Situating Multimodal Learning Analytics

Marcelo Worsley, University of Southern California, worsley@usc.edu
Dor Abrahamson, University of California, Berkeley, dor@berkeley.edu
Paulo Blikstein, Stanford University, paulob@stanford.edu
Shuchi Grover, SRI International, shuchi.grover@sri.com
Bertrand Schneider, Stanford University, schneibe@stanford.edu
Mike Tissenbaum, University of Wisconsin, Madison, miketissenbaum@gmail.com

Abstract: The digital age has introduced a host of new challenges and opportunities for the learning sciences community. These challenges and opportunities are particularly abundant in multimodal learning analytics (MMLA), a research methodology that aims to extend work from Educational Data Mining (EDM) and Learning Analytics (LA) to multimodal learning environments by treating multimodal data. Recognizing the short-term opportunities and long-term challenges will help develop proof cases and identify grand challenges that will help propel the field forward. To support the field's growth, we use this paper to describe several ways that MMLA can potentially advance learning sciences research and touch upon key challenges that researchers who utilize MMLA have encountered over the past few years.

Introduction

Multimodal learning analytics (MMLA) (Blikstein & Worsley, in press.; Blikstein, 2013; Worsley, 2012) sits at the intersection of three ideas: multimodal teaching and learning, multimodal data, and computer-supported analysis. At its essence, MMLA utilizes and triangulates among non-traditional as well as traditional forms of data in order to characterize or model student learning in complex learning environments. However, as we describe later, the ways that researchers utilize multimodal data vary widely.

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.

However, the ability to tractably study multimodal teaching and learning is largely limited by how well we can perceive, qualify, and quantify multimodal data. Multimodal data includes anything from gestures, to speech, to emotions, to data about social interactions. (Note: For more detailed examples of multimodal data used in MMLA see Blikstein & Worsley (in press), Blikstein (2013) or Worsley (2012)). In years past, analyzing multimodal teaching and learning was completed using human inference and coding. Human observation and analysis provide the innate ability to situate, contextualize and interpret the multimodal data that is emerging from a given learning scenario. MMLA aims to maintain the richness and highly contextualized nature of traditional human-mediated qualitative analysis but with the added benefits of: (1) quantifying those data in news ways: and (2) leveraging innovative sensors to capture data that is not easily perceivable through human observation (e.g., galvanic skin response, eye gaze). To this end, recent developments in non-invasive multimodal sensors have made the process of capturing rich qualitative process data more tractable.

Finally, MMLA is heavily influenced by emerging capabilities in Big Data and computational analysis. For example, the resurgence of machine learning and artificial intelligence provides new ways for analyzing the wealth of multimodal data that can be collected through low-cost, non-invasive sensors. However, as we will describe in the next section, the affordances of computer-supported analysis are not merely in the form of black-box predictions. Instead, computer-supported analysis can mean anything from normalizing the data, to making it easier to interpret by human inference, to a fully automated system that provides real-time help or feedback.

Applications of multimodal learning analytics

The different ways for using MMLA that we describe below are not exhaustive nor are they limited to multimodal analysis. Instead, we attempt to provide a glimpse of some various ways that researchers are using MMLA so as to suggest how MMLA, and learning analytics writ large, could be used more broadly.

Visualizing/Representing information for human inference

 © ISLS

While many would assume that MMLA necessarily eliminates the need for human inference, this is absolutely not the case. As the reader will note from this sub-section and others, human inference and judgment play a very important role in MMLA. For example, MMLA could be used to support more traditional qualitative analysis by presenting the researcher with normalized bio-physiology data to interpret alongside the audio transcripts and raw video footage. In particular, if a researcher were utilizing electro-dermal activation data, it would be beneficial to transform those data so as to account for individual differences in stress response and allow the researcher to visualize those data in a synchronous fashion with the other data streams. The researcher could then go about interpreting each participant's experience using data that would otherwise be unknown to an observer. Similarly, teachers and students could use synchronized views across modalities to better interpret learner and instructor experiences. One example of software that has specifically been developed for the purpose of supporting improved visualization of multimodal data is Chronoviz (Fouse, 2011).

Prediction of indicators
A very common strategy within computational analysis is to acquire a data set, code that data for something in particular (e.g., emotions from videos: happy, sad, angry, confused, surprised, etc.), and then utilize those hand-coded data to automatically code a future dataset. The ability to automatically make a prediction about someone's perceived emotional state, for example, can then become a piece of information used by a qualitative researcher, the focal point of a computational analysis, or used to control a data-driven application (described in the next section). Worsley & Blikstein (2015) is an example of work that uses predictions of indicators. Specifically, they used automatically derived facial expressions to compare the frequency that students seemed to express confusion in a dyadic hands-on learning activity.

