

A New Era in Multimodal Learning Analytics: Twelve Core Commitments to Ground and Grow MMLA

Marcelo Worsley¹, Roberto Martinez-Maldonado², Cynthia D'Angelo³

Abstract

Multimodal learning analytics (MMLA) has increasingly been a topic of discussion within the learning analytics community. The Society of Learning Analytics Research is home to the CrossMMLA Special Interest Group and regularly hosts workshops on MMLA during the Learning Analytics Summer Institute (LASI). In this paper, we articulate a set of 12 commitments that we believe are critical for creating effective MMLA innovations. Moreover, as MMLA grows in use, it is important to articulate a set of core commitments that can help guide both MMLA researchers and the broader learning analytics community. The commitments that we describe are deeply rooted in the origins of MMLA and also reflect the ways that MMLA has evolved over the past 10 years. We organize the 12 commitments in terms of (i) data collection, (ii) analysis and inference, and (iii) feedback and data dissemination and argue why these commitments are important for conducting ethical, high-quality MMLA research. Furthermore, in using the language of commitments, we emphasize opportunities for MMLA research to align with established qualitative research methodologies and important concerns from critical studies.

Notes for Practice

- Multimodal learning analytics (MMLA) is a set of analytic techniques that can be used to better contextualize a given learning analytics project.
- The commitments described in this paper constitute best practices that researchers should follow in collecting, analyzing, and disseminating MMLA research that may be conducted in schools and laboratories.
- These commitments highlight the goal of making MMLA relevant to practice and ensuring that educators' and students' voices are authentically taken into consideration.
- Practitioners can use these commitments to hold researchers accountable for the ways that they enact MMLA innovations in educational settings.

Keywords

Ethics, data collection, data mining, artificial intelligence, data dissemination, multimodal, sensor data

Submitted: 18/09/20 — **Accepted:** 25/05/21 — **Published:** 16/11/21

Corresponding author ¹Email: 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>

²Email: roberto.martinezmalDONADO@monash.edu Address: Building H, Caulfield Campus, Monash University, 900 Dandenong Road, Caulfield East, VIC 3145, Australia. ORCID ID: <https://orcid.org/0000-0002-8375-1816>

³Email: cdangelo@illinois.edu Address: College of Education, Education Building, University of Illinois, Champaign, 1310 S. Sixth St., Champaign, IL, USA. ORCID ID: <https://orcid.org/0000-0001-6998-1052>

1. Introduction

It has been nearly 10 years since the introduction of multimodal learning analytics (MMLA) as an organized discipline (Scherer, Worsley, & Morency, 2012). What originated from an observed need for learning analytics to authentically reflect the various ways that learners may demonstrate their knowledge has grown into a dedicated and influential special interest group within the learning analytics community. As a sign of this growth, MMLA was the first special interest group within the Society of Learning Analytics Research (SoLAR) and has been influential in scores of research papers across several journals and conferences (Sharma & Giannakos, 2020; Worsley, 2018). However, as a relatively new community, research under the label of MMLA has, at times, seemed disparate, with projects that involve a variety of contexts, research populations, analytic techniques, and purposes. The relatively inclusive nature of the discipline has made some important contributions to the direction of this field. For instance, MMLA merged with the CrossLAK community (Martinez-Maldonado et al., 2016) to

realize the ways that the two approaches could support rich analyses of learning across various physical and digital learning spaces. However, as MMLA becomes increasingly prevalent within the learning analytics and education research communities, it is crucial to identify a set of core commitments that can serve to maintain the driving intent behind this approach.

Throughout this manuscript we are intentional about using the language of commitments rather than guiding principles, best practices, or tenets. This choice in language is motivated by a deliberate goal of acknowledging the important contributions of qualitative research methodologies and critical theory to the commitments that we describe. Additionally, we use this language to invite conversations between MMLA and other research methodologies. For example, interaction analysis (Jordan & Henderson, 1995), a core methodology within the learning sciences, follows a set of *commitments* that speak to the collection, analysis, and dissemination of data. Some of these include commitments to (1) analyze data “in use,” “in action,” or “in practice”; (2) capture data to allow close, repeated analysis and alternative interpretations; and (3) use analytic approaches that account for temporality of events and interactions among participants, materials, and places (Hall & Stevens, 2015). These commitments have become pillars that interaction analysis researchers use to reflect on the design, analysis, and dissemination of their work. One proposition is that such a set of commitments would be a meaningful addition to the MMLA community. We also note that the language of commitments is utilized within critical theory disciplines. For example, within the disability community, Hamraie and Fritsch (2019) define four *commitments* that lay the foundation for bridging technology and disability. As yet another example of commitments within academic research, Wise and Schwarz (2017) explain that the “CSCL [computer-supported collaborative learning] community is self-consciously founded on a commitment to the value of collaborative learning as an educational goal and focus of research, but also on a commitment to science as a means of shared inquiry” (p. 433). Our use of the term *commitments* follows a similar vein. Namely, commitments provide a clear demarcation of important stances for different disciplines.

The organization of this paper bears some similarity to Hall and Stevens’s (2015) paper, by laying out the commitments in terms of a common data analysis pipeline. We begin with commitments that relate to *data collection* (Section 3). This specifically considers the kind of data that are collected, the contexts from which they are collected, and the reasons for collecting them. We then move into commitments concerning *data analysis and inference* (Section 4). These commitments cover core ideas about how to work with multimodal data and ways to develop suitable inferences from those data. The final set of commitments relates to *feedback and dissemination* (Section 5). Once the analyses have been completed, it is important to consider how that information should be communicated to different stakeholders using various kinds of user interfaces. Following our presentation of the commitments, we reiterate and synthesize some common threads that pervade MMLA research (Section 6). We then discuss ways that the commitments reflect past, present, and future opportunities within MMLA and suggest how the commitments might be utilized by researchers and educators (Section 7).

2. Commitments Overview

This manuscript draws on prior work within the MMLA research community and closely related disciplines to articulate a set of core commitments that can ground future research in MMLA. Many of these core commitments have been mentioned in talks and tutorials given by the authors. They also reflect conversations over the years at international workshops on MMLA and CrossMMLA, where the authors served as organizers and participants. Finally, the commitments are implicit in many of the publications that have surfaced during these first 10 years of MMLA research. To this point, the authors have collectively reviewed hundreds of papers, published and unpublished, related to MMLA. This prior literature informs various dimensions and tensions that are salient to MMLA though not necessarily present across the entire field. That said, we do not position this paper as a traditional review of the field. Reviews tend to focus on the broad contributions of a given field and highlight existing gaps. We refer to prior work in the discussion of each commitment, but the focus of this paper is on the substance of the commitments, and not a quantification of MMLA contributions over the years.

Table 1 presents an overview of the commitments by category. We include a total of 12 commitments, which are roughly equally distributed across data collection, data analysis and inference, and feedback and dissemination. We position the commitments as reflective, prospective, and provocative. They are reflective in the sense that many of the commitments are drawn from early work in MMLA (e.g., Blikstein, 2013; Worsley, 2012). At the same time, many of the commitments also reflect the ways that the discipline has grown to a larger number of modalities (Vrzakova, Amon, Stewart, Duran, & D’Mello, 2020; Worsley & Blikstein, 2018), learning contexts (Di Mitri et al., 2020; Echeverria, Martinez-Maldonado, Power, Hayes, & Buckingham Shum, 2018), and analytic approaches (Cukurova, Luckin, Millán, & Mavrikis, 2018; Di Mitri, Schneider, Specht, & Drachler, 2018). More broadly, the community has seen several important practical, technological, and theoretical developments. Many of these developments are reflected in the articulation of the 12 commitments. The commitments are prospective from the perspective that they are based on an imagined future for MMLA where the research has an increased impact on scholarship research and student learning. Finally, the commitments are provocative in that we see them as an important space for conversation and a space for growth within this subdiscipline.

Table 1. MMLA commitments by category

Data Collection	Analysis and Inference	Feedback and Dissemination
Multimodality	Multimodal data and human inference	Transparency and benefit
Expansive learning experiences	Limitations in prediction from multimodal data	Multimodal feedback
Making learner’s complexity visible	Participatory interpretation of multimodal data	Meaningful, usable feedback
Learning across spaces	Representation and multimodal data analysis	
Multimodal data control		

3. Data Collection Commitments

Considerable intentionality and planning are needed to effectively undertake an MMLA study that can tractably answer a given set of research questions. The increasing availability of low-cost multimodal sensors makes it easy to lose sight of the guiding intent of the proposed study. The commitments presented in this section help to lay the foundation for collecting data in ways that take advantage of the affordances of MMLA and maintain standards for high-quality, ethical research.

3.1. Commitment 1: Multimodality

Students demonstrate and communicate knowledge, interests, and intent through a plurality of modalities.

“A key tenet of MMLA is the recognition that teaching and learning are enacted through multiple modalities. Even in a traditional classroom, teachers engage in significant multimodal behaviors (voice inflections, gestures, etc.) in order to emphasize and de-emphasize different ideas in a lecture. Similarly, students draw upon a host of modalities in order to demonstrate their knowledge, and, more importantly, to gain in their understanding of a given subject area. In this way, MMLA draws upon certain ideas from Constructionism (Papert, 1980), namely the importance of conceptualizing and constructing using a broad set of modalities, and Embodied Cognition (Kirsh, 2011), the ability for embodied experiences to spur cognition” (Worsley et al., 2016, p. 1346).

