger students, one respondent wrote, "Find people who work well together, these people may not necessarily be your closest friends." Another student advised "be friends with your teammates," suggesting an awareness that compatibility (or friendship) is, again, something that can be developed.

In terms of prior research, to our knowledge, compatibility has not been a focal topic within research on collaboration in education, or collaboration analytics. Research in pair programming (Grover et al., 2016) and group formation strategies (Boaler, Wiliam, & Brown, 2000) is likely the most pertinent to this topic, but seldom does the work explicitly discuss team building. There is a complementary body of research on strategies for team building, but we scarcely see this work discussed within the collaboration or MMLA communities. Part of this may be connected to the contexts we typically use to study student collaboration, something that we discuss in more detail shortly.

4.2.4. Conflict

The Conflict dimension encapsulates student approaches and needs for responding to challenging or contentious situations that might arise during a collaborative experience. Some of the most common words associated with this category include the following: conflict, different, bad, team, and professor. One student's response, in particular, seems to encapsulate frequently mentioned approaches for dealing with conflict: "I usually switch up which people I am collaborating with if possible. If not, I will confront the responsible people in the group or else just deal with it." Hence, the term "different" was often used to describe a common approach that students take when faced with conflict: "finding a different team." In addition to exiting the collaboration, students talked about ways for confronting the individual(s) responsible, "just dealing with it," and bringing their concerns to the course instructor or external mediator. On this topic, one student voiced that, in the ideal world they would like, "having experienced educators facilitating collaboration would be nice, someone being there to make sure everyone is heard and gets a chance to speak, also sort of a conflict mediator." Another student similarly noted, "If conflict arises, it would be helpful to have someone more experienced mitigate this to show how issues can be resolved." This is a significant observation because the student is not satisfied with simply having the conflict resolved. Instead, they are interested in learning how to address these types of challenging situations.

Existing research on the topic of classroom orchestration (Dillenbourg et al., 2011; Roschelle, 1999) begins to scratch the surface of this dimension by alerting instructors to the needs of students. Other research has documented the emergence of conflict in collaborative situations (Jung, Toprani, Yan, & Borge, 2017; Roschelle, 1999; Barron, 2003), but little of the leading research on collaboration analytics appears to think about conflict as an important dimension within the collaboration landscape. Put differently, researchers have approaches for identifying that conflict has arisen, but this is seldom included within collaboration analytics tools, nor are tools developed to support conflict resolution.

4.2.5. Context

Context refers to considerations of who is participating within a given group, why they are participating, and where the collaboration takes place. Some of the words associated with Context responses include learn, experience, problem, different, and taught. "Problem," refers to learning collaboration on real-world problems. "Learn" and "taught" are emblematic of students wanting to learn collaboration more explicitly. One student notes, "In an ideal world, there would be a mandatory class teaching students how to collaborate using research-backed ideas." Part of what this student is suggesting is that collaboration is never explicitly taught as an end goal, but is instead always treated as a secondary skill. For instance, one student responded, "I would want to learn general principles in collaboration through some form of class/lecture, but focus

more on actually collaborating with other people and learn from experience." Context also touches upon questions of why and where students learn to collaborate.

Students consistently mentioned wanting to learn collaboration across a variety of settings. For example, a student responded, "I think [collaboration] is best learned by doing it both in class and out of class." Another student echoed that sentiment but proposed another dimension of Context that is not explicitly in or out of class. The student wrote, "I would want to learn how to collaborate in a setting where I feel comfortable, since sometimes I am shy and being forced into collaboration can be counterproductive." This is another important consideration since the Context can oftentimes influence how willing people are to contribute. To this point, one person noted that they would ideally "[a]lways do it with friends, where you can communicate your needs and ideas comfortably. Strangers are frustrating." A final sub-dimension of context has to do with the ways that actions and interactions within a collaborative setting may be interpreted. One student's comment makes this point explicitly. The student wrote, "oftentimes outside perspectives on group collaborations do not reflect the opinions of the collaboration members." This student is concerned with how an external person or tool may interpret what is happening within that group. Namely, that an external evaluator may make judgments about a group that reflect a lack of understanding of the context for those actions. Accordingly, the context for the collaboration has important implications for how these tools and research studies are conducted. It also has implications for how we choose to surface information to collaborators or instructors.

With regard to Context, much of the existing conversation within the MMLA community has been in thinking about the diverse places where collaboration occurs. Compared to some of the early MMLA research that looked at a very decontextualized mathematics learning environment (Oviatt & Cohen, 2014), more recent work has begun to move outside of laboratory environments and into Fablabs and Makerspaces (Schneider & Blikstein, 2015; Spikol, Avramides, & Cukurova, 2016; Worsley & Blikstein, 2018). Within many of these studies, students are solving open-ended learning problems with tangibles and microcontrollers, which necessarily shift the context away from a mandated set of tasks, and more towards self-directed tasks. Other work has moved away from traditional areas of education towards hands-on healthcare learning environments (Martinez-Maldonado et al., 2017). Less common, however, are studies and platforms designed for user-selected contexts. This need for user-selected contexts is corroborated by student survey responses that found them reporting more time collaborating outside of class than in class. Ochoa's multimodal selfie (Domínguez, Chiluiza, Echeverria, & Ochoa, 2015), a mobile device with camera and microphone, could satisfy these cross-context needs, although it was designed with a much more individualistic perspective in mind.

4.2.6. Contribution

The Contribution dimension highlights what individual participants are, and are not, bringing to the collaboration. Among the most common words for this dimension were work, ideas, contribute, make, time, effort, and project. Students talked about contributions along two main sub-dimensions. First, there were examples of responses related to verbal contributions to group discussions, in terms of ideas and solutions to problems. At times these were in regard to people adding ideas. In other instances they had to do with ideas being discredited. For example, a student talked about bad collaboration involving "[o]ne person dominating, ideas being shut-down." In other instances, contribution has more to do with "work" or "effort" towards a group project. To this point, another student wrote, "a good collaboration involves all parties contributing equally to complete a task to a degree that satisfies all parties involved." These two examples surface the existence of people over-contributing and under-contributing, themes that pervade the survey responses.

Much of the prior research on collaboration, and collaboration analytics itself, looks at this dimension (Borge & Rosé, 2016; Tse et al., 2007; Worsley & Blikstein, 2018). Students' verbal and gestural contributions to a given collaboration are quantified and represented for students to reflect upon (Praharaj et al., 2018; Dillenbourg et al., 2011). In the case of collaborative problem solving, researchers may wish to uncover who contributes the eventual solution or idea that the group eventually used. Dependent measure aside, assessing student contributions has been a driver of collaboration analytics.

