

The Promise and Challenges of Multimodal Learning Analytics

Mutlu Cukurova, University College London. m.cukurova@ucl.ac.uk

Michail Giannakos, Norwegian University of Science and Technology. michailg@ntnu.no

Roberto Martinez-Maldonado, Monash University. roberto.martinezmaldonado@monash.edu

Introduction

Learning outcomes are the result of interactions between intraindividual factors, such as prior knowledge, emotions, and motivation; as well as the contextual factors in which the learning takes place (Dumas, McNeish, & Greene, 2020). If any, perhaps only a few Learning Analytics (LA) researchers and practitioners would disagree with this shared understanding about learning among the modern learning scientists. After all, as stated in the definition of the field, LA aims to understand and optimise learning and the environments in which learning occurs through the measurement, collection, analysis and reporting of data about learners and their contexts. However, most LA research and practice today relies upon logged data from digital platforms, which oftentimes fall short to help us accurately and timely interpret and account for factors influencing learning. This is particularly the case for factors influencing the contexts in which learners perform and the intraindividual factors that are regularly categorised in the affective domain. Recently, Multimodal Learning Analytics (MMLA) provided promising results to make progress in this space to help us better understand, model, and support learning *in situ*. However, MMLA also poses significant methodological, practical, and ethical challenges. We are aware that MMLA research and practice can lead to significant concerns among educational stakeholders particularly with regards to the invasion of their privacy; bias, fairness, accountability and transparency of the MMLA models; as well as the risk of enabling a “surveillance” culture through constant monitoring of learner behaviours. In an attempt to raise awareness to some of these issues as well as providing some ways forward, we proposed this special issue on the promise and challenges of MMLA.

The Promise

MMLA research leverages various modalities of data both from physical and digital spaces, use computational approaches to process and analyse these multimodal data, and utilise and contribute to theories about the analysis of human behaviours in learning contexts to help us achieve LA’s aspirational goals. Some proponents of MMLA research have conceptually suggested that multimodal data can create more complete models of learning. For example, Drachler and Schneider (2018) suggested that multimodal data in learning analytics can

provide a more holistic picture of learning processes and success factors than the current form of knowledge extracted using individual data sources (e.g. clickstreams or keystrokes). More recent empirical research has demonstrated that fusion of multimodal data brings significantly better prediction of learning outcomes and helps us to interpret complex learning processes. For instance, Giannakos et al., (2019) identified that the prediction of skill-acquisition was far better using fused multimodal data from eye-tracking, EEG and facial video than the prediction using any of the individual streams. Similarly, Cukurova et al. (2019) showed that multimodal data can bring in more accurate predictions of tutor performance in ill-structured learning situations. Two studies in this special issue further contribute to demonstrate the added value of MMLA approaches compared to unimodal approaches. The first study by Emerson et al. (2020) presents an MMLA approach based on system interaction, eye tracking, and facial expression data to predict intrinsic motivation and learning outcomes of students using a game-based learning environment. The authors found that predictive models using two different streams of data often outperform unimodal predictors. Surprisingly, they also found that models that rely on three modalities of data featured a decreased performance. The authors point at the need for more work to understand how features extracted from each modality can (or cannot) contribute to effectively modelling educational constructs. Similarly, the study presented by Olsen et al. (2020), focuses on predicting learning gains in the context of a collaborative intelligent tutoring system by combining features extracted from students' gaze, system log, audio, and dialogue data. The authors extracted temporal features from each data source and created predictive models considering all the possible combinations of the four modalities. They found the inclusion of gaze data in most multimodal models contributed to improved prediction performance. They also found that combining data with different time scales also contribute to improving prediction of learning outcomes.

In addition to more accurate predictions of learning in single observations, MMLA uses the advances in machine learning and sensor technologies to monitor factors that are argued to be significant for learning but are often ignored due to challenges in their dynamic measurement and interpretations. Many constructs behind the cognitive and affective factors that influence how students interact with devices, people, and other resources are critical for learning (e.g., stress levels, prior knowledge, cognitive levels, engagement, motivation, emotions). However, they are challenging to be captured continuously and unobtrusively with traditional educational research methods. To address this, MMLA research focuses on the use of physical sensors, due to their capacity to capture physiological and behavioural manifestations of learners' emotions and traits that cannot be easily seen by the human naked eye such as learners' emotions, gaze, cognitive states and bodily responses. For instance, recent work in MMLA has enabled performing text, speech, handwriting, sketch, gesture, affect, or eye-gaze analysis (e.g. Sharma & Giannakos, 2020; Blikstein & Worsley, 2016, Spikol et al., 2018), in order to extract indicators and use these to model higher-order cognitive, meta-cognitive and affective factors for research purposes or to provide automated feedback to enable learner reflections (e.g. Ochoa et al, 2018). As it appears in this special issue, Sharma and Giannakos (2020) summarise the state-of-the-art in various learning capacities of multimodal data as well as the critical insights about MMLA in the form of a systematic literature review.