As described by Worsley and colleagues (2016), multimodality is central to the idea of MMLA. This, we presume, should be obvious from the name. However, what may be less apparent is that this approach is used to recognize the plurality of approaches that learners may use to demonstrate or experience learning. Nigay and Coutaz (1993) defined multimodality as both (i) the various types of communication channels used to express ideas or convey information, including the ways in which an action can be performed, and (ii) the various ways in which information can be interpreted to generate new meaning. In other words, multimodality in the context of human-computer interaction (HCI) can involve both modelling the complexity of human communication (Obrenovic & Starcevic, 2004) and supporting the automated generation of feedback that people can receive in various ways (e.g., through haptic, audio, and visual mechanisms) (Freeman et al., 2017). It is thus fundamentally about expanding the set of modalities that we might use to understand student learning, without narrowly assuming that one modality (e.g., typing) should be a more acceptable form of knowledge demonstration than others (e.g., speaking, gesturing, making, or enacting).

Worsley and Blikstein (2011b) make this point in their discussion of “markers of expertise.” Within their specific study, these markers provide a way to more judiciously recognize the expertise that a novice is bringing in describing a solution to an engineering design problem. Concretely, Worsley and Blikstein highlight the ways that a novice and an expert created nearly identical solutions to the same engineering problem, albeit using starkly different representations of that idea. Conventionally, the expert’s representations, which included more articulate and specific statements as well as systematic tables, would be afforded greater recognition than the work of the novice, whose language was less precise and representations less systematic. While the novice used non-conventional ways to demonstrate her knowledge, her ideas should still be recognized as representing complex and sophisticated reasoning. Hence, multimodal data are used as a means to provide evidence of the multiple ways learners express themselves and the multiple ways they learn.

3.2. Commitment 2: Expansive Learning Experiences

MMLA centres learning experiences that may be collaborative, hands on, and face to face, de-emphasizing the computer screen as the primary form or object of interaction.

While multimodality plays an important role in MMLA, the current name overlooks a primary goal of MMLA of supporting the field’s ability to support teaching and learning within open-ended, complex learning environments. To this point, the incorporation of computational multimodal data analytics vastly pre-dates the naming of MMLA (Obrenovic & Starcevic,

2004; Oviatt, DeAngeli, & Kuhn, 1997). For instance, part of the emergence of MMLA was facilitated through the existing International Conference on Multimodal Interaction, which includes researchers exploring novel ways to use multimodal analytics and interfaces across a variety of contexts (see Scherer et al., 2012). Yet, a real point of distinction between MMLA and prior work in educational data mining or artificial intelligence in education is the orientation away from computer screen-mediated learning experiences. These open-ended, complex learning environments have been less studied by data-intensive initiatives, partly because they impose more challenges in terms of data collection and data interpretation (Baker, 2019). Put differently, MMLA advances experiences that ask researchers to think differently about both the form and the function of computers. Instead of the computer being the central point of learner focus, the computer operates as a tool that supports the learner as they interact with others and in the material world. To examine this further, consider the following quotation from Worsley & Blikstein (2011a):

“[W]e have worked to develop Learning Analytics—a set of multi-modal sensory inputs that can be used to predict, understand, and quantify student learning. Central to the efficacy of Learning Analytics is the belief that educators will be able to more easily adhere to learning recommendations when they are given the proper tools; in this case, tools for more accurately assessing student knowledge in open-ended learning tasks” (p. 1).

Within this first mention of MMLA, which served as the springboard for this research community, the use of multimodal input was seen as an obvious part of what learning analytics should constitute. Instead, the real emphasis was on providing “tools for more accurately assessing student knowledge in open-ended learning tasks.” Blikstein and Worsley (2016) made this point more explicit:

“To date most of the work on learning analytics and educational data mining has been focused on online courses and cognitive tutors, both of which provide a high degree of structure to the tasks, and are restricted to interactions that occur in front of a computer screen. In this paper, we argue that multimodal learning analytics can offer new insights into students’ learning trajectories in more complex and open-ended learning environments” (p. 222).

Again, the goal of MMLA is not to ignore or lament the power of computers. Much of the prior work on cognitive tutors and digital games has made apparent the ability of computing to help us grow in our understanding and design of digital learning. MMLA, however, aims to extend that computational power to learning experiences that are more expansive and support learners, teachers, and other stakeholders as they work to master many of the skills needed to face increasingly complex problems. Furthermore, this goal suggests that learning analytics research should creatively consider ways that the former factors of computers have substantially changed in recent history.

Computers are no longer devices that just sit on a desk with large screens that students look at. Instead, we should also realize the computing power of our mobile phones, microcomputers, smart watches, wearables, etc. Multimodal selfies (Domínguez, Echeverría, Chiluiza, & Ochoa, 2015) are a prime example of bringing the tools of computation, in the form of a microcomputer, to face-to-face learning experiences. These small devices are mobile and can be placed on a student’s desk to capture digital pen, audio, and video data of individuals or small groups. Hence, the computer, while still present, is not the focus of the learning experience. Several projects by MMLA researchers exemplify this point by conducting data-intensive research on collaborative (Cukurova et al., 2018; D’Angelo et al., 2019; Schneider & Pea, 2014), open-ended (Schneider & Blikstein, 2015; Worsley & Blikstein, 2018), and face-to-face (Donnelly et al., 2016; Echeverría, Martínez-Maldonado, & Buckingham Shum, 2019) learning contexts. Through these analyses, researchers have drawn new insights into increasingly complex learning environments. While not explicitly noted in the name MMLA, the support of open-ended, collaborative, face-to-face learning experiences that de-emphasize the centrality of computers as the primary point of interaction is a core commitment within the formulation of MMLA. That said, we also note that the current commitments can and should be applied across research from a variety of learning settings.

3.3. Commitment 3: Making Learners’ Complexity Visible

MMLA helps make the invisible visible. This is in terms of invisible modalities and, more broadly, in terms of patterns or practices that may exist within the data but may remain hard to see without computational aids.

Pervasive across many MMLA studies is the goal of surfacing aspects of learning that are hard to see. Existing work with eye-tracking and electro-dermal activation is a hallmark of this commitment, in that they incorporate modalities that offer a level of specificity not easily attained through human observation (Abrahamson, Shayan, Bakker, & Van Der Schaaf, 2016; Dindar, Järvelä, & Haataja, 2020; Huang, Bryant, & Schneider, 2019; Jermann, Gergle, Bednarik, & Brennan, 2012; Sharma,

Giannakos, & Dillenbourg, 2020). In the case of eye gaze, humans can broadly perceive where someone is looking, but typically not at the level or frequency provided by eye-tracking technology. Electro-dermal activation and other physiological sensors provide a window into the participant's experience that is wholly unavailable without these sensors. At times this requires specialized wearable sensors, while in other instances it can be achieved without individual instrumentation. For example, Eulerian magnification video data can be analyzed to provide estimates of a participant's heart rate (Wu et al., 2012). All of these represent access to relatively raw data points. In other instances, the analytic tools of MMLA support the identification, quantification, or qualification of higher-level patterns of engagement or behaviours that may only emerge when observing large quantities of data, or across multiple participants. Many examples commonly used within the collaborative problem-solving literature exemplify this point. For instance, constructs like turn-taking, turn management, body synchrony, and convergent conceptual change require tracking nuanced constructs across multiple participants and multiple modalities and at different time scales (Worsley & Ochoa, 2020).

MMLA aims to support researchers, practitioners, and learners as they attempt to make sense of these constructs, which can otherwise feel invisible (Cukurova, Kent, & Luckin, 2019; Di Mitri et al., 2018). This inclination toward making the invisible visible has important implications for expanding the researcher's ability to properly contextualize the participant's experience and to draw deeper inferences about a given learning experience. However, because of the amount of aid provided in drawing these inferences, and because they can otherwise seem invisible, justly interpreting these data is a significant challenge. Importantly, a commitment to exploring the invisible must also coincide with an increased commitment to ethically interpreting and using these data. Later we describe an associated commitment about including participants in the data interpretation process to reduce the propensity for misinterpretation of these, and other, signals.

Another key part of this commitment is recognizing the ways that computational analysis, or using artificial intelligence, on qualitatively annotated or computer-annotated data might surface hard-to-see patterns in a given data set. Prior research in MMLA includes examples where scholars have applied machine learning to data that were manually coded (Smith et al., 2016; Worsley & Blikstein, 2014, 2018) (see Di Mitri, Schneider, Klemke, Specht, and Drachler (2019) for example data annotation tools). Across different papers, the application of machine learning to manually coded data was intentionally undertaken because of the ways that it could aid the researcher in unearthing hard-to-see patterns. This is an important element of MMLA that is often overlooked and serves as a way to bridge valid data annotation and the affordances of computation. However, because this part of the commitment is more closely tied to data analysis, it will be further expanded upon later.

3.4. Commitment 4: Learning across Spaces

Learning is practised and evidenced across spaces, and context influences manifestations and conceptualizations of learning.