4.2.7. Constructive

The Constructive dimension relates to how students conceptualize or discuss the overall goals of collaboration. Some of the words commonly used in responses about this dimension include ideas, team, goal, help, make, and better. This dimension typically aligned with the idea of team progress towards individual or group learning goals or project outcomes. To this point, one student wrote, "Good collaboration is mutually beneficial. Every individual involved in the collaboration has something unique to contribute, and the net result is that every person gains insights that they would not have had on their own through the collaboration."

The end goals of a collaboration can vary widely based on many of the other dimensions previously mentioned, Context, in particular. Much of the prior research on collaboration in engineering designs contexts has endeavoured to chronicle how well project teams progress towards a specific final project. This work will often involve documenting the different phases of collaboration a group is in based on the types of discussions they are having. For instance Gero and Kannengiesser's (2004) work on Function-Behavior-Structure mapping has been applied to studying groups during collaborative problem solving. The

orientation towards products or outcomes is also evident across other research on collaborative problem solving (Hmelo-Silver & Barrows, 2008; Praharaj et al., 2018). In fact, many prior research studies have the objective of improving student learning or helping them produce better artifacts (Schneider et al., 2018; Worsley & Blikstein, 2018; Spikol, Superiore, Anna, & Landolfi, 2020; Martinez-Maldonado et al., 2017), but are not necessarily concerned with team dynamics or how much each person is contributing.

4.2.8. Dimension summary

We conclude this section with the relative frequency of student responses associated with each category (Table 1), recalling that multiple codes may be applied to a given response. We also present a co-occurrence matrix of the different dimensions.

Table 1. Percentage of Responses Coded with Each Dimension

	Percentage
Climate	22%
Communication	50%
Compatibility	20%
Conflict	15%
Context	26%
Contribution	43%
Constructive	35%

References to Communication were the most prevalent (50%), followed closely by Contribution (43%) and Constructive (35%). Context was mentioned in just over a quarter of the responses (26%). Meanwhile, Conflict (20%) and Compatibility (15%) were the least frequent. In reviewing these percentages, we want to reiterate that our goal in presenting these dimensions is not to definitively say what students think is more or less important. Instead, our objective is to identify a set of dimensions that span student perspectives and prior experiences with collaboration. The distribution and relative importance of each dimension may vary based on context and student population. That said, given the prevalence of responses coded as communication and contribution, it is not surprising that much of the prior research on collaboration analytics focuses on these two areas (e.g., D'Angelo et al., 2019; Blanchard et al., 2015). As an additional point of consideration, a quick glance at the percentage of responses assigned to each code should make it clear that many multiple dimensions were applied to most of the student responses. To quantify this, we looked at the number of times each code was the only code represented within a given student response and the frequency that different dimensions co-occurred with one another. Data describing the number of times a given category was the sole category for a given response is summarized in Table 2. We present this information both in terms of the raw count and as a percentage of the number of response codes with that category. In total, approximately 25% of student responses were coded with just one category, with Context being the most frequent (29%) in terms of count and percentage. In contrast, nearly all (~97%) of the 124 responses coded as Climate co-occurred with other categories.

Table 2. Number and Percentage of Responses by Dimension
(With only that dimension code applied)

	Count as single code	Total coded	Percentage
Context	43	149	29%
Contribution	40	245	16%
Communication	27	284	10%
Constructive	13	200	7%
Compatibility	9	112	8%
Climate	4	124	3%
Conflict	4	82	5%

Data about which dimensions most frequently co-occurred with one another (Table 3 and Table 4) provide some additional insights about the nature of these categories. Table 3 highlights the count of pairwise co-occurrences between different dimensions. From this table we see that the two most frequently occurring categories, Communication and Contribution, also frequently co-occurred with other categories. Table 4 looks at pairwise co-occurrences, but as a percentage of a given category's total count. This gives a more nuanced picture of code co-occurrence. For example, as we look across the row labelled Communication, we see that Contribution (48%) and Constructive (41%) most frequently appeared alongside the

Communication code. We also see that while Communication was generally the highest co-occurring dimension for most other categories, its relative rate of co-occurrence with Compatibility lags behind both Contribution and Constructive. Looking across the row labelled Compatibility, we see that Contribution and Constructive appeared in 41% of the responses coded as Compatibility, while Communication was 34%. This might suggest that Communication is less essential to Compatibility than the dimensions of Contribution and Constructive. Similarly, one might look at the Conflict row and note that Communication appeared alongside approximately three-fourths of the responses that included Conflict. This could corroborate the general idea that communication is a primary component of conflict resolution. There is, potentially, a lot to be gleaned from examining these tables. However, an in-depth discussion of these interactions is beyond the scope of this paper. Our purpose in noting the co-occurrence of these categories, however, is to reiterate that they are not mutually exclusive, and to motivate the idea that a single interface design element may simultaneously address, or reflect, multiple dimensions. To explicate this point and answer our second research question, the next section will outline some concrete strategies for instantiating these different dimensions into the design of a collaboration analytics platform. The strategies that we described are based on the BLINC platform and will be presented alongside a discussion of how each element aligns with ideas of collaboration literacy.

Table 3. Count of Pairwise Dimension Co-Occurrence

	Climate	Communication	Compatibility	Conflict	Context	Contribution	Constructive
Climate	124	75	28	19	30	44	56
Communication	75	284	38	61	42	135	112
Compatibility	28	38	113	10	23	46	46
Conflict	19	61	10	82	16	31	24
Context	30	42	23	16	149	34	35
Contribution	44	135	46	31	34	245	81
Constructive	56	112	46	24	35	81	200

Table 4. Frequency of Dimension Co-Occurrence re. Total Number of Responses Code by Dimension
(Each row corresponding to proportion of dimension co-occurring)

	Climate	Communication	Compatibility	Conflict	Context	Contribution	Constructive
Climate	100%	6%	23%	15%	24%	35%	45%
Communication	26%	100%	13%	21%	15%	48%	39%
Compatibility	25%	34%	100%	9%	2%	41%	41%
Conflict	23%	74%	12%	100%	2%	38%	29%
Context	2%	28%	15%	11%	100%	23%	23%
Contribution	18%	55%	19%	13%	14%	100%	33%
Constructive	28%	56%	23%	12%	18%	41%	100%

4.3. BLINC Design

There are many ways to instantiate the above dimensions within a given collaboration analytics tool. Moreover, many prior research projects in teacher analytics, classroom orchestration, and MMLA include design components that resonate with the dimensions that emerged from student responses. The examples that follow are intended to surface some of the ways we have integrated these dimensions into the design of our BLINC platform.