Data-driven interventions
In the era of data-driven applications, one cannot help but consider the potential for using behavior-level predictions to determine an appropriate form of intervention or feedback that a learner should receive. For example, a system could analyze the combination of speech, video, and text-based responses that a student provides in order to determine if the student has achieved proficiency with a given skill. Based on that calculation of proficiency, the system could then recommend a future set of activities for the student to complete, much in line with existing intelligent tutoring systems.

Constructing models of interaction
Instead of automating feedback, a common objective with MMLA is to model learner experiences. Modeling can inform the design of a given space or improve one's understanding of a given theory or conjecture. In such cases, indicator predictions can be utilized within a Hidden Markov Model (HMM), for example, to model the ways that students transition from one behavior to the next. For example, recent work by Tissenbaum, Kumar, and Berland (under review) developed an HMM of learner productivity while participating in a multi-touch, tabletop experience. In this work, they also created a Markov Model to provide insight into the types of challenges learners face when participating with tabletop applications.

Evaluating conjecture-based learning designs
As briefly suggested in the previous section, a particular opportunity that comes about through MMLA is the ability to study, with an enhanced level of complexity, multimodal embodied learning experiences. Rather than rely merely on the measurement of pre- to post-test learning gains, or on human coding, MMLA can provide a means for an increasingly detailed analysis. One example of this is recent work by Abrahamson, Shayan, Bakker, and Van der Schaaf (in press), in which the convergence of action logging, eye-tracking, and video analysis created a means for deconstructing learner experiences with a sufficiently high level of specificity that the authors were able to confirm their conjecture of students creating, seeing, and manipulating "attentional anchors," imaginary objects that the students were visualizing on a computer screen.

The above paradigms range from being purely in support of a qualitative analysis to fully automated analyses. Hence, the purpose of MMLA is not to prescribe a certain set of research questions but, instead, to support effective responding to the diversity of research questions examined through the learning science.

Connecting learning with multimodal data
One prerequisite for using MMLA is having multimodal data. However, going from data to results will typically require the use of new analytic tools and techniques. To help elucidate this process, we use the following paragraphs to provide a short overview of how we conceptualize the relationship between learning-related constructs and the multimodal data that we capture and analyze.

Learning constructs

As education researchers, our primary objective is to conduct analyses that ultimately contribute to the field's understanding of teaching and learning. In particular, we are often looking to make a conjecture about some learning-related skill, attribute or construct. For example, in a given study we may wish to study student conceptual change, or student identity. For the purposes of this paper, we will refer to these skills, attributes or constructs as "learning constructs."

Indicators

Similar to any analysis, in order to ascertain the development of a given "learning construct" we will look for one or more indicators. For example, in the case of documenting conceptual change, one indicator could be a change in the explanation that a student provides for a given phenomenon. This could be language-based (i.e., in the words and justification that the student provides) or pictorial (i.e., demonstrated through comparing two drawings that the student made of the phenomenon, one before an intervention and another after the intervention). The point here is simply that for a given "learning construct" there can be any number of potential indicators that we use to prove or disprove a given "learning construct." These indicators are not synonymous with the learning construct, but tend to help us in making inferences related to the construct in question. In the same way that we have indicators in traditional education research, we also have indicators in MMLA. These indicators could be system-generated predictions from text, audio, stress levels, emotions, or any number of other data streams.

Analytic techniques, tools, and data

Indicators are typically generated by one or more analytic techniques, as made available through a given analytic tool and one or more forms of data. Hence, what MMLA is providing is not necessarily a completely new approach for conducting research as much as it is enabling researchers to tap into a new, or complementary, set of indicators, as derived through computational analyses of multimodal data.

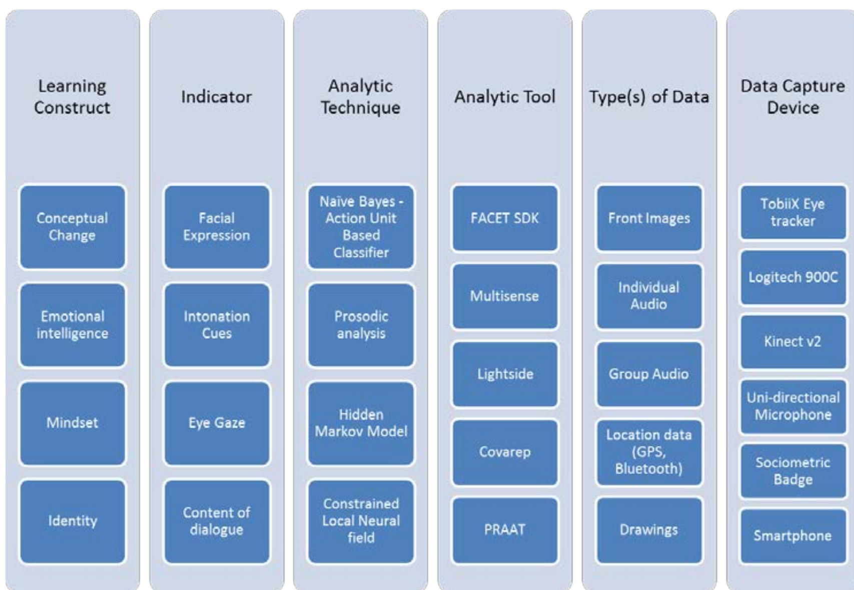

Figure 1. Sample Landscape of Constructs, Indicators, Techniques, Tools, Data Types and Data Capture Tools.