“Learning is a complex, mostly invisible process that happens across spaces, occurring in the physical world but also increasingly in virtual worlds or web-based spaces. In order to explore what happens in such blended learning experiences, there is a need for multiple data sources that bring evidence from these different spaces, including logs, learning resources, or even physical sensors. The combination of different data sources often generates multimodal datasets, with data representing different views of the same learning event” (Prieto et al., 2017, p. 1).

Learning ultimately happens where the student is, not in a particular digital or physical environment (Lave & Wenger, 1991; Looi, Wong, & Milrad, 2015). Pervasive and mobile technologies are enabling students to access educational resources from different physical spaces (e.g., ubiquitous/mobile learning support) and to enrich their learning experiences in the classroom in ways that were not previously possible (e.g., face-to-face/blended learning support). These technologies are increasingly embedded in everyday objects, enhancing their communication and data capture capabilities. As a result, new possibilities are emerging for creating richer models that can capture the complexity of learners' journeys in the increasingly hybrid learning spaces (Delgado Kloos, Hernández Leo, & Asensio-Pérez, 2012; Muñoz-Cristóbal et al., 2017; Pérez-Sanagustin, Santos, Hernández-Leo, & Blat, 2012). This was the initial aim of the initiative related to MMLA called Learning Analytics across Spaces (CrossLAK), organized at the International Conference on Learning Analytics and Knowledge in 2016 and 2017 (Martinez-Maldonado, Hernandez-Leo, Pardo, & Ogata, 2017; Martinez-Maldonado et al., 2016). This later merged with the MMLA community and was re-launched as the CrossMMLA SoLAR Special Interest Group on Multimodal Learning Analytics across Spaces.

This emphasizes the original intention of MMLA to focus on educational questions or problems to capture, model, and analyze learning in non-traditional spaces. It embraces the complexity of learning phenomena as human activity that is distributed across spaces, people, tools (both digital and physical), and time. Once the *learning problem and educational contexts* have been identified, an MMLA initiative can focus on assessing the feasibility of using learning analytics and modelling to tackle the research question at hand. These analytics may be simple (e.g., just focused on the analysis of one

modality of interaction that is collected from several spaces or that is relevant to achieve personalization across several spaces) or quite sophisticated (e.g., requiring the capture of traces such as eye gaze, posture, positioning, speech, and physiological markers).

The ultimate purpose is to take advantage of emerging sensing and data-processing technologies to gain a deeper understanding of learning, moving beyond the analysis of clickstreams and keystrokes in one context by also considering other sources of evidence such as speech, handwriting, sketching, gestures, postures, affective states, or eye gazing, which could be captured across educational contexts. By looking at educational experiences across spaces it can also be more natural to rely on educational theory or the learning designs to explain findings and disentangle the intertwined features that can be obtained by analyzing multiple streams of heterogeneous data.

It should, however, be noted that despite these goals of understanding student learning across contexts, it is important for researchers to also recognize that the research conducted in laboratories is materially different from that in ecological settings. While wanting to understand ecological settings, we have to be careful that we do not simply treat classrooms as extensions of the laboratory. As the ecological nature of research studies increases, so too, does the burden of collecting data in ways that are not overly invasive or overly exploratory. To this end, we are not advocating for mass deployment of multimodal sensors to every classroom. Instead, we should use the insights from MMLA to improve practice and develop data collection tools that provide the most benefit with the least amount of instrumentation. Put differently, even though we often affix an overwhelming number of sensors to each student, this is not desirable. As previously noted, the justification for these modalities must clearly align with the learning goals and research questions being explored. Moreover, as research moves into increasingly ecological settings, there is also an increased responsibility to add value to the research participants, a commitment that we describe in more detail later.

3.5. Commitment 5: Multimodal Data Control

Research participants control their data to preserve data privacy and ethics.

“While privacy guidelines for systems that expose student data exist, most focus on online systems in which the exploration of the data is often detached from the physical spaces in which it was collected. But capturing multimodal team data raises some acute concerns. For example, sensor data has a personal dimension to it not found in the more abstracted data from clickstreams, such as physiology, posture, gaze, and movement” (Martinez-Maldonado, Echeverria, Fernandez Nieto, & Buckingham Shum, 2020, p. 10).

The inclusion of multimodal data introduces heightened concerns about data privacy. Several codes of ethics have been proposed to mitigate the potential risks of misusing student data (see recent review by Kitto & Knight, 2019), which can serve as a basis to identify effective strategies for learners to control their data and ensure data privacy. Some specific design recommendations have also been suggested to address data ownership and access (Corrin et al., 2019; Drachsler & Greller, 2016; Ifenthaler & Schumacher, 2019). However, little work has explored how these issues intertwine with each other in the context of MMLA. The risks of pervasive surveillance in deploying sensors in classrooms or in asking teachers and students to use wearable devices are evident. In online learning systems, it may be easier for students to understand that all the actions they perform in the system can be monitored for the purpose of supporting their learning. For them, this translates into consenting to the recording of the clicks and keystrokes they perform in a particular digital environment or device. Yet capturing data using physical sensors and tracking devices can be a different story. For example, data from many inconspicuous sensors, such as accelerometers, gyroscopes, and barometers, can be used to make unexpected inferences, such as detecting daily life activities or personal habits, unrelated to learning tasks (Kröger, 2019). Respecting the privacy of participants is a crucial consideration.

The ubiquitous computing field has emphasized the critical importance of data protection and fair information practices to deal with sensor data (Camp & Connelly, 2008). This includes challenges considering how insights from sensor data are inferred and communicated; who gains access to the data, in which form and for what purpose; and how people can communicate their privacy and data-disclosure preferences. These are concerns that students have reported as critical in various learning analytics systems (Tsai, Whitelock-Wainwright, & Gašević, 2020). But sensor data in particular have a personal dimension not found in clickstream data. For example, posture, gaze, and gesture data relate to bodily interaction of students. Physiological and electroencephalography data, which have been increasingly captured in some MMLA projects (Worsley, 2018), can be strongly related to emotional and cognitive states unrelated to the learning task at hand. Additionally, many of these data would otherwise remain largely invisible to educators, or even students themselves. MMLA data can also have a strong social component in collaborative learning scenarios, particularly if MMLA interfaces are intended to be used for team reflection (e.g., Echeverria et al., 2019), revealing individuals’ information at least to other team members. To address these potential issues, Ochoa (2017) argues that “even if highly personal information is captured, privacy concerns are defused if the

decision of what and when to share it remains in the control of the learner.” Yet, the critical question is how can this be achieved in practice?

Two recent MMLA systematic reviews have revealed the urgent need to carefully consider the potential privacy issues of MMLA. The first review, conducted by Crescenzi-Lanna (2020) in the context of MMLA studies with children, contrasted the special ethical considerations for working with children (in terms of potential surveillance issues and parental involvement) with the potential value of multimodal data to enrich children’s learning. A second, more general, MMLA review (Sharma & Giannakos, 2020) pointed at the vacuum of ethics and privacy-related studies and warned against the use of several off-the-shelf sensors that entail privacy issues that can become further threats to adoption. To the best of our knowledge, only one MMLA-related paper has explicitly addressed a privacy issue related to students’ consent. Beardsley, Martinez Moreno, Vujovic, Santos, and Hernández-Leo (2020) recently presented a test to measure learners’ understanding of informed consent forms for MMLA research. It is becoming evident that protocols for ensuring transparency and communication will be key to increasing MMLA adoption.

The added privacy challenges that working with multimodal sensors may entail should not be used as arguments to diminish the potential that MMLA can bring to education. Instead, these challenges should drive discussion within the community of research and practice about the ethical issues that need to be addressed when working with sensor data and learning analytics. In sum, ethics and privacy integrity in MMLA research must comply with the same basic benchmarks as with any learning analytics project in terms of purpose, access, and anonymity (Tsai et al., 2020). But for pervasive data collection, we still need to create mechanisms to ensure data ownership and control over the sensors, devices, and systems collecting and processing the data. Horizontal practices to design for data-intensive innovation in intelligent physical spaces (e.g., see participatory surveillance by Albrechtslund and Ryberg (2011)) can certainly provide channels for including students’ and teachers’ voices in the design of such mechanisms. Yet, this is an area that has scarcely been explored (Oviatt, 2018).

4. Analysis and Inference Commitments

The purpose of data analysis is to transform the available multimodal data into useful insights that reflect an awareness of the learning context and knowledge of how to appropriately interpret the different modalities. Raw data, especially from many of the devices used in MMLA studies, seldom provide immediate answers to the proposed research questions. Instead, the raw data need to be filtered, reshaped, and summarized. Other times, data need to be collapsed into meaningful units (e.g., a class session or a problem) before being analyzed. This process involves several decisions on the part of the researcher. It also requires a good understanding of the local context in which the data are collected, who the participants are, and the analytic techniques available. Following data preparation, data can undergo any number of analyses and subsequently be used to draw inferences. The commitments in this section address different elements of this process by broadly considering data cleaning, data representation, and the role of participants in supporting data analysis and interpretation.

4.1. Commitment 6: Thorough, Consistent, and Transparent Data Modelling

MMLA research includes thorough, consistent, and transparent decision making with regard to data modelling options (e.g., data normalization, multimodal data fusion, and unit of analysis), because these choices greatly influence interpretations of data.