It is important to note that the MMLA community is not the first or the only research community engaging with multimodal data and computational approaches to understand and support learning. Neighbouring fields to LA community such as Artificial Intelligence in Education (AIED), Educational Data Mining (EDM), and User Modelling, Adaptation and Personalisation (UMAP) all have examples of research that utilises computationally processed multimodal data. More specifically affective computing research from these communities has a long tradition of utilising more than one data modality to analyse and support learning. So, what makes MMLA unique and different from other relevant research areas?

MMLA: A Sweet-spot at the Intersection of Learning Sciences, Affective Computing, and Human-Computer Interaction

Social reality is the main subject of educational research, but it is interpreted in multiple ways by different paradigms of research. A research paradigm is referred to as a set of beliefs, values, and assumptions that are shared by the members of the research community. Regardless of the specific technological solution we design and use, it is not possible to undertake research without committing to a paradigm and their transparency and consistency are paramount to ensure research quality (Cukurova, 2018). In MMLA, research paradigms are frequently derived from the roots of research in learning sciences, affective computing, and human-computer interaction (HCI).

To start with, the interdisciplinary research from the learning sciences has helped us understand a great deal about the way that humans learn, and as a result, we now have an improved understanding of how best to teach and train people. This knowledge from the learning sciences should inform the collection, analysis, and presentation of multimodal data (Luckin, & Cukurova, 2019). In that way, we are better placed to design MMLA systems to improve teaching and learning practices, but also we can further improve our understanding of the way humans learn. For instance, in this special issue, Larmuseau et al. (2020) utilise the cognitive load theory (Sweller, 1994) from the learning sciences and leverage MMLA with the aim of optimising online problem-solving processes. The authors attempted to measure the cognitive load of learners while they were engaged in online problem-solving activities at different levels of difficulty using physiological data from galvanic skin response, skin temperature, heart rate and heart rate variability. Although, their results showed that heart rate and skin temperature are significantly related to self-reported cognitive load; only physiological data was not enough to distinguish differences in learners' cognitive load in online problem-solving conditions. The authors argued for the value of combining self-reports of associated mental states with physiological data to facilitate interpretations of the results from physiological data. Similarly, Ahn et al., (2020) drew upon the Control Value Theory of Achievement Emotion (Pekrun, 2006) and aligned facial recognition data with eye-tracking data to understand emotional processes and their relationship with learning gains in the context of game-based learning. The authors provided insights into the

frequency and the types of emotions elicited in learning scenarios dealing with a focus on typically undesirable emotions in learning (e.g., anger). The authors showed that the learners may not always behaviorally express emotions through facial expressions, therefore, it is helpful to supplement granular with larger time-frame analyses to examine emotional profiles. Their results not only are driven with theories from the learning sciences, but also contribute back to the theory showing that emotions with positive valence might not always have a positive impact on learning, and vice versa. MMLA's key affordances such as temporality or potential direct access to new indicators of cognitive and affective processes consist of relatively new forms of information in the field. In that vein, MMLA can challenge established "truths" and allow different sensemaking theories to advance the field. In turn, these new theoretical lenses can allow us to enrich our current understanding and further foundational and contemporary properties on how people learn in technology-rich learning environments.

Correspondingly, learning sciences-driven multimodal data's potential to interpret complex social phenomena encouraged many affective computing researchers to try and build models that can process information from multiple modalities through computational techniques and social signal processing (Vinciarelli, Pantic, & Bourlard, 2009). The literature on affective computing is rich with examples of audio-visual speech recognition; multimedia content indexing and retrieval, as well as multimodal affect recognition. Learning from multimodal data provides opportunities to acquire an in-depth understanding of complex processes. Therefore, for affective computing research to make progress, it makes sense to focus on multimodal models that can process and relate information from multiple modalities. This work situated in affective computing research often aims to create systems that in essence replicate the human affect processing and interpretation by making machines capable of processing multimodal data at a similar accuracy. Affective computing research led to great examples of showcasing the potential of machine learning approaches to unravel complicated manifolds in complex educational data including non-linear effects and multivariate interactions. Many affective computing studies also present the superiority of multimodal data over unimodal approaches in terms of predicting learning performance automatically (Baltrušaitis, Ahuja, & Morency, 2019). Such automated affective computing research work is particularly useful for the provision of automated personalised support to learners through intelligent tutoring systems and adaptive learning platforms.