At a high-level, the BLINC platform consists of a distributed system of wireless audio/video data collection devices, a centralized server, and a secure, web-based interface for accessing data in real-time or post-hoc. The tool is designed to simultaneously be useful for researchers, practitioners, and students. Embedded within this tool are a variety of capabilities that can provide useful and actionable information for different stakeholders. We will not describe all of these technical components in detail within this paper, but will present a general overview of the platform, followed by the description of two core capabilities that align with three of the categories: Climate, Communication, and, Context. As we discuss each feature, we also describe ways that it might be used to support collaboration literacy. We recognize, however, that effective use of these features towards collaboration literacy takes more than simply including them within the interface. Instead, there needs to be a complementary set of intentionally designed experiences for students that incorporate this data.

4.3.1. BLINC Interface Overview

The participant-facing component of BLINC consists of four main views. All of these views are user specific and require that

the user log into their account before gaining access to data. The first view after logging in is a user-specific home screen that allows the user to adjust to different configurations (Figure 1). Some configurations include defining relevant keywords, uploading documents to create topic models, configuring WiFi settings, and organizing previous data collection sessions. From this home screen, users are also able to look at previous data-capture sessions or start a new data-capture session by selecting the Discussions button.

After clicking the Discussions button, and starting a new session, users are given the option to specify keywords and a topic model to be associated with this new session. Users can subsequently start synchronous data collection across any available data collection devices. This includes microphone arrays that run custom software to process multi-dimensional audio and video data or smartphones that are simultaneously logged into the BLINC interface through the user's account. After selecting the appropriate devices and starting data collection, users are taken to a screen where they can selectively look at the real-time data coming from each device. This group-specific data is presented using four primary representations (Figure 2). The data depicted within these representations can be adjusted using the Timeline control at the top of the interface. Adjusting this slider lets the user specify which time interval they want to review.
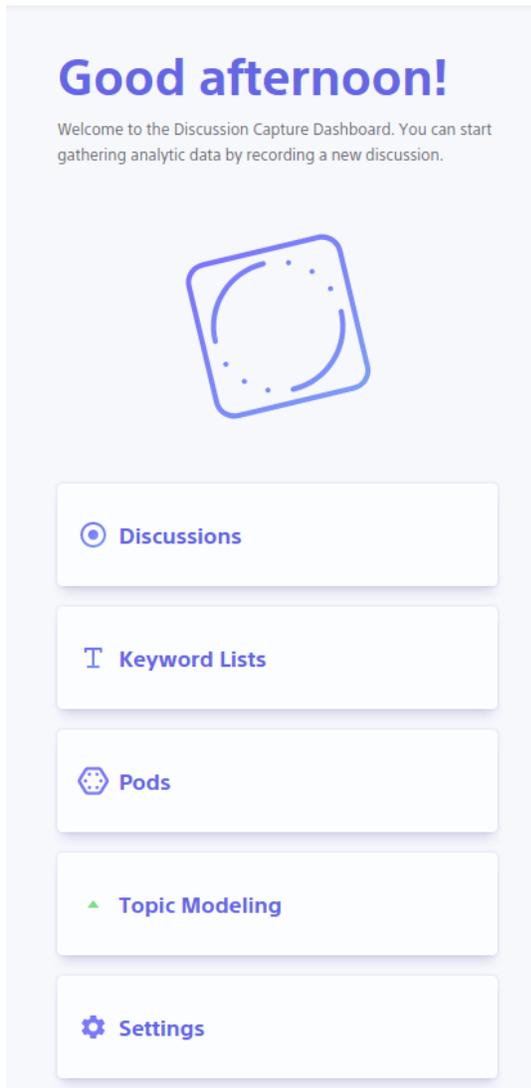

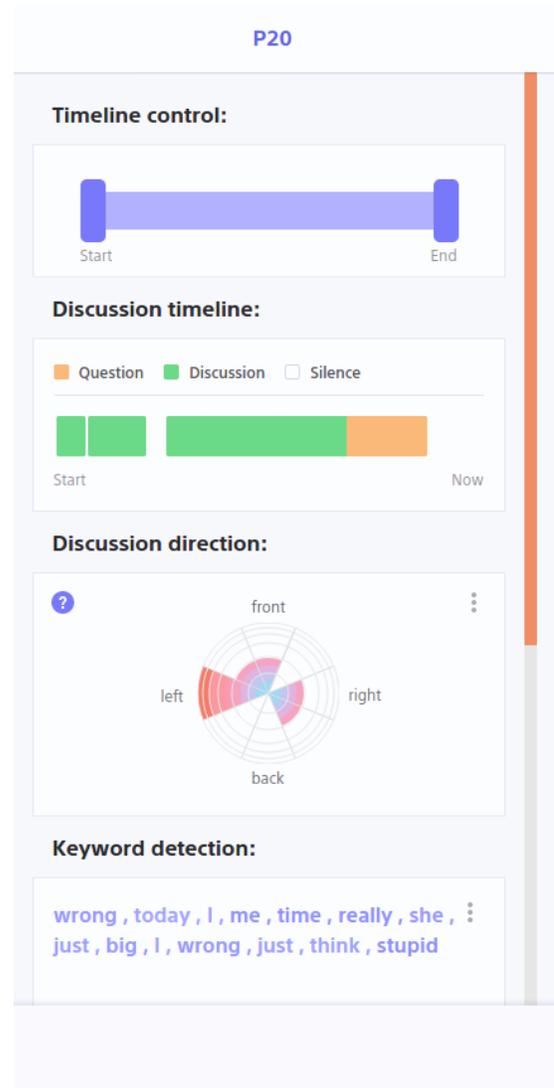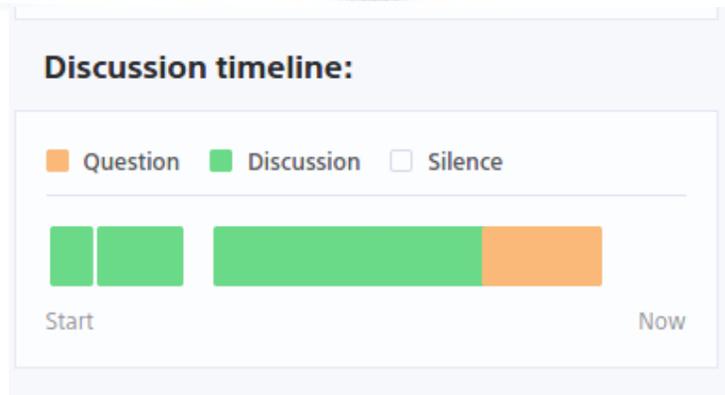**Figure 1.** BLINC platform home screen.