The data analysis process can include what might seem like an overwhelming number of decisions. Consistent with work within the broader learning analytics and education research communities, researchers must attend to questions about the appropriate unit of analysis (Häkkinen, 2013; Lehman, D’Mello, & Strain, 2011; Martinez-Maldonado, Kay, Buckingham Shum, & Yacef, 2019) and an appropriate time scale (Anderson, 2002; Richey, D’Angelo, Alozie, Bratt, & Shriberg, 2016). For example, work in the context of collaboration analytics states that “Varying the unit of analysis (e.g. individuals, groups, cohorts, devices, artifacts/objects) widens the possible insights that can be gained from the [...] analytics” (Martinez-Maldonado et al, 2019). Somewhat more specialized to the area of MMLA, however, are concerns with data normalization and multimodal fusion. For the case of data normalization, this is a step that is common across learning analytics research. However, the practice of normalizing multimodal data may require greater awareness of the underlying science about how the multimodal technology or modelling algorithms work. Within the MMLA community, this normalization process has frequently been applied with audio (Bassiou et al., 2016), electro-dermal activation (Dindar et al., 2020; Worsley & Blikstein, 2018), and facial expression (Grafsgaard et al., 2014; Worsley, Scherer, Morency, & Blikstein, 2015) analysis, as well as gesture classification (Schneider & Blikstein, 2015; Worsley & Blikstein, 2013). Given the extensive research on bias in facial expression and face recognition analysis (Xu, White, Kalkan, & Gunes, 2020), based on race, gender, and ethnicity, for example, there is an unmistakable need to effectively normalize the data and account for individual and group differences.

Hence, part of this commitment is about ensuring that the multimodal data modelling techniques being used are attentive to these biases and properly reflect the uniqueness of each modality.

Additionally, the modalities used are seldom analyzed individually. Instead, one of the driving opportunities with multimodal data is the ability to model different phenomena using a combination of modalities, as well as the opportunity to deconstruct a single modality into a divergent set of features (e.g., extracting pitch, voice intensity, emotional tone, timbre, and creak from the same audio sample). In early MMLA research, the common paradigm was to conduct feature extraction across several modalities and then fuse the output of those features within a machine learning classifier to predict the accuracy of a given model (Ochoa et al., 2013; Worsley & Blikstein, 2011b, 2011a). This type of analysis takes a fairly high level approach, in contrast to lower-level data fusion approaches where multimodal data are combined before passing the data through a classifier (see Lahat, Adali, & Jutten (2015) for a more detailed discussion of multimodal fusion approaches). These approaches operate at different ends of a spectrum (or a multidimensional space). Similarly, unique research questions of interest can be addressed by each approach. Regardless of the eventual modelling decisions, these choices should be clearly articulated and justified based on the research questions and context for a given project. Supporting this level of transparency is important from scientific and ethical perspectives and also serves to advance the education of future MMLA researchers.

4.2. Commitment 7: Multimodal Data and Human Inference

Computationally derived features often require human inference, knowledge of the algorithms used, and deep awareness of the context.

One of the allures of computational analysis is the ease with which it can generate labelled data. For example, within the previous commitment, we mentioned the ability to generate text from speech data, recognize gestures, and measure electro-dermal activation and facial expression. All of these processes are facilitated through advancements in artificial intelligence and signal processing. It is important to recognize, however, that these features should not be confused as absolutes, nor should they be equated to learning constructs. The features may be a proxy for a given learning construct but will often require human inference and knowledge of the context to be properly interpreted. Their interpretation may also require knowledge of the algorithm and an appreciation of the participant epistemology. Additionally, there is uncertainty within our measurement of all modalities. To provide a couple of concrete examples, consider using speech recognition as a proxy for student knowledge or expertise (e.g., Chandrasegaran, Bryan, Shidara, Chuang, & Ma, 2019). On one level, we may think that we hear a student say certain words, when, in fact, they said something different. This is to say that there may be uncertainty in what they articulated. With speech, this tends to be less of a concern when listening to our native languages, but it can often be the case when listening to languages that we are less familiar with. This is to say that human language processing is a complex endeavour that can be fallible for both humans and machines. On another level, a student saying a given collection of words does not mean that they understand that concept or idea. In the simplest sense, they could be repeating a comment or idea that someone else generated, or they could be posing a question. Hence, additional contextual information is essential for interpreting the features that may emerge from a given analysis. Moving to a somewhat more provocative example, a researcher may be interested in studying student emotions from electro-dermal activation or facial expression analysis software (e.g., Noroozi et al., 2019). The data from these two devices can help surface underlying changes in student physiology, but it would be inaccurate to say that a facial expression of confusion or frustration is identical to the student feeling frustrated or confused. The data might suggest that they are expressing confusion or frustration, but these features should be acknowledged as imprecise and interpreted measures.

4.3. Commitment 8: Limitations in Prediction from Multimodal Data

Learners are not defined by the behaviours that they exhibit within a given context. Any “predicted” label or dependent variable used within an algorithm or generated by data analysis reflects an interpretation of what a participant did, not who they are.

This commitment concerns how we use predicted labels or data about participants. Many of the algorithms used within MMLA can predict labels. These labels might be associated with a facial expression, how much someone talks, or pre- and post-test learning gains. While labelling is often a natural tendency, these labels reflect an interpretation of an analysis from a given context, or maybe even multiple contexts. Even so, researchers should be careful not to define a participant by an interpretation of their actions or performance in that context. The following example will highlight some important considerations about this.

It is not uncommon for studies to conduct a median split of the research participants based on their pre- and post-test learning gains. This is a reasonable approach for trying to differentiate between participants along the dimension of learning gains, for example. Furthermore, having established a performance difference between the two groups, researchers can begin

to interrogate how the patterns of engagement differed between the groups, with an understanding that some students performed better than others on the activity. This difference in performance, however, should not result in one group being labelled as low achievers and the other as high achievers. Assigning these static labels fails to represent the contextual and interpretative nature of the data and observations. With MMLA we prefer to use language along the lines of “students who scored lower” and “students who scored higher.” The difference between this suggested language and that of “low/high achiever” is subtle but important. A similar discussion easily translates across the different modalities that appear within MMLA research. A student who may frequently have been interpreted as having an angry facial expression should not be labelled as the “angry student” because, again, all of these are interpretations, and we, as researchers, are in no position to project identities onto research participants. We must also be careful about viewing learners in this light.

4.4. Commitment 9: Participatory Interpretation of Multimodal Data

MMLA research gives participants the opportunity to contribute to the researcher’s understanding of the data and can be an important way to provide transparency and improve the validity of claims that may emerge from an analysis.

MMLA incorporates several emerging low-cost sensors and computational techniques. In analyzing these different modalities, researchers may often face challenges in tractably and accurately interpreting the different data streams. Some MMLA works have addressed this challenge by reducing or oversimplifying the complexity of the learning phenomena by slicing the activity, that is, only looking at certain aspects of the task, such as speech (Mohan, Sun, Lederman, Full, & Pentland, 2018) or gaze (Schneider & Pea, 2013), to find correlations with some higher-order indicator, such as learning gain or collaboration. Others have taken a more holistic approach and let theory drive the analytics. This pathway has been encouraged by Worsley and Blikstein (2018), who propose epistemic framing to typify certain high-level activity (e.g., a person is discussing) based on a combination of low-level behaviours that can be detected via sensors (e.g., prolific gestures, an upright posture, gaze at peers, and animated talk and facial expressions). Both approaches can support insightful findings but would have potentially benefited from engaging research participants in the interpretation of the data.

Inclusion of research participants can take on various forms. One way to achieve this is to give educational stakeholders (i.e., teachers and students) the opportunity to share their expert knowledge and contribute to the researcher’s understanding of how to map from multimodal data to indicators of salient aspects of the learning activity. For example, Echeverria (2020) proposes a co-design approach, in the form of a template to be used by a researcher, to conduct interviews with teachers to identify the indicators that could be extracted from sensor data that could point at critical elements of their learning designs or the pedagogical approaches they follow. In her work she identified what evidence could be collected from indoor positioning trackers, system logs, and physiological wristbands to identify errors and stress levels of nursing students during face-to-face team simulations. Another approach involves presenting research data or findings to participants as a vehicle for reflections and annotation. For instance, researchers can use video playback of actions alongside participant facial expressions to help seed discussions about what they were feeling and thinking during different parts of a given learning experience.

In sum, because MMLA aims at embracing the complexity of learning, the analysis and inference of multimodal data can easily become a complex and challenging task. Identifying the complex interconnections between multiple sources of data is a key meaning-making challenge for the MMLA community. We suggest that enabling communication between researchers and participants can partly address the burden of interpretation of such relationships while providing transparency to the multimodal meaning-making process and validity to the kinds of claims that can emerge from the analysis.

5. Feedback and Data Dissemination Commitments

Equipped with a collection of analyses that elucidate their understanding of the multimodal practices of an individual or a group of participants, researchers typically aim to conclude by taking some form of action based on their findings. The commitments described here reflect the need to consider who should have access to the analytic findings, how the analyses should be represented, and how they might be utilized.

5.1. Commitment 10: Transparency and Benefit

MMLA research provides transparency and meaningful benefits to the participants as quickly as possible.