Figure 1 provides a simple representation of the relationship between learning constructs, indicators, analytic techniques, analytic tools, data, and data-capture devices. Identifying the appropriate mapping from "learning construct" to data and data capture tool is an important consideration, as each decision point can impact the resultant analysis. Similarly, recognizing that a single construct can comprise several indicators, and that a given indicator can involve multiple analytic techniques and/or multiple types of data, adds to the complexity of MMLA. Some of the other challenges that have been raised by the MMLA community are related to questions of data privacy, cost, data synchronization, and how to keep the data capture process sufficiently naturalistic.

Examples of emerging research

Even with the aforementioned challenges, a number of researchers have begun to identify important findings from MMLA-related studies. For example, Schneider et al. (in press) conducted innovative research that includes the use of mobile eye-trackers and audio data with dyads using a tangible user interface. In their study they compared low and high-performing dyads, and were able to draw important insights about how to automatically analyze collaboration quality using MMLA techniques. Grover et al. (in press) are conducting a similar line of research around collaborative problem solving in K-12 by combining clickstream, gaze and gesture with measures of proximity, engagement, and turn-taking during pair programming. Through this work, they are modeling the practices associated with high performance collaborations, paying particular attention to the constituents of high collaboration interactions.

Conclusion

There is a host of ways for utilizing MMLA to advance the learning sciences. In recent years researchers have used MMLA to model student performance, predict student learning, and construct models of student–student/student–artifact interactions (Grafsgaard, 2014; Schneider & Blikstein, in press; Worsley & Blikstein, 2015a). Utilizing these techniques enables researchers to study complex learning environments through a different set of lenses that could serve as a strong complement to existing work in the learning sciences.

References

Abrahamson, D., Shayan, S., Bakker, A., & Van der Schaaf, M. F. (in press). Eye-tracking Piaget: Capturing the emergence of attentional anchors in the coordination of proportional motor action. *Human Development*.

Blikstein, P. (2013). Multimodal learning analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13*, 102. doi:10.1145/2460296.2460316

Blikstein, P., & Worsley, M. (n.d.). Multimodal Learning Analytics: a methodological framework for research in constructivist learning. *Journal of Learning Analytics*.

Fouse, A. S. (2011). ChronoViz : A system for supporting navigation of time-coded data. *Chi*, 1–6. doi:10.1145/1979742.1979706

Grafsgaard, J. F. (2014). Multimodal Analysis and Modeling of Nonverbal Behaviors During Tutoring. In *Proceedings of the 16th International Conference on Multimodal Interaction* (pp. 404–408). New York, NY, USA: ACM. doi:10.1145/2663204.2667611

Grover, S., Bienkowski, M, Tamrakar, A., Siddiquie, B., Salter, D. & Divakaran, A. (2016). Multimodal Analytics to Study Collaborative Problem Solving in Pair Programming. *To be presented at the 6th International Learning Analytics & Knowledge Conference (2016)*, Edinburgh, UK. ACM.

Kirsh, D. (2011). Thinking with the body. *Proceedings of the 9th ACM SIGCHI Italian Chapter International Conference on Computer-Human Interaction Facing Complexity - CHItaly*, 15. doi:10.1145/2037296.2037303

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc.

Schneider, B., & Blikstein, P. (in press). Unraveling Students' Interaction Around a Tangible Interface using Multimodal Learning Analytics. *Journal of Educational Data Mining*.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (accepted). Detecting Collaborative Dynamics Using Mobile Eye-Trackers. In *Proceedings of the 12th International Conference of the Learning Sciences.*

Tissenbaum, M., Kumar, V. & Berland, M. (*submitted*) Modeling Visitor Behavior in a Game-Based Engineering Museum Exhibit with Hidden Markov Models. *Submitted to the The Ninth International Conference on Educational Data Mining*. Raleigh, NC.

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* (pp. 353–356).

Worsley, M., & Blikstein, P. (2015a). Leveraging Multimodal Learning Analytics to Differentiate Student Learning Strategies. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 360–367). New York, NY, USA: ACM. doi:10.1145/2723576.2723624

Worsley, M., & Blikstein, P. (2015b). Using learning analytics to study cognitive disequilibrium in a complex learning environment. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15*, 426–427. doi:10.1145/2723576.2723659

 © ISLS