Figure 2. BLINC group discussion view.

The first representation is of a discussion timeline that depicts an approximation of when the group was discussing, asking questions, and being silent (Figure 3). This data is derived from automatically transcribed audio data.
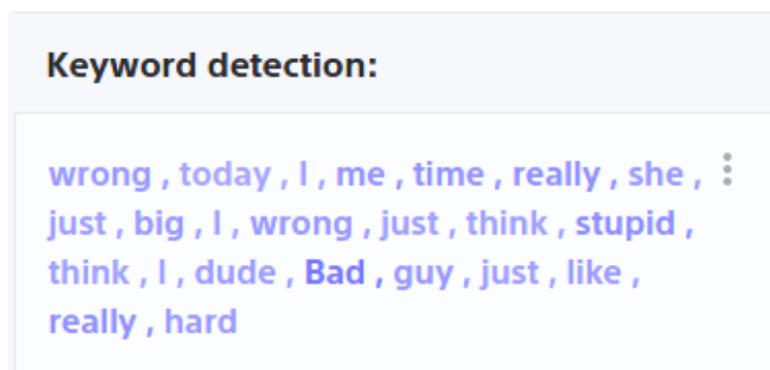
Discussion timeline:

☐ Question   ☐ Discussion   ☐ Silence

Start   Now

Figure 3. Discussion timeline component.

The second representation is called "Discussion direction" and provides the directional of audio arrival captured by the microphone array (Figure 4). This specific representation can provide an approximate distribution of participant talk within the group. This representation can be adjusted to different levels of granularity. For example, it can divide the 360 degree space surrounding the microphone array into 2, 4, 6, 8, or 12 equal-sized parts.

Discussion direction:

front

left   right

back
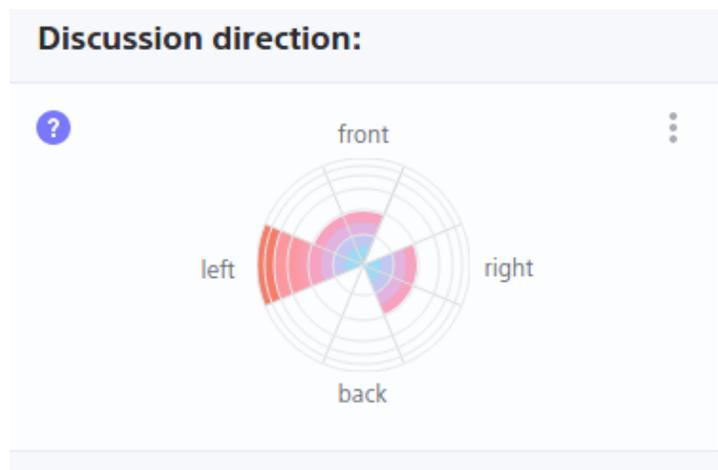
Figure 4. Discussion direction component that describes the direction-of-arrival.

The next data representation is of the keywords that have been stated (Figure 5). Recall that the desired set of keywords were something that users could select prior to starting the discussion. Each occurrence of a keyword is highlighted within the "Keyword detection" component.

Keyword detection:

wrong , today , I , me , time , really , she , ⋮
just , big , I , wrong , just , think , stupid ,
think , I , dude , Bad , guy , just , like ,
really , hard

Figure 5. Customizable keyword detection component.

The fourth representation is a customizable set of sentiment classes, labelled "Discussion features" (Figure 6). This component appears just below the keywords at the bottom of the main view. Automatically transcribed audio is used to conduct basic sentiment analysis by leveraging extended forms of the Harvard General Inquire (Stone, Bales, Namenwirth, & Ogilvie, 1962) and the Linguistic Inquiry Word Count (LIWC; Tausczik & Pennebaker, 2010). Hovering over any of the sentiment categories provides a tooltip that specifies which words are associated with that specific sentiment class.
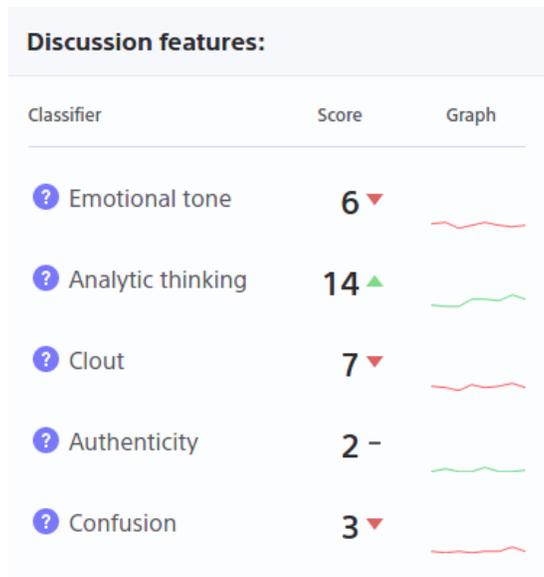

Figure 6. Customizable discussion features component.