In the data collection section, we mentioned the need for MMLA to provide increasing benefits as the work moves more into ecological settings. Concretely, this commitment is about moving beyond the perspective that it is sufficient to simply avoid

doing harm. Instead, this commitment is about actively supporting teaching and learning by providing users access to the data, in useful representations, and using multimodal input to allow for more naturalistic interactions. In the case of the former, MMLA research should work to surface reliable data to participants in as close to real time as possible. Existing work on collaborative gaze awareness (Schneider & Pea, 2015) and group discussion analysis (Anderson et al., 2019) is a prime example of tools that provide real-time multimodal data to participants. Importantly, within each of these two examples, the platforms do not suggest what should be considered positive or negative multimodal behaviours. Instead, the tools leave that inference to the individuals or group to consider since there can frequently be contextual or situational factors that inform how to interpret this data. Hence, the expectation should not be that the platform provides researcher-level inferences, but it should make as much descriptive data available to participants as possible. Moreover, tools like Discussion Capture (Anderson et al., 2019) involve many intermediate features about group collaboration that can be informative to the participants. For example, real-time transcripts, while at times inaccurate, can be a resource that participants use to review or revisit elements of their prior conversation. Visual attention is another example of a modality that can easily be beneficial to groups collaborating with one another. If collaborators' gaze points are displayed, participants can more easily converge on key ideas or topics (Schneider & Pea, 2015). This is in contrast to simply using the eye gaze data to study learner behaviour from a research perspective. MMLA researchers should be intentional about considering which data streams can be beneficial to participants and put forth effort to design for participant needs alongside researcher needs. At the same time, the data must be presented in ways that acknowledge the uncertainty of the analyses. This is an essential part of maintaining transparency.

The examples of real-time transcription and eye tracking also elevate a second dimension for providing value to participants. Namely, participants can use these modalities as alternative input streams. Instead of engaging with an interface using text, users might use eye gaze or speech. This aligns with Commitment 1 on the multimodality of learning and also helps support the accessibility of different computer-supported learning environments.

5.2. Commitment 11: Multimodal Feedback

MMLA feedback leverages multimodality.

One of the early commitments that we mentioned highlighted the crucial role that multimodality brings to teaching and learning. Incorporating multiple modalities for demonstrating or experiencing knowledge was the overarching idea for that commitment. It is not surprising, then, that this same principle should carry over into how feedback or multimodal data should be shared with researchers and participants. Development of multimodal feedback has been prevalent within the HCI community (Ciordas-Hertel, 2020; Freeman et al., 2017; Limbu, Jarodzka, Klemke, & Specht, 2019) but has scarcely been explored within the MMLA community (Worsley & Ochoa, 2020). Instead, there has been a tendency to forget about multimodality as soon as the data have been collected and analyzed, resorting to traditional charts and figures in a dynamic dashboard in order to display data. This limits the ability of researchers, practitioners, and learners to effectively make sense of the data in meaningful ways. As an extreme example of what might be possible, consider a feedback, or data, representation of gestures that actually moves one's arm based on the user data. While seemingly far-fetched, current capabilities in HCI make this a possibility (Lopes & Baudisch, 2017). As a somewhat less extreme example, consider the opportunity to provide real-time feedback to students during a group collaboration session using vibrations on a phone or smartwatch (e.g., Ciordas-Hertel, 2020). Instead of highlighting student over-participation or distracting behaviour in a shared group display, the student might receive an individual notification in the form of a vibration or text-based notification. Prior work has considered some of these ideas within the HCI research community, but these approaches need to be more heavily examined within the MMLA community. Simply put, the incorporation of multimodal feedback not only reflects a recognition that multimodality is central to teaching and learning, but it also provides an additional set of dimensions that MMLA researchers can use to positively and creatively influence the learning environment.

5.3. Commitment 12: Meaningful, Usable Feedback

End-user MMLA interfaces deliver meaningful educational information that non-data savvy users can understand.

The community is due to deliver MMLA end-user interfaces. The latest MMLA systematic review (Sharma & Giannakos, 2020) confirms the dearth of interfaces that teachers or students can actually utilize, with a few exceptions. For example, Ochoa and colleagues (2018) proposed the Multimodal Transcript, which is a prototype that visualizes included logged actions, verbal participation, gaze direction, and emotional traits from groups of students working at an interactive tabletop in an experimental setting. Echeverria and colleagues (2019) designed four visualizations aimed at serving as proxies, each related to one modality (speech, arousal, positioning, and logged actions) of team activity. But these proxies were not fused into a single interface to facilitate reflection. Martinez-Maldonado and colleagues (2020) presented a multimodal layered approach to extract and

visualize data stories from multimodal data (positioning, physiological, and log data) to make it easier for teams of students to reflect on data captured while they collaborated in one face-to-face training session. A similar approach was followed by Ochoa and Dominguez (2020) by automatically capturing speech and pose data to automatically provide feedback to students giving oral presentations. Preliminary work by Vujovic and Hernández-Leo (2019) evaluated prototypes of MMLA interfaces to investigate how to compress electrodermal activity and noise data during meetings, but this was for interpretation by learning scientists. Anderson and colleagues (2019) describe another platform designed to directly provide collaboration analytics information to students and teachers. The platform includes many opportunities to drill down into specific moments in a conversation and opportunities for teachers and students to reflect on their collaboration contributions retrospectively. Another tool, the CPR Trainer (Di Mitri et al., 2020), consists of a multimodal data collection and feedback system that supports proper administration of CPR. The platform is an informative example of ways to provide users with relevant and actionable insights that can be extremely important to society.

A potential explanation for this dearth of MMLA end-user interfaces is that MMLA tools can very easily generate complex interfaces. Integrating data streams from multiple modalities can result in rather complex, hard-to-interpret interfaces (Ceneda et al., 2016). In MMLA that targets collaborative learning situations, this complexity is further multiplied by the number of students and educators involved in the activity and the interactions between them. It is thus imperative to deconstruct the data representations in meaningful ways that align with the needs of users. This means recognizing the particularities of the learning activity, the pedagogical intentions of the educator, and the needs that users without formal data analysis training have for interpreting data representations. Buckingham Shum, Ferguson, and Martinez-Maldonado (2019) suggest that instead of training teachers and students to use complex dashboards that require a high level of digital literacy, “it is more sensible to change the tools to suit their users, rather than changing the users to suit the tools” (p. 5). This of course requires a paradigm shift from imposing research prototypes and products entirely created by designers or researchers to embracing human-centred design approaches for teachers and students to become active agents in the design of the MMLA interfaces they use.

6. Discussion

Prior research in MMLA consists of considerable diversity in methodologies, research contexts, and analytic approaches. We envision that much of this diversity will continue throughout the upcoming decade but maintain that articulating some core commitments is an important part of growing the field. Thus far, this paper has focused on a seemingly broad set of ideas. In the subsections to follow, we more concretely outline a few of the potential challenges and opportunities that future MMLA research might address, specifically motivated by contemporary concerns with ethics and data privacy.

6.1. Better Interfaces and Support for Participants to Control Data Collection across Contexts

The elevation of privacy concerns has been a central area of discussion among MMLA researchers. As noted across a number of commitments, MMLA research should move toward giving participants more control over the multimodal data that are collected about them. This, however, will require the development of new tools and more sophisticated methodological approaches. First, on the topic of technological developments, many of the current technological tools for data collection are oriented toward instrumentation of entire learning environments and very seldom allow individuals to easily control data collection. Similar to the challenge of obtaining consent for entire classrooms, we need better tools for respecting these privacy concerns either in real time or post hoc. This technological challenge, however, will be much easier to address than the methodological challenge that this commitment presents. Considerable research and data mining innovation have been based on the assumption that the data being collected are voluminous and sampled without bias. If participants determine which data should be shared with researchers, these assumptions are no longer met. Because of this, many of the existing analytic techniques may need to be reconsidered or entirely abandoned. For example, it is difficult to compare two populations of users when you only have short, self-selected snippets of data from each group. To this end, adhering to a commitment about data privacy could mean re-examining the types of research questions that we study. For example, instead of looking at how a given learning intervention differentially impacted two groups of students, we might ask about the differences in the learning moments that students found to be worth sharing with peers and researchers.

As a complementary point about data collection across contexts, there is a need to develop tools that are more mobile and require less calibration and explicit normalization. Eye tracking and electro-dermal activation are two examples of technology that has become increasingly mobile in recent years but that can still be cumbersome to use effectively in authentic settings. Calibration, by having participants complete a few standardized tasks, is still a requirement in most cases. Additionally, few data collection devices support both rich multimodality and mobility.

6.2. In Expanding to Ecological Settings, Less Is More

In thinking about conducting research across contexts and at larger scales, MMLA researchers should be aware that less instrumentation and the avoidance of rigid predictions will provide greater mileage in terms of growing MMLA research. As we have noted in multiple places, research conducted in ecological settings is not the same as research conducted in laboratory

settings. MMLA began with a spirit of exploration in supporting expansive learning. However, the manner of scaling should not involve the same intent toward extensive exploration. We must avoid the dystopian futures that involve entire schools and classrooms persistently under the watchful eye of technology and multimodal sensors. Selectively employing these strategies to help inform learning theory could be within reason but must still be subject to considerable input from teachers and learners. Moreover, it requires a high level of intentionality and explicit discussions of questions of student and teacher data privacy.