Nevertheless, in essence, most MMLA approaches aim to provide explicit and comprehensible ways of presenting information to learners and teachers to make them more informed decision-makers (Cukurova et al., 2019). This great potential can offer a new form of transhumanist technologies that can enable students and instructors to perceive important insights that can augment their learning and teaching capacities (Eisenberg, 2017). Therefore, MMLA systems by nature should be designed to be tightly coupled with humans to become part of their extended cognition, and ultimately enhance their capabilities in teaching and learning (Cukurova, 2019). Accordingly, as researchers, it is important to create an intellectual track that allows MMLA advances to support learners and be utilised in ways that serve our values and further our abilities. This work requires a greater emphasis on teacher and learner interfaces of MMLA systems as these interfaces need to be internalized by humans. It is important to note that the

interfaces do not necessarily explain how the subsystem of the MMLA exactly works to teachers and learners, but they provide opportunities for cognitive processes to be instantiated and internalized by humans to make the most of the MMLA affordances. In this sense, MMLA systems are less autonomous systems than the aspirations of most affective computing research. Due to the MMLA system's strong emphasis on their tightly coupled nature, MMLA studies require and benefit from fundamental methodologies and principles of HCI research. As a good example of the approach, Martinez-Maldonado et al. (2020) recently explored how data storytelling principles can be applied to the design of an MMLA visual interface to enable students and teachers to gain insights into the complexity of large amounts of sensor and log data.

In addition to the nuances on the intended autonomy of the learning support platform, MMLA research is not limited to affective aspects of computing and equally relevant to the cognitive and metacognitive aspect of learning. For instance, in this special issue, Dindar et al. (2020) investigated the relationship between electrodermal activity (EDA) and a number of self-reported metacognitive experiences and learning performance among small collaborative problem-solving groups with the help of MMLA. The authors found that the only consistent result was the measure of physiological synchrony and the mental effort required by a group of learners during a collaborative problem-solving task. This suggests that further research could use MMLA affordances to provide suggestions for momentary changes from complementary to synchronous group activities, and vice versa, during collaborative problem-solving activities.

At last but not least, perhaps more than ever in this digital era we live in, it is important to keep in mind that learning is not constrained to a digital learning environment. Learning occurs in situ, where the learner is (Sharples, & Roschelle, 2010). Moreover, it happens as a continuous and dynamic process. This requires careful orchestration of the learning in different physical and digital contexts, to enable the learning in one context to be integrated into and complement to another. The increasingly blended nature of learning settings and the inherent temporality of learning processes make it essential to move beyond approaches that rely solely on a unidimensional single data source generated from learner interactions within a digital context to multidimensional and multimodal data sources generated from learners' interactions with all the resources in situ. Although there is plenty of previous articulations of the learner interactions in physical contexts, the empirical evidence from reliable evaluations is limited. However, HCI research is robust and covers all aspects of the interaction between humans and computers (e.g., social, behavioural, cognitive, embodiment) (see the full review in Rogers, 2012). More specific to MMLA research, this includes theories that have been used to understand the complexity of the social learning and interaction ecosystems (e.g. Bardram, 2005; Jetter et al., 2015; Martinez-Maldonado et al., 2016), observable multimodal interaction aspects (e.g. Norris, 2004), and how non-observable cognitive processes and emotional states can play a significant role in reflection and learning (e.g. DeVane, & Squire, 2012; Picard, 1997).

This means that MMLA innovations can be designed based on strong theories that lay at the intersection of HCI and learning sciences with the purpose of guiding not only the data analysis but also the modelling approaches to point at higher-order educational constructs. For example,

Echeverria, et al. (2019) implemented a data structure called the multimodal matrix to map from indoor positioning, physiological, audio and system logging data to a holistic view of collaboration that includes dimensions of collaboration (epistemic, physical, social and affective) as per a theoretical model of networked learning, the ACAD framework (Goodyear, & Carvalho, 2014). A similar mapping was conceptually proposed by Worsley *et al.* (2016) to navigate from low-level sensor data (e.g. extracted from eye-trackers, video/depth cameras and microphones) to learning constructs (e.g. conceptual change, emotional intelligence, mindset and identity) using a range of analytic techniques such as classifiers, prosodic analysis, Hidden Markov Models and constrained local neural fields.

In this special issue, Vujovic et al., (2020) investigated the potential of MMLA for studying the impact of the contextual factors with new analytical methods and new data sets in the domain of collaborative learning for particular characteristics in its learning design. With the help of using a combination of qualitative and quantitative data, the authors captured elementary students' motions, level of participation and movement range in a physical collaborative formal learning setting. Their results showed that the shape of classroom tables has a significant effect on students' learning behaviours. The paper exemplifies the natural connection between the learning space design and the various insights that can be drawn with regards to learning outcomes with the help of MMLA. However, as pointed out by Ochoa and Domingues (2020) MMLA systems are rarely evaluated in physical real-world spaces they are designed for; and existing evaluations often have significant limitations on how they are conducted and reported. As the authors showed in their randomised controlled evaluation of the impact of an MMLA system that aims to help students develop their presentation skills, indeed MMLA tools have the potential to make significant real-world impact on students' skills development. We expect that as the maturity of MMLA systems improves, there will be extensive evaluations of MMLA systems in real-world settings to better understand how such systems are operated with learners and to further improve MMLA systems with insights gained from their robust evaluations.

There is a clear need for further work in the implementation of MMLA systems in authentic spaces where learning occurs (e.g., homes, classrooms, museums), an endeavour that is already acknowledged as challenging (Baker, and