

Chapter 14

Collecting Unobtrusive Data: What Are the Current Challenges?



Roberto Martinez-Maldonado 

Abstract Sensing technologies are rapidly dropping in price and improving the quality of data acquisition. It is therefore expected that sensing technologies, paired with artificial intelligence algorithms, will become a common part of the educational researcher’s toolkit to unobtrusively measure learning phenomena in years to come. In this section, we learned about the potential of using multimodal and multichannel data to create rich models of higher-order constructs, namely engagement, self-regulated learning (SRL) and collaboration. The five chapters showcased various applications of sensing technologies and logging mechanisms to generate indicators that can be critical for studying and supporting learning.

Keywords Unobtrusive observation · Data collection · Self-regulated learning · Collaboration · Multimodal data

1 A Critical Overview of the Chapters

In Chap. 8, Prahraj et al. described some critical steps towards automating the generation of collaboration indicators based on audio data, making inroads into designing end-user interfaces that teachers and students could use. The provision of end-user multimodal analytics interfaces is rare, partly because of the complexity of transforming low-level data into meaningful information. Prahraj et al.

R. Martinez-Maldonado (✉)
Monash University, Clayton, Australia
e-mail: Roberto.MartinezMaldonado@monash.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

V. Kovanovic et al. (eds.), *Unobtrusive Observations of Learning in Digital Environments*, Advances in Analytics for Learning and Teaching,
https://doi.org/10.1007/978-3-031-30992-2_14

demonstrated how this could be done for the case of microphone data by cleaning less important words (e.g., stop words) and distilling keywords that can indicate topics that may mean something to educational stakeholders.

As sensor data is increasingly being used in educational research, a much-needed discussion about the challenges of extracting meaningful information from sensor data, such as electrodermal activity (EDA) data, has been provided by Lee in Chap. 9. Most of the EDA sensors used in educational research have not (so far) been designed for highly dynamic educational activities such as those that commonly occur in maker spaces. Therefore, it is crucial to recognise that sensor data is commonly noisy and incomplete. Deriving indicators from such fuzzy data sources requires lots of exploration, like the one described in Lee's chapter. To provide context, Lee explored the use of wearable physiological wristbands and wearable still image cameras to enrich the engagement analysis in a physical learning space and complement the sensor data.

As shown by Salehian Kia et al. in Chap. 10, multichannel data itself could serve researchers to establish the validity of the mapping from low-level trace data to higher-order constructs, such as the phases of the SRL process. If the same learning event can be observed on multiple data channels, the data streams can validate one another or add meaning to low-level data. This approach can help researchers perform a deeper analysis of logged activity happening in various digital spaces (e.g., at the learning management system, an assessment platform and a digital textbook) and provide effective interventions in the future.

With the growing amount and diversity of educational data, it is critical to ensure that the use of such data is grounded in sound educational theories. In Chap. 11, Wiedbusch et al. provided a theoretically grounded approach for measuring engagement with multimodal data originating from self-reports, log streams, oculometrics, physiological sensors, facial expressions, body gestures, think-aloud and emotelouds. Their goal is to measure SRL engagement by considering all the aspects of students, including what happens inside their heads as well as emotional and social aspects that are often hard to inspect with the 'naked eye'. Authors envision a set-up where multiple sensors, computer vision algorithms and educational tools are part of a synchronised ecosystem capable of recognising the behavioural, cognitive, emotional and agentic engagement states of learners to provide some specific support while students work in front of the computer.

In this regard, in Chap. 12, Malmberg et al. also discussed their stance on using multimodal data to study engagement from a collaborative learning perspective as students experience socio-emotional interactions. Like Lee, the authors also use EDA sensors, but this time, to analyse how physiological synchrony can aid in understanding the relationship between cognitive, socio-emotional and interaction episodes in group-level regulation. The authors illustrate how multimodal data can augment conversation analysis which has been a staple technique used to study collaborative learning.