As a complementary line of inquiry, MMLA research should look at a more extensive set of features that might be derived from commonly utilized data sources like audio and video. Using simpler modalities may decrease the cost and size of different systems. Furthermore, the more MMLA research can employ these more traditional data collection paradigms, the greater traction it will gather. For example, MMLA can certainly be applied to online, or remote, learning that is facilitated through multi-person video conference platforms. Although this speaks to part of the allure of clickstream data, it tends to be fairly innocuous to collect, but also provides reasonable signals for certain research questions. Hence, one approach with MMLA is to determine the level of support that multimodal data can provide given the practical constraints of schooling, minimizing the number of sensors being utilized (Lang, Woo, & Sinclair, 2020).

6.3. Interfaces That Privilege the Participant alongside the Researcher

Broadly speaking, researchers, practitioners, and learners are often interested in many of the same constructs and ideas. Researchers are typically able to devote more time to data analysis and are trained to analyze and interpret data that are quite complex and nuanced. Practitioners and learners tend to have less time to study data. Building on the alignment between researchers and practitioners, researchers should not overlook the opportunity for their work to also privilege the needs of learners and practitioners. This is with regard to developing interfaces and data representations that are pertinent and useful in practice. To this point, by thinking about different interfaces and data representations from the perspective of learners and practitioners, researchers will grow their ability to derive meaningful implications from their data. Furthermore, these representations might serve as important tools for onboarding new MMLA researchers into this community. At the same time, there is an opportunity for the multimodal interfaces to promote new forms of interactions. Doing so, however, means centring the needs of practitioners alongside those of researchers. Importantly, we suggest that taking this approach will grow the impact of MMLA and also result in a more robust and sustained research community.

7. Conclusion

We are encouraged by the broad set of researchers that are embracing the power of multimodality as an important lens for studying and supporting student learning. Ten years ago, this was uncharted space within the education research community. However, with the variety of research that has emerged in MMLA, the field has made significant and important advances. As we continue to advance this work, it is essential for the field to have a set of guiding commitments to ensure that MMLA is being used in ways that are ethical and meaningful. In putting forth these commitments, we do not envision that any existing MMLA research projects will be able to fulfill them all. Nor do we anticipate that future projects will be able to exemplify all of the commitments. Nonetheless, researchers should carefully consider these different commitments as they begin a given MMLA project, periodically reviewing them throughout the project timeline. Moreover, the research community should collectively work to develop solutions to the methodological, technical, and practical challenges presented by these commitments. The first 10 years of MMLA research has provided a strong complement to many of the existing practices and analytic techniques used in educational data mining and learning analytics. The commitments included in this paper are aimed to further push the MMLA community and subsequently translate into better research across learning analytics, educational data mining, and education research more broadly.

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

This material is based upon work supported by the US National Science Foundation under Grant No. 1832234. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Roberto Martinez-Maldonado's research is partially funded by the Jacobs Foundation.

References

- Abrahamson, D., Shayan, S., Bakker, A., & Van Der Schaaf, M. (2016). Eye-tracking Piaget: Capturing the emergence of attentional anchors in the coordination of proportional motor action. *Human Development*, 57(4-5), 218–244. <https://doi.org/10.1159/000443153>
- Albrechtslund, A., & Ryberg, T. (2011). Participatory surveillance in the intelligent building. *Design Issues*, 27(3), 35–46. https://doi.org/10.1162/DESI_a_00089
- Anderson, J. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science*, 26(1), 85–112. https://doi.org/10.1207/s15516709cog2601_3
- Anderson, K., Dubiel, T., Tanaka, K., Poultney, C., Brenneman, S., & Worsley, M. (2019). Chemistry pods: A multimodal real time and retrospective tool for the classroom. In *Proceedings of the 2019 International Conference on Multimodal Interaction (ICMI 2019)*, 14–18 October 2019, Suzhou, China (pp. 506–507). New York: ACM. <https://doi.org/10.1145/3340555.3358662>
- Baker, R. S. (2019). Challenges for the future of educational data mining: The Baker learning analytics prizes. *Journal of Educational Data Mining*, 11(1), 1–17. <https://doi.org/10.5281/zenodo.3554745>
- 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. In *Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH 2016)*, 8–12 September 2016, San Francisco, CA, USA (pp. 888–892). ISCA. <https://doi.org/10.21437/Interspeech.2016-1569>
- Beardsley, M., Martínez Moreno, J., Vujovic, M., Santos, P., & Hernández-Leo, D. (2020). Enhancing consent forms to support participant decision making in multimodal learning data research. *British Journal of Educational Technology*, 51(5), 1631–1652. <https://doi.org/10.1111/bjet.12983>
- Blikstein, P. (2013). Multimodal learning analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK 2013)*, 8–13 April 2013, Leuven, Belgium (pp. 102–106). New York: ACM. <https://doi.org/10.1145/2460296.2460316>
- Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of Learning Analytics*, 3(2), 220–238. <https://doi.org/10.18608/jla.2016.32.11>
- Buckingham Shum, S., Ferguson, R., & Martínez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, 6(2), 1–9. <https://doi.org/10.18608/jla.2019.62.1>
- Camp, J., & Connelly, K. (2008). Beyond consent: Privacy in ubiquitous computing (UbiComp). In A. Acquisti, S. de C. di Vimercati, S. Gritzalis, & C. Lambrinoudakis (Eds.), *Digital privacy: Theory, technologies, and practices* (pp. 1–17). Boca Raton, FL, USA: Auerbach Publications. <https://doi.org/10.1201/9781420052183-26>
- Ceneda, D., Gschwandtner, T., May, T., Miksch, S., Schulz, H.-J., Streit, M., & Tominski, C. (2016). Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1), 111–120. <https://doi.org/10.1109/TVCG.2016.2598468>
- Chandrasegaran, S., Bryan, C., Shidara, H., Chuang, T.-Y., & Ma, K.-L. (2019). TalkTraces: Real-time capture and visualization of verbal content in meetings. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI 2019)*, 4–9 May 2019, Glasgow, UK (pp. 1–14). New York: ACM. <https://doi.org/10.1145/3290605.3300807>
- Ciordas-Hertel, G.-P. (2020). How to complement learning analytics with smartwatches? Fusing physical activities, environmental context, and learning activities. In *Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI 2020)*, 25–29 October 2020, online (pp. 708–712). New York: ACM. <https://doi.org/10.1145/3382507.3421151>
- Corrin, L., Kennedy, G., French, S., Buckingham Shum, S., Kitto, K., Pardo, A., ... Colvin, C. (2019). *The ethics of learning analytics in Australian higher education*. Discussion Paper. Retrieved from http://utscic.edu.au.s3.amazonaws.com/wp-content/uploads/2021/03/22150303/LA_Ethics_Discussion_Paper.pdf
- Crescenzi-Lanna, L. (2020). Multimodal learning analytics research with young children: A systematic review. *British Journal of Educational Technology*, 51(5), 1485–1504. <https://doi.org/10.1111/bjet.12959>
- Cukurova, M., Kent, C., & Luckin, R. (2019). Artificial intelligence and multimodal data in the service of human decision-making: A case study in debate tutoring. *British Journal of Educational Technology*, 50(6), 3032–3046. <https://doi.org/10.1111/bjet.12829>
- Cukurova, M., Luckin, R., Millán, E., & Mavrikis, M. (2018). The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. *Computers & Education*, 116(Jan 2018), 93–109. <https://doi.org/10.1016/j.compedu.2017.08.007>