Across all of these four representations, users are able to click on the desired items, and will subsequently be taken to the fourth primary view, which includes a complete transcript of audio session data from that specific audio capture device (Figure 6). When entering the transcript view from any of these four representations, the transcript immediately loads into the corresponding context of the selected keyword, sentiment, direction of arrival, or question segment, for example.
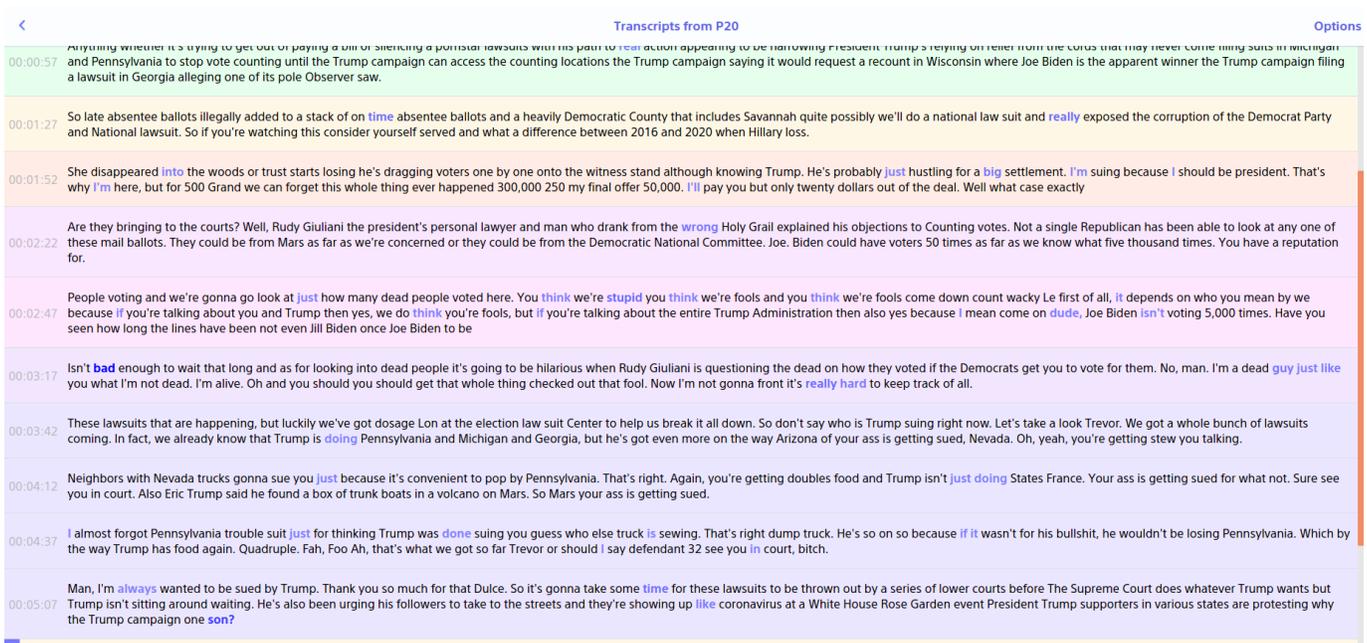

Figure 7. Color-coded transcript view, based on audio direction of arrival. This view also included highlighted keywords.

The transcript view is colour-coded by audio direction of arrival and also highlights keywords in the transcripts. The transcripts can also be labelled with the detected speaker using speaker diarization. The inclusion of anonymized speaker labels or colour-coded direction of arrival helps to provide pertinent, individual-specific information even if participants may not be using a microphone array with direction-of-arrival capabilities, and are instead using a smartphone.

Given this general overview of the BLINC platform, we now discuss connections to the different dimensions that we covered in previous sections.

4.3.2. Communication and Contribution

As previously noted, certain design components can feasibly align with multiple dimensions. When we look broadly at the different representations, particularly the discussion timeline and the direction-of-arrival chart, we see how these may be used to support questions about Communication and Contribution. Both allow the user to see a quantification of how much communication is taking place within the group, both collectively, in the timeline, and on a more fine-grained bases in the direction-of-arrival chart. These components also include useful information about the types of verbal contributions being made and, broadly speaking, how equally distributed those contributions are across the group.

The transcript view similarly lets users see what kind of communication is taking place, and the origins of those different contributions. By following along the transcripts, users can also see how much the different verbal contributions being made are being taken up within subsequent conversations.

4.3.3. Constructive

The different representations also speak to the Constructive category in that they provide information that can be used to review a group's stated goals and expectations, from the timeline, and ascertain how well the group is moving towards those goals. The keyword representations can be used as a means for seeing how well students are utilizing domain-specific language, for example, or, at the very least, discussing content relevant to the activity at hand.

4.3.4. Context

The salience of context emerges across several different platform features. First, the home screen, which allows users to specify the keywords and topic models utilized within the backend, allows for the analytics being used to be easily adapted to the various disciplinary contexts in which students may practise collaboration. The platform also utilizes contextualized data representations by allowing users to specify the interval of data they want to be displayed, and by supporting the ability to drill down into the specific utterances where a given keyword or question was used. This closely aligns with Echevarria's suggestion to contextualize information. Additionally, the platform is intentional about not presuming what constitutes a good or bad collaboration, but instead provides some numeric and graphical data that users can make sense of based on the context. Perhaps a more important feature is the capability for participants to capture collaboration data on their own terms. The opportunity to collect and receive real-time analysis in any internet-enabled space using a smartphone means that students can examine their collaboration data outside of classroom experiences. As we saw from the Likert scale questions and student quotations, students thought this would be extremely helpful.

4.3.5. Climate

With regard to Climate, the text-based sentiment analysis features mean that both students and teachers can have a window into the emotional tone of group interactions. This information could be useful in real-time to redirect group conversation to address concerning behaviour. They may also use this information post-hoc to better realize how changes in the climate propagate across different participants in a group, and for students to reflect upon how these changes in climate impacted their own participation within the group. As described by the students in the survey, this is often an important dimension that is not always captured by only looking at the content of a given group's discussion or the quality of the work product created.

4.3.6. Compatibility and Conflict Resolution

The current BLINC implementation supports the dimensions of compatibility and conflict resolution by providing a combination of transcript data and sentiment analysis data that teachers and students can use to reason about group dynamics, and to see how well students are participating in conflict resolution. For example, teachers might utilize the transcript information to see how well different students' ideas are being taken up by their peers. The direction-of-arrival chart also gives a high-level view of how well a group may be working together and/or the different roles that students are taking up within their groups. In terms of conflict resolution, the transcripts of discussion, while still anonymous, can help external facilitators intervene appropriately and effectively when conflicts arise. It also provides an external individual with important contextual information in the prior conversation that took place; in this way, it can be a useful tool for engaging with the group.

While there was not space here to enumerate each feature of the BLINC platform and how it aligns to the goals of collaboration literacy, the information above should have highlighted three key ideas: 1) there are many ways to instantiate the different dimensions within a collaboration analytics platform; 2) a single interface design component can feasibly address

multiple dimensions; and 3) the design components, can be used to support collaboration literacy, albeit this may require intentional connections to instructional strategies and/or teacher interventions.