2 Using Multimodal Data for Unobtrusive Measurement of Learning: Where Are We Now?

Overall, using multimodal data to unobtrusively measure learning can be seen as an emerging and exciting area still in its infancy. One of the key challenges that researchers in this area often face concerns the ecological *validity of indicators* and the process of imbuing the data captured from sensors with meaningful constructs relevant to a specific learning context for educational stakeholders to make sense of resulting algorithmic outputs (Cukurova et al., 2020; Yan et al., 2022; Di Mitri et al., 2018). We can minimise noisy data if we conduct controlled experiments and take the time to carefully analyse and manually fill the gaps in multimodal data. However, eventually, we want to close the analytics loop by supporting teachers and students where learning happens. Lab studies will continue to be critical for investigating the validity of specific indicators and advancing educational research (Sharma & Giannakos, 2020). However, to make a practical impact, multimodal studies need to make it into the actual classroom, where multiple confounding variables can be introduced as they reflect the authentic (often ‘messy’) conditions in which learning ultimately occurs (Worsley et al., 2021). Chejara et al. (2023) and Martinez-Maldonado et al. (2023) have recently discussed the several practical, logistic and ethical challenges that can be identified only when multimodal technologies are deployed *in-the-wild*, which can strongly shape the ultimate effectiveness of multimodal data-enhanced educational interventions. Nonetheless, enriching authentic learning spaces with multimodal analytics and data sensing capabilities can hold the potential to help researchers study the umbrella of expected and unexpected events that can shape learning.

As multimodal data can help us create a richer picture of the learning activity, it can also increase the *complexity of the potential pedagogical intervention*. This points to a second key challenge: how can we create fully automated multimodal tools that provide some direct benefit to teachers and students? To address this challenge, there is a need to design end-user interfaces more carefully to help audiences who usually are not formally trained in data analysis gain insight into multimodal educational data. In several studies where multimodal data is used, humans still need to be part of the pre-processing data process (i.e., see review by Praharaj et al., 2021). While this adds validity and rigour to a study from an educational research standpoint, it also makes innovation *in the wild* more challenging. Fully automating the whole analytics process, from multimodal or multichannel data acquisition to creating interfaces that are ‘easy to use’ and meaningful to end-users in the educational sector, requires a multidisciplinary team of experts in data science, human-computer interaction and education (Yan et al., 2022). Unfortunately, not all research centres have the resources to form such multidisciplinary teams. As a result, the collaboration between researchers and practitioners in multimodal learning analytics is crucial for advancing the field and keeping it thriving.

A third challenge illustrated across these previous chapters is related to the *critical role of the educational context* in giving meaning to multimodal data. While one

data channel can help provide context to another channel, the ultimate meaning of any indicator extracted from data depends on the learning task and, hence, on the pedagogical intentions of the learning design (Ochoa, 2022). For example, the detection of the quality of collaboration is highly contextual. Thus, collaboration indicators can be identified based on the learning goals (the learning design) and using educational and teamwork theories. Theoretical constructs, such as those found in SRL and collaborative learning theory, can give meaning to the indicators obtained from fuzzy physiological data (e.g., Azevedo et al., 2022). Ultimately, education is highly contextual. Therefore, it is not expected to treat multimodal learning analytics innovations as one type of solution that can be applied to multiple contexts but as an approach for embracing the complexity and particularities of each educational context. These multimodal innovations also highlight the limitations of just analysing the clickstreams and keystrokes that students perform in the learning management system by considering the broader context of using the socio-technical context where learning happens (Echeverria et al., 2019).

These and other challenges in multimodal learning analytics research can be seen as opportunities yet to be explored. In any case, the current technical limitations in sensing technologies and analysis approaches are being addressed by the rapid progress of artificial intelligence (AI). For example, the automated transcription problem that previously hindered educational researchers from developing fully automated tools to aid face-to-face collaboration is now close to being entirely resolved (Southwell et al., 2022). Improvements in human voice detection, noise filtering, speaker diarisation and automated transcription algorithms generate conversation logs similar to those of professional human transcription services. The physiological wristbands currently used in educational research, not specifically created for educational purposes, will soon be replaced by better sensors less susceptible to the physical activity and ambient conditions found in most classrooms. Advances in ubiquitous and pervasive computing will only further augment our capacity to gather more and more data. Hence, it will be critical to further advance approaches for extracting meaning from multimodal data while also considering the ethical implications of using such data.