- D'Angelo, C. M., Smith, J., Alozie, N., Tsiartas, A., Richey, C., & Bratt, H. (2019). Mapping individual to group level collaboration indicators using speech data. In *Proceedings of the 13th International Conference on Computer Supported Collaborative Learning—A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings* (CSCL 2019), 17–21 June 2019, Lyon, France. ISLS. <https://repository.isls.org/handle/1/1637>
- Delgado Kloos, C., Hernández Leo, D., & Asensio-Pérez, J. I. (2012). Technology for learning across physical and virtual spaces. *Journal of Universal Computer Science*, 18(15): 2093–2096. http://www.jucs.org/jucs_18_15/technology_for_learning_across/abstract.html
- Di Mitri, D., Schneider, J., Klemke, R., Specht, M., & Drachsler, H. (2019). Read between the lines: An annotation tool for multimodal data for learning. In *Proceedings of the Ninth International Conference on Learning Analytics and Knowledge* (LAK 2019), 4–8 March 2019, Tempe, AZ, USA (pp. 51–60). New York: ACM. <https://doi.org/10.1145/3303772.3303776>
- Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning*, 34(4), 338–349. <https://doi.org/10.1111/jcal.12288>
- Di Mitri, D., Schneider, J., Trebing, K., Sopka, S., Specht, M., & Drachsler, H. (2020). Real-time multimodal feedback with the CPR Tutor. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial Intelligence in Education* (pp. 141–152). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-52237-7_12
- Dindar, M., Järvelä, S., & Haataja, E. (2020). What does physiological synchrony reveal about metacognitive experiences and group performance? *British Journal of Educational Technology*, 51(5), 1577–1596. <https://doi.org/10.1111/bjet.12981>
- Domínguez, F., Echeverría, V., Chiluíza, K., & Ochoa, X. (2015). Multimodal selfies: Designing a multimodal recording device for students in traditional classrooms. *Proceedings of the 2015 ACM International Conference on Multimodal Interaction* (ICMI 2015), 9–13 November 2015, Seattle, WA, USA (pp. 567–574). New York: ACM. <https://doi.org/10.1145/2818346.2830606>
- Donnelly, P. J., Blanchard, N., Samei, B., Olney, A. M., Sun, X., Ward, B., ... D'Mello, S. K. (2016). Automatic teacher modeling from live classroom audio. *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization* (UMAP 2016), 13–17 July 2016, Halifax, NS, Canada (pp. 45–53). New York: ACM. <https://doi.org/10.1145/2930238.2930250>
- Drachsler, H., & Greller, W. (2016). Privacy and analytics: It's a DELICATE issue. A checklist for trusted learning analytics. In *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge* (LAK 2016), 25–29 April 2019, Edinburgh, UK. New York: ACM. <https://doi.org/10.1145/2883851.2883893>
- Echeverría, V. (2020). *Designing and Validating Automated Feed-back for Collocated Teams Using Multimodal Learning Analytics*. PhD Thesis. University of Technology Sydney (UTS), Sydney, Australia.
- Echeverría, V., Martínez-Maldonado, R., & Buckingham Shum, S. (2019). Towards collaboration translucence: Giving meaning to multimodal group data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI 2019), 4–9 May 2016, Glasgow, UK (pp. 1–16). New York: ACM. <https://doi.org/10.1145/3290605.3300269>
- Echeverría, V., Martínez-Maldonado, R., Power, T., Hayes, C., & Buckingham Shum, S. (2018). Where is the nurse? Towards automatically visualising meaningful team movement in healthcare education. In *Proceedings of the International Conference on Artificial Intelligence in Education* (AIED 2018), 27–30 June 2018, London, UK (pp. 74–78). Cham: Springer. https://doi.org/10.1007/978-3-319-93846-2_14
- Freeman, E., Wilson, G., Vo, D.-B., Ng, A., Politis, I., & Brewster, S. (2017). Multimodal feedback in HCI: Haptics, non-speech audio, and their applications. In *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations—Volume 1* (pp. 277–317). New York: ACM. <https://doi.org/10.1145/3015783.3015792>
- Grafsgaard, J. F., Wiggins, J. B., Vail, A. K., Boyer, K. E., Wiebe, E. N., & Lester, J. C. (2014). The additive value of multimodal features for predicting engagement, frustration, and learning during tutoring. In *Proceedings of the Sixteenth ACM International Conference on Multimodal Interaction* (ICMI 2014), 12–16 November 2014, Istanbul, Turkey (pp. 42–49). New York: ACM. <https://doi.org/10.1145/2663204.2663264>
- Häkkinen, P. (2013). Multiphase method for analysing online discussions. *Journal of Computer Assisted Learning*, 29(6), 547–555. <https://doi.org/10.1111/jcal.12015>

- Hall, R., & Stevens, R. (2015). Interaction analysis approaches to knowledge in use. In A. A. diSessa, M. Levin, and N. J. S. Brown (Eds.), *Knowledge and interaction: A synthetic agenda for the learning sciences* (Chapter 3). Routledge. Retrieved from https://peabody.vanderbilt.edu/departments/tl/teaching_and_learning_research/space_learning_mobility/Hall_Stevens_2016.pdf
- Hamraie, A., & Fritsch, K. (2019). Crip technoscience manifesto. *Catalyst: Feminism, Theory, Technoscience* 5(1). <https://doi.org/10.28968/cftt.v5i1.29607>
- Huang, K., Bryant, T., & Schneider, B. (2019). *Identifying collaborative learning states using unsupervised machine learning on eye-tracking, physiological and motion sensor data*. Paper presented at the 12th International Conference on Educational Data Mining (EDM 2019), 2–5 July 2019, Montreal, Canada. <https://eric.ed.gov/?id=ED599214>
- Ifenthaler, D., & Schumacher, C. (2019). Releasing personal information within learning analytics systems. In D. Sampson, J. M. Spector, D. Ifenthaler, P. Isaías, and S. Sergis (Eds.), *Learning Technologies for Transforming Large-Scale Teaching, Learning, and Assessment* (pp. 3–18). Springer. https://doi.org/10.1007/978-3-030-15130-0_1
- Jermann, P., Gergle, D., Bednarik, R., & Brennan, S. (2012). Duet 2012: Dual eye tracking in CSCW. In *Proceedings of the 2012 ACM Conference on Computer Supported Cooperative Work. Companion (CSCW 2012)*, 11–15 February 2012, Seattle, WA, USA (pp. 23–24). ACM: New York. <https://doi.org/10.1145/2141512.2141525>
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1), 39–103. https://doi.org/10.1207/s15327809jls0401_2
- Kitto, K., & Knight, S. (2019). Practical ethics for building learning analytics. *British Journal of Educational Technology*, 50(6), 2855–2870. <https://doi.org/10.1111/bjet.12868>
- Kröger, J. (2019). Unexpected inferences from sensor data: A hidden privacy threat in the Internet of Things. In L. Strous & V. G. Cerf (Eds.), *Internet of things. Information processing in an increasingly connected world* (pp. 147–159). Springer International Publishing. https://doi.org/10.1007/978-3-030-15651-0_13
- Lahat, D., Adali, T., & Jutten, C. (2015). Multimodal data fusion: An overview of methods, challenges, and prospects. *Proceedings of the IEEE*, 103(9), 1449–1477. <https://doi.org/10.1109/JPROC.2015.2460697>
- Lang, C., Woo, C., & Sinclair, J. (2020). Quantifying data sensitivity: Precise demonstration of care when building student prediction models. In *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK 2020)*, 23–27 March 2020, Frankfurt, Germany (pp. 655–664). New York: ACM. <https://doi.org/10.1145/3375462.3375506>
- Lave, J., & Wenger, E. (1991). *Situated learning*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511815355>
- Lehman, B., D’Mello, S., & Strain, A. (2011). Inducing and tracking confusion with contradictions during critical thinking and scientific reasoning. In *Proceedings of the International Conference on Artificial Intelligence in Education (AIED 2011)*, 28 June–11 July 2011, Auckland, New Zealand (pp. 171–178). Springer. https://doi.org/10.1007/978-3-642-21869-9_24
- Limbu, B. H., Jarodzka, H., Klemke, R., & Specht, M. (2019). Can you ink while you blink? Assessing mental effort in a sensor-based calligraphy trainer. *Sensors*, 19(14), 3244. <https://doi.org/10.3390/s19143244>
- Looi, C.-K., Wong, L.-H., & Milrad, M. (2015). Guest editorial: Special issue on seamless, ubiquitous, and contextual learning. *IEEE Computer Architecture Letters*, 8(01), 2–4. <https://doi.org/10.1109/TLT.2014.2387455>
- Lopes, P., & Baudisch, P. (2017). Interactive systems based on electrical muscle stimulation. *Computer*, 50(10), 28–35. <https://doi.org/10.1109/MC.2017.3641627>
- Martinez-Maldonado, R., Echeverria, V., Fernandez Nieto, G., & Buckingham Shum, S. (2020). From data to insights: A layered storytelling approach for multimodal learning analytics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI 2020)*, 25–30 April 2020, Honolulu, HI, USA (pp. 1–15). New York: ACM. <https://doi.org/10.1145/3313831.3376148>
- Martinez-Maldonado, R., Hernandez-Leo, D., Pardo, A., & Ogata, H. (2017). Second cross-LAK: Learning analytics across physical and digital spaces. In *Proceedings of the Seventh International Conference on Learning Analytics and Knowledge (LAK 2017)*, 13–17 March 2017, Vancouver, BC, Canada (pp. 510–511). New York: ACM. <https://doi.org/10.1145/3027385.3029432>
- Martinez-Maldonado, R., Hernandez-Leo, D., Pardo, A., Suthers, D., Kitto, K., Charleer, S., ... Ogata, H. (2016). Cross-LAK: Learning analytics across physical and digital spaces. In *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge (LAK 2016)*, 25–20 April 2016, Edinburgh, UK (pp. 486–487). <https://doi.org/10.1145/2883851.2883855>
- 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. <https://doi.org/10.1080/07370024.2017.1338956>