5. Discussion

The seven categories that we identified based on student survey responses help us consider the ways that students think about collaboration. These seven categories include apparent ties back to prior research, but also feature a number of novel considerations. The dimensions of Communication, Contribution, and Constructive are readily apparent across a broad spectrum of research projects in collaboration analytics, MMLA, and educational dashboards. Concretely, these three dimensions speak to primary questions about what people are talking about within a given collaborative group, who is contributing, and how well the group is achieving their goals. On the other hand, some of the most intriguing considerations come among the dimensions of Context, Climate, and Compatibility, which move outside of the content of group interactions and, instead, focus more on questions of where collaboration takes place, the affective dimensions of group activities, and questions of overall group dynamics. While less prevalent within responses to our survey questions, these dimensions also appear to be relevant for supporting student collaboration both in terms of direct information for learners, and in terms of information that teachers might use to support student collaboration and collaboration literacy. Moreover, within our analysis of category co-occurrence, we see that while Communication and Contribution frequently co-occurred with most of the other dimensions, the Context code was a noticeable outlier. Namely, in nearly 30% of the instances that the Context code was utilized, it was the only code applied to that response. This seems to suggest that thinking about the Context dimension is an especially important consideration.

In the second portion of the analysis, we looked at ways that these dimensions might be instantiated in a collaboration analytics platform. We situated these examples within the design of the BLINC tool. In looking through those examples, we demonstrated ways that a given design or interface component can satisfy multiple categories. We also demonstrate that there are multiple ways to address the different dimensions, and that the objective is not necessarily to explicitly name each dimension within the interface. The goal with considering the different dimensions is to ensure that the data being provided can align with the ways that students may want to think about or reflect upon their collaboration experiences. Additionally, as we described with our use of user smartphones as data collection devices, are the ways that intentionality within the design process can enable users to practise and reflect upon collaboration analytics across a broader set of contexts. The discussion of implementing the different components as supports for collaboration literacy also reiterated that these types of tools may reasonably need to be utilized alongside intentional instruction or interventions by an educator.

6. Conclusion

The current analysis has been an opportunity to closely consider the ways that students think about collaboration and collaboration literacy. We position collaboration literacy as an important re-orientation for the collaboration analytics community. Considering collaboration as a literacy can push us to think about how to help people learn to read and respond to different collaborative settings. Developing such a skill set requires that we support more thorough reflection of collaborative experiences. Importantly, this reflection should carefully align with the seven categories, or dimensions, that we identified in this paper: Climate, Compatibility, Communication, Conflict, Context, Contribution, and Constructive. These categories are not meant to suggest the specific language to utilize within a collaboration analytics interface. Instead, researchers and designers should reflect on the categories and ensure that the different dimensions are reflected in the design, or use of the tool. Additionally, we observed that in many ways the categories are interrelated and appear to span a diverse set of ideas that students have about collaboration. From this perspective, researchers and designers may find that a single cleverly designed interface component can address multiple dimensions.

We also note that the categories are not prescriptive in identifying what good or bad collaboration looks like. Instead, they touch on dimensions over which someone may think about how they, or their group, is collaborating as desired. While we recognize that these dimensions are based on a relatively targeted sample of students, we are encouraged by the ways that the emergent categories build on existing research in the collaboration analytics and collaborative problem-solving communities. The confluence of novel data collection tools and global interest in collaboration present several new and exciting opportunities for this research community. We suggest that through the seven dimensions outlined in this paper, and tools like BLINC that aim to support collaboration literacy, we can grow important pedagogical innovations in collaboration across a variety of contexts.

Declaration of Conflicting Interest

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

Funding

References

Alhadad, S. S. (2018). Visualizing data to support judgement, inference, and decision making in learning analytics: Insights from cognitive psychology and visualization science. *Journal of Learning Analytics*, *5*(2), 60–85. https://doi.org/10.18608/jla.2018.52.5

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, *12(3)*, 307–359. https://doi.org/10.1207/S15327809JLS1203_1

Bassiou, N., Tsiartas, A., Smith, J., Bratt, H., Richey, C., Shriberg, E., & Alozie, N. (2016). Privacy-preserving speech analytics for automatic assessment of student collaboration. *Proceedings of the 17th Annual Conference of the International Speech Communication Association* (INTERSPEECH 2016), 8–12 September 2016, San Francisco, CA, USA (pp. 888–892). International Speech Communication Association. https://doi.org/10.21437/Interspeech.2016-1569

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. https://doi.org/10.1007/s11412-015-9217-z

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 17–66). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-2324-5_2

Blanchard, N., D'Mello, S., Olney, A. M., & Nystrand, M. (2015). Automatic classification of question & answer discourse segments from teacher's speech in classrooms. In O. C. Santos et al. (Eds.), *Proceedings of the 8th International Conference on Educational Data Mining* (EDM2015), 26–29 June 2015, Madrid, Spain (pp. 282–288). International Educational Data Mining Society. Retrieved from https://files.eric.ed.gov/fulltext/ED560555.pdf

Boaler, J., Wiliam, D., & Brown, M. (2000). Students' experiences of ability grouping-disaffection, polarisation and the construction of failure. *British Educational Research Journal*, *26*(5), 631–648.

Borge, M., & Rosé, C. P. (2016). Automated feedback on group processes: An experience report. In T. Barnes et al. (Eds.), *Proceedings of the 9th International Conference on Educational Data Mining* (EDM2016), 29 June–2 July 2016, Raleigh, NC, USA (pp. 573–574). International Educational Data Mining Society.

Bos, M. C. (1937). Experimental study of productive collaboration. *Acta Psychologica, 3*, 315–426. https://doi.org/10.1016/S0001-6918(01)90007-1

Carbone, E. (1999). Students behaving badly in large classes. *New Directions for Teaching and Learning*, *1999*(77), 35–43.

Cukurova, M., Luckin, R., Mavrikis, M., & Millan, E. (2017). Machine and human observable differences in groups collaborative problem-solving behaviours. *Proceedings of the 12th European Conference on Technology Enhanced Learning* (EC-TEL 2017), 12–15 September 2017, Tallinn, Estonia (pp. 17–29). Lecture Notes in Computer Science, Springer.

Cukurova, M., Luckin, R., Millan, E., & Mavrikis, M. (2018). The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. *Computers & Education, 116*, 93–109. https://doi.org/10.1016/j.compedu.2017.08.007

Cukurova, M., Zhou, Q., Spikol, D., & Landolfi, L. (2020). Modelling collaborative problem-solving competence with transparent learning analytics: Is video data enough? *Proceedings of the 10th International Conference on Learning Analytics and Knowledge* (LAK '20), 23–27 March 2020, Frankfurt, Germany (pp. 270–275). New York: ACM. https://doi.org/10.1145/3375462.3375484

D'Angelo, C., Smith, J., Alozie, N., Tsiartas, A., Richey, C., & Bratt, H. (2019). Mapping individual to group level collaboration indicators using speech data. In K. Lund, G. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings. Proceedings of the 13th International Conference on Computer Supported Collaborative Learning* (CSCL 2019), 17–21 June 2019, Lyon, France (Vol. 2, pp. 628–631). International Society of the Learning Sciences.