Indeed, a topic that has not been deeply covered in the chapters presented in this section involves the ethical and privacy implications of unobtrusively gathering multimodal sensor and log data from students. Just because we can capture more data does not mean we should do it. If we do it, much more discussion about who owns these data is required. The danger of excessive surveillance is genuine, and it is still difficult to predict all the possible scenarios concerning what education providers can do with such detailed and frequently sensitive data. What limitations will be placed on the utilisation of these data? Most importantly, if the aim is to create more accurate learning models, what will happen for students that can be considered 'outliers' from a data science perspective but may be demonstrating unique learning pathways from a social science perspective? Finally, there is a pressing need for more dialogue on how to involve teachers, students and other educational stakeholders in the design process of multimodal learning analytics (Echeverria et al., 2022). Several important decisions are taken during the design process of any programmable tool. How can the values of educational stakeholders be considered

in the design process? This calls for a human-centred design perspective that provides a channel for relevant stakeholders easily affected by data-intensive initiatives to voice their concerns and remain active agents in the learning process rather than passive receivers. Since AI and sensing technologies are already impacting educational practices, we must create systems that exploit these technologies with integrity.

References

- Azevedo, R., Bouchet, F., Duffy, M., Harley, J., Taub, M., Trevors, G., Cloude, E., et al. (2022). Lessons learned and future directions of metatutor: Leveraging multichannel data to scaffold self-regulated learning with an intelligent tutoring system. *Frontiers in Psychology, 13*.
- Chejara, P., Prieto, L. P., Rodríguez-Triana, M. J., Ruiz-Calleja, A., Kasepalu, R. & Shankar, S. K. (2023). Multimodal Learning Analytics research in the wild: challenges and their potential solutions. In *LAK 2023 workshop: Leveraging multimodal data for generating meaningful feedback* (pp. 1–5).
- Cukurova, M., Giannakos, M., & Martinez-Maldonado, R. (2020). The promise and challenges of multimodal learning analytics. *British Journal of Educational Technology, 51*(5), 1441–1449.
- Di Mitri, D., Schneider, J., Specht, M., & Drachslar, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning, 34*(4), 338–349.
- Echeverria, V., Martinez-Maldonado, R., & Buckingham Shum, S. (2019). Towards collaboration translucence: Giving meaning to multimodal group data. In *Proceedings of the 2019 SIGCHI conference on human factors in computing systems* (pp. 1–16).
- Echeverria, V., Martinez-Maldonado, R., Yan, L., Zhao, L., Fernandez-Nieto, G., Gašević, D., & Shum, S. B. (2022). HuCETA: A framework for human-centered embodied teamwork analytics. *IEEE Pervasive Computing, 1–11*.
- Martinez-Maldonado, R., Echeverria, V., Fernandez-Nieto, G., Yan, L., Zhao, L., Alfredo, R., Li, X., Dix, S., Jaggard, H., Wotherspoon, R., Osborne, A., Gasevic, D., & Buckingham Shum, S. (2023). Lessons learnt from a multimodal learning analytics deployment in-the-wild. *ACM Transactions on Computer-Human Interaction*. In press.
- Ochoa, X. (2022). Multimodal learning analytics – Rationale, process, examples, and direction. In C. Lang, G. Siemens, A. F. Wise, D. Gašević, & A. Merceron (Eds.), *The handbook of learning analytics* (pp. 54–65). SOLAR.
- Praharaj, S., Scheffel, M., Drachslar, H., & Specht, M. (2021). Literature review on co-located collaboration modeling using multimodal learning analytics—Can we go the whole nine yards? *IEEE Transactions on Learning Technologies, 14*(3), 367–385.
- 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.
- Southwell, R., Pugh, S., Perkoff, E. M., Clevenger, C., Bush, J. B., Lieber, R., Ward, W., Foltz, P. & D’Mello, S. (2022). Challenges and feasibility of automatic speech recognition for modeling student collaborative discourse in classrooms. In *Proceedings of the 15th international conference on educational data mining* (pp. 302–315).
- Worsley, M., Martinez-Maldonado, R., & D’Angelo, C. (2021). A new era in multimodal learning analytics: Twelve core commitments to ground and grow MMLA. *Journal of Learning Analytics, 8*(3), 10–27.
- Yan, L., Zhao, L., Gasevic, D., & Martinez-Maldonado, R. (2022). Scalability, sustainability, and ethicality of multimodal learning analytics. In *LAK22: 12th international learning analytics and knowledge conference* (pp. 13–23).