- Mohan, A., Sun, H., Lederman, O., Full, K., & Pentland, A. (2018). Measurement and feedback of group activity using wearables for face-to-face collaborative learning. In *Proceedings of the IEEE 18th International Conference on Advanced Learning Technologies (ICALT 2018)*, 9–13 July 2018, Mumbai, India (pp. 163–167).
- Muñoz-Cristóbal, J. A., Rodríguez-Triana, M. J., Bote-Lorenzo, M. L., Villagrà-Sobrino, S. L., Asensio-Pérez, J. I., & Martínez-Monés, A. (2017). Toward multimodal analytics in ubiquitous learning environments. *CEUR Workshop Proceedings*, 1828, 60–67.
- Nigay, L., & Coutaz, J. (1993). A design space for multimodal systems: Concurrent processing and data fusion. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, Amsterdam, Netherlands (pp. 172–178). New York: ACM. <https://doi.org/10.1145/169059.169143>
- Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2019). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior*, 100, 298–304. <https://doi.org/10.1016/j.chb.2018.12.019>
- Obrenovic, Z., & Starcevic, D. (2004). Modeling multimodal human-computer interaction. *Computer*, 37(9), 65–72. <https://doi.org/10.1109/MC.2004.139>
- Ochoa, X. (2017). Multimodal learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.), *Handbook of learning analytics* (pp. 129–141). Society for Learning Analytics Research (SoLAR). <https://doi.org/10.18608/hla17.011>
- Ochoa, X., Chiluitza, K., Granda, R., Falcones, G., Castells, J., & Guamán, B. (2018). Multimodal transcript of face-to-face group-work activity around interactive tabletops. *CEUR Workshop Proceedings*, 2163, 1–6.
- Ochoa, X., Chiluitza, K., Méndez, G., Luzardo, G., Guamán, B., & Castells, J. (2013). Expertise estimation based on simple multimodal features. In *Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI 2013)*, 9–13 December 2013, Sydney, Australia (pp. 583–590). New York: ACM. <https://doi.org/10.1145/2522848.2533789>
- Ochoa, X., & Dominguez, F. (2020). Controlled evaluation of a multimodal system to improve oral presentation skills in a real learning setting. *British Journal of Educational Technology*, 51(5), 1615–1630. <https://doi.org/10.1111/bjet.12987>
- Oviatt, S. (2018). Ten opportunities and challenges for advancing student-centered multimodal learning analytics. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI 2018)*, 16–20 October 2018, Boulder, CO, USA (pp. 87–94). New York: ACM. <https://doi.org/10.1145/3242969.3243010>
- Oviatt, S., DeAngeli, A., & Kuhn, K. (1997). Integration and synchronization of input modes during multimodal human-computer interaction. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 1997)*, 22–27 March 1997, Atlanta, GA, USA (pp. 415–422). New York: ACM. <https://doi.org/10.1145/258549.258821>
- Pérez-Sanagustín, M., Santos, P., Hernández-Leo, D., & Blat, J. (2012). 4SPPIces: A case study of factors in a scripted collaborative-learning blended course across spatial locations. *International Journal of Computer-Supported Collaborative Learning*, 7(3), 443–465. <https://doi.org/10.1007/s11412-011-9139-3>
- Prieto, L. P., Martínez-Maldonado, R., Spikol, D., Hernández Leo, D., Rodríguez-Triana, M. J., & Ochoa, X. (2017). Editorial: Joint Proceedings of the Sixth Multimodal Learning Analytics (MMLA) Workshop and the Second Cross-LAK Workshop. *CEUR Workshop Proceedings*, 1828, 1–3.
- Richey, C., D'Angelo, C., Alozie, N., Bratt, H., & Shriberg, E. (2016). The SRI speech-based collaborative learning corpus. In *Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH 2016)*, 8–12 September 2016, San Francisco, CA, USA. (pp. 1550–1554). <https://doi.org/10.21437/Interspeech.2016-1541>
- Scherer, S., Worsley, M., & Morency, L.-P. (2012). First international workshop on multimodal learning analytics. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction (ICMI 2012)*, 22–26 October 2012, Santa Monica, CA, USA (pp. 609–610). New York: ACM. <https://doi.org/10.1145/2388676.2388803>
- 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. <https://doi.org/10.5281/zenodo.3554729>
- Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning*, 8(4), 375–397. <https://doi.org/10.1007/s11412-013-9181-4>
- Schneider, B., & Pea, R. (2014). The effect of mutual gaze perception on students' verbal coordination. In *Proceedings of the Seventh International Conference on Educational Data Mining (EDM 2014)*, 4–7 July 2014, London, UK (pp. 138–144). Retrieved from https://educationaldatamining.org/EDM2014/uploads/procs2014/long%20papers/138_EDM-2014-Full.pdf

- 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>
- Sharma, K., & Giannakos, M. (2020). Multimodal data capabilities for learning: What can multimodal data tell us about learning? *British Journal of Educational Technology*, 51(5), 1450–1484. <https://doi.org/10.1111/bjet.12993>
- Sharma, K., Giannakos, M., & Dillenbourg, P. (2020). Eye-tracking and artificial intelligence to enhance motivation and learning. *Smart Learning Environments*, 7, 1–19. <https://doi.org/10.1186/s40561-020-00122-x>
- Smith, J., Bratt, H., Richey, C., Bassiou, N., Shriberg, E., Tsiartas, A., ... Alozie, N. (2016). Spoken interaction modeling for automatic assessment of collaborative learning. In *Proceedings of Speech Prosody 2016*, 31 May–3 June 2016, Boston, MA, USA (pp. 277–281). <http://doi.org/10.21437/SpeechProsody.2016-57>
- Tsai, Y.-S., Whitelock-Wainwright, A., & Gašević, D. (2020). The privacy paradox and its implications for learning analytics. In *Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK 2020)*, 23–27 March 2020, Frankfurt, Germany (pp. 230–239). New York: ACM. <https://doi.org/10.1145/3375462.3375536>
- Vrzakova, H., Amon, M. J., Stewart, A., Duran, N. D., & D'Mello, S. K. (2020). Focused or stuck together: Multimodal patterns reveal triads' performance in collaborative problem solving. In *Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK 2020)*, 23–27 March 2020, Frankfurt, Germany (pp. 295–304). New York: ACM. <https://doi.org/10.1145/3375462.3375467>
- Vujovic, M., & Hernández-Leo, D. (2019). Shall we learn together in loud spaces? Towards understanding the effects of sound in collaborative learning. In *Proceedings of the Computer-Supported Collaborative Learning Conference (CSCL 2019)*, 17–21 June 2019, Lyon, France (pp. 891–892). <https://repository.isls.org/handle/1/4551>
- Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, 12(4), 423–467. <https://doi.org/10.1007/s11412-017-9267-5>
- Worsley, M. (2012). Multimodal learning analytics: Enabling the future of learning through multimodal data analysis and interfaces. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction (ICMI 2012)*, 22–26 October 2012, Santa Monica, CA, USA (pp. 353–356). New York: ACM. <https://doi.org/10.1145/2388676.2388755>
- Worsley, M. (2018). Multimodal learning analytics' past, present, and potential futures. In *Companion Proceedings of the Eighth International Conference on Learning Analytics and Knowledge (LAK 2018)*, 5–9 March 2018, Sydney, Australia. SoLAR. <http://bit.ly/lak18-companion-proceedings>
- Worsley, M., Abrahamson, D., Blikstein, P., Grover, S., Schneider, B., & Tissenbaum, M. (2016). Situating multimodal learning analytics. In *Proceedings of the 2016 International Conference for the Learning Sciences (ICLS 2016)*, 20–24 June 2016, Singapore (pp. 1346–1349). https://www.isls.org/icls/2016/docs/ICLS2016_Volume_2.pdf
- Worsley, M., & Blikstein, P. (2011a). *Toward the development of learning analytics: Student speech as an automatic and natural form of assessment*. Paper presented at the Annual Meeting of the American Education Research Association (AERA 2010), Denver, CO, USA. Retrieved from http://marcelworsley.com/papers/aera_2011.pdf
- Worsley, M., & Blikstein, P. (2011b). What's an expert? Using learning analytics to identify emergent markers of expertise through automated speech, sentiment and sketch analysis. In *Proceedings of the Fourth Annual Conference on Educational Data Mining (EDM 2011)*, 6–8 July 2011, Eindhoven, Netherlands (pp. 235–240). https://educationaldatamining.org/EDM2011/wp-content/uploads/proc/edm2011_paper18_short_Worsley.pdf
- Worsley, M., & Blikstein, P. (2013). Towards the development of multimodal action based assessment. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK 2013)*, 8–13 April 2013, Leuven, Belgium (pp. 94–101). New York: ACM. <https://doi.org/10.1145/2460296.2460315>
- Worsley, M., & Blikstein, P. (2014). Analyzing engineering design through the lens of computation. *Journal of Learning Analytics*, 1(2), 151–186. <https://doi.org/10.18608/jla.2014.12.8>
- Worsley, M., & Blikstein, P. (2018). A multimodal analysis of making. *International Journal of Artificial Intelligence in Education*, 28(3), 385–419. <https://doi.org/10.1007/s40593-017-0160-1>
- Worsley, M., & Ochoa, X. (2020). Towards collaboration literacy development through multimodal learning analytics. In *Companion Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK 2020)*, 23–27 March 2020, Frankfurt, Germany (pp. 585–595). SoLAR. https://www.solaresearch.org/wp-content/uploads/2020/06/LAK20_Companion_Proceedings.pdf
- Worsley, M., Scherer, S., Morency, L.-P., & Blikstein, P. (2015). Exploring behavior representation for learning analytics. In *Proceedings of the 2015 International Conference on Multimodal Interaction (ICMI 2015)*, 9–13 November 2015, Seattle, WA, USA (pp. 251–258). New York: ACM. <https://doi.org/10.1145/2818346.2820737>
- Wu, H.-Y., Rubinstein, M., Shih, E., Gutttag, J., Durand, F., & Freeman, W. (2012). Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (TOG)*, 31(4), 1–8. <https://doi.org/10.1145/2185520.2185561>

Xu, T., White, J., Kalkan, S., & Gunes, H. (2020). *Investigating bias and fairness in facial expression recognition*. arXiv:2007.10075. https://doi.org/10.1007/978-3-030-65414-6_35