Dillenbourg, P., & Bachour, K. (2010). An interactive table for supporting participation balance in face-to-face collaborative learning. *IEEE Transactions on Learning Technologies, 3*(3), 203–213. https://doi.org/10.1109/TLT.2010.18

Dillenbourg, P., Nussbaum, M., Dimitriadis, Y., & Roschelle, J. (2013). Design for classroom orchestration. *Computers & Education*, *69*, 485–492.

Dillenbourg, P., Zufferey, G., Alavi, H., Jermann, P., Do-lenh, S., & Bonnard, Q. (2011). Classroom orchestration: The third circle of usability. *Proceedings of the 9th International Conference on Computer-Supported Collaborative Learning* (CSCL 2011), 4–8 July 2011, Hong Kong, China (Vol. 1, pp. 510–517). International Society of the Learning Sciences.

Domínguez, F., Chiluiza, K., Echeverria, V., & Ochoa, X. (2015). Multimodal selfies: Designing a multimodal recording device for students in traditional classrooms. *Proceedings of the 17th ACM International Conference on Multimodal Interaction* (ICMI '15), 9–13 November 2015, Seattle, WA, USA (pp. 567–574). New York: ACM.

Echeverria, V., Martinez-Maldonado, R., Shum, S. B., Chiluiza, K., Granda, R., & Conati, C. (2018). Exploratory versus explanatory visual learning analytics: Driving teachers' attention through educational data storytelling. *Journal of Learning Analytics, 5*(3), 72–97. https://doi.org/10.18608/jla.2018.53.6

Echeverria, V., Martinez-Maldonado, R., Shum, S. B., & Sydney, T. (2019). Towards collaboration translucence: Giving meaning to multimodal group data. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI '19), 4–9 May 2019, Glasgow, Scotland, UK (Paper No. 39, pp. 1–16). New York: ACM.

Edmondson, A. C., & Lei, Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 23–43.

Furuichi, K., & Worsley, M. (2018). Using physiological responses to capture unique idea creation in team collaborations. *Companion of the 2018 ACM Conference on Computer-Supported Cooperative Work & Social Computing* (CSCW '18) 3–7 November 2018, Jersey City, NJ, USA (pp. 369–372). New York: ACM.

Gero, J., & Kannengiesser, U. (2004, 07). The situated function-behaviour-structure framework. *Design Studies, 25*(4), 373–391. https://doi.org/10.1016/j.destud.2003.10.010

Gibbs, G., Jenkins, A., & Alan, J. (Eds.). (1992). *Teaching large classes in higher education: How to maintain quality with reduced resources*. Hove, East Sussex, UK: Psychology Press.

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory. *International Journal of Qualitative Methods, 5*, 1–10. Retrieved from http://www.sxf.uevora.pt/wp-content/uploads/2013/03/Glaser_1967.pdf

Grover, S., Bienkowski, M., Tamrakar, A., Siddiquie, B., Salter, D., & Divakaran, A. (2016). Multimodal analytics to study collaborative problem solving in pair programming. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge* (LAK '16), 25–29 April 2016, Edinburgh, UK (pp. 516–517). New York: ACM. https://doi.org/10.1145/2883851.2883877

Hmelo-Silver, C. E., & Barrows, H. S. (2008). Facilitating collaborative knowledge building. *Cognition and Instruction, 26*(1), 48–94. https://doi.org/10.1080/07370000701798495

Jung, Y. J., Toprani, D., Yan, S., & Borge, M. (2017). Children's participation in rulemaking to mitigate process problems in CSCL. In B. K. Smith, M. Borge, E. Mercier, & K. Y. Lim (Eds.), *Making a Difference: Prioritizing Equity and Access in CSCL. Proceedings of the 12th International Conference on Computer Supported Collaborative Learning* (CSCL 2017) 18–22 June 2017, Philadelphia, PA, USA (Vol. 2, pp. 652–655). International Society of the Learning Sciences.

Marsh, A. A., Elfenbein, H. A., & Ambady, N. (2003). Nonverbal "accents": Cultural differences in facial expressions of emotion. *Psychological Science*, *14*(4), 373–376.

Martinez-Maldonado, R., Kay, J., Buckingham Shum, S., & Yacef, K. (2019). Collocated collaboration analytics: Principles and dilemmas for mining multimodal interaction data. *Human–Computer Interaction*, *34*(1), 1–50.

Martinez-Maldonado, R., Power, T., Hayes, C., Abdiprano, A., Vo, T., Axisa, C., & Buckingham Shum, S. (2017). Analytics meet patient manikins: Challenges in an authentic small-group healthcare simulation classroom. *Proceedings of the 7th International Conference on Learning Analytics and Knowledge* (LAK '17), 13–17 March 2017, Vancouver, BC, Canada (pp. 90–94). New York: ACM. http://doi.acm.org/10.1145/3027385

Mayo, E. (1939). Routine interaction and the problem of collaboration. *American Sociological Review*, *4*(3), 335–340. https://doi.org/10.2307/2084920

Ochoa, X., Chiluiza, K., Mendez, G., Luzardo, G., Guamán, B., & Castells, J. (2013). Expertise estimation based on simple multimodal features. *Proceedings of the 15th ACM International Conference on Multimodal Interaction* (ICMI '13), 9–13 December 2013, Sydney, Australia (pp. 583–590). New York: ACM. https://doi.org/10.1145/2522848.2533789

Oviatt, S., & Cohen, A. (2014). Written activity, representations and fluency as predictors of domain expertise in mathematics. *Proceedings of the 16th ACM International Conference on Multimodal Interaction* (ICMI '14), 12–16 November 2014, Istanbul, Turkey (pp. 10–17). New York: ACM. https://doi.org/10.1145/2663204.2663245

Oviatt, S., Cohen, A., & Weibel, N. (2013). Multimodal learning analytics: Description of math data corpus for ICMI grand challenge workshop. *Proceedings of the 15th ACM International Conference on Multimodal Interaction* (ICMI '13), 9–13 December 2013, Sydney, Australia (pp. 563–568). New York: ACM. https://doi.org/10.1145/2522848.2533790

Praharaj, S., Scheffel, M., Drachsler, H., & Specht, M. (2018). Multimodal analytics for real-time feedback in co-located collaboration. *Proceedings of the 13ᵗʰ European Conference on Technology Enhanced Learning* (EC-TEL 2018)*,* 3–5 September 2018, Leeds, UK (pp. 187–201). Lecture Notes in Computer Science, Springer.

Reimers, G., & Neovesky, A. (2015). Student focused dashboards. In S. Zvacek, M. T. Restivo, J. Uhomoibhi, M. Helfert (Eds.), *Proceedings of the 7ᵗʰ International Conference on Computer Supported Education* (CSEDU 2015), 23–25 May 2015, Lisbon, Portugal (Vol. 1, pp. 399–404). Springer.

Roschelle, J. (1999). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences, 2*(3), 235–276.

Schneider, B., & Blikstein, P. (2015). Unraveling students' interaction around a tangible interface using multimodal learning analytics. *Journal of Educational Data Mining, 7*(3), 89–116.

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. https://doi.org/10.18608/jla.2015.22.9

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2018). Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning, 13*(3), 241–261.

Spikol, D., Cukurova, M., & Ruffaldi, E. (2017). Using multimodal learning analytics to identify aspects of collaboration in project-based learning. In B. K. Smith, M. Borge, E. Mercier, & K. Y. Lim (Eds.), *Making a Difference: Prioritizing Equity and Access in CSCL. Proceedings of the 12ᵗʰ International Conference on Computer Supported Collaborative Learning* (CSCL 2017) 18–22 June 2017, Philadelphia, PA, USA (Vol. 1, pp. 263–270). International Society of the Learning Sciences. https://doi.org/10.22318/cscl2017.37

Spikol, D., Avramides, K., & Cukurova, M. (2016). Exploring the interplay between human and machine annotated multimodal learning analytics in hands-on STEM activities. *Proceedings of the 6ᵗʰ International Conference on Learning Analytics and Knowledge* (LAK '16), 25–29 April 2016, Edinburgh, UK (pp. 522–523). New York: ACM. https://doi.org/10.1145/2883851.2883920

Spikol, D., Superiore, S., Anna, S., & Landolfi, L. (2020). Modelling collaborative problem-solving competence with transparent learning analytics: Is video data enough? *Proceedings of the 10ᵗʰ International Conference on Learning Analytics and Knowledge* (LAK '20), 23–27 March 2020, Frankfurt, Germany (pp. 270–275). New York: ACM. https://doi.org/10.1145/3375462.3375484

Stone, P. J., Bales, R. F., Namenwirth, J. Z., & Ogilvie, D. M. (1962). The general inquirer: A computer system for content analysis and retrieval based on the sentence as a unit of information. *Systems Research and Behavioral Science, 7*(4), 484–498.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1), 24–54.

Teasley, S. D. (2017). Student facing dashboards: One size fits all? *Technology, Knowledge and Learning*, *22*(3), 377–384.

Terken, J., & Sturm, J. (2010). Multimodal support for social dynamics in co-located meetings. *Personal and Ubiquitous Computing, 14*(8), 703–714. https://doi.org/10.1007/s00779-010-0284 -x

Tse, E., Greenberg, S., Shen, C., & Forlines, C. (2007). Multimodal multiplayer tabletop gaming. *Computers in Entertainment, 5*(2), 12. https://doi.org/10.1145/1279540.1279552

Van Leeuwen, A. V., Janssen, J., Erkens, G., & Brekelmans, M. (2015). Teacher regulation of cognitive activities during student collaboration: Effects of learning analytics. *Computers & Education, 90,* 80–94. http://doi.org/10.1016/j.compedu.2015.09.006

Vrzakova, H., Amon, M. J., Stewart, A., Duran, N. D., & Mello, S. K. D. (2020). Focused or stuck together: Multimodal patterns reveal triads' performance in collaborative problem solving. *Proceedings of the 10ᵗʰ International Conference on Learning Analytics and Knowledge* (LAK '20), 23–27 March 2020, Frankfurt, Germany (pp. 295–304). New York: ACM. https://doi.org/10.1145/3375462.3375467

Williams, R. (2015). The non-designer's design book: Design and typographic principles for the visual novice. London: Pearson Education.

Wise, A. F., & Chiu, M. M. (2011). Analyzing temporal patterns of knowledge construction in a role-based online discussion. *International Journal of Computer-Supported Collaborative Learning, 6*(3), 445–470. https://doi.org/10.1007/s11412-011-9120-1

Worsley, M., & Blikstein, P. (2015). Leveraging multimodal learning analytics to differentiate student learning strategies. *Proceedings of the 5ᵗʰ International Conference on Learning Analytics and Knowledge* (LAK '15), 16–20 March 2015, Poughkeepsie, NY, USA (pp. 360–367). New York: ACM. https://doi.org/10.1145/2723576.2723624

Worsley, M., & Blikstein, P. (2018). A multimodal analysis of making. *International Journal of Artificial Intelligence in Education, 28*(3), 385–419.

Worsley, M., & Blikstein, P. (2011). What's an expert? Using learning analytics to identify emergent markers of expertise through automated speech, sentiment and sketch analysis. In M. Pechenizkiy et al. (Eds.), *Proceedings of the 4th Annual Conference on Educational Data Mining* (EDM2011), 6–8 July 2011, Eindhoven, Netherlands (pp. 235–240). International Educational Data Mining Society.

Worsley, M., & Ochoa, X. (2020). Towards collaboration literacy development through multimodal learning analytics. In M. Giannakos, D. Spikol, I. Molenaar, D. Di Mitri, K. Sharma, X. Ochoa, & R. Hammad (Eds.), *Proceedings of CrossMMLA in Practice: Collecting, Annotating and Analyzing Multimodal Data Across Spaces*. Co-located with the 10th International Conference on Learning Analytics and Knowledge (LAK '20), 24 March 2020, Cyberspace (Vol. 2610, pp. 53–63). New York: ACM. http://ceur-ws.org/Vol-2610/

Zhou, J., Hang, K., Oviatt, S., Yu, K., & Chen, F. (2014). Combining empirical and machine learning techniques to predict math expertise using pen signal features. *Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge* (MLA '14), 12–16 November 2014, Istanbul, Turkey (pp. 29–36). New York: ACM. https://doi.org/10.1145/2666633.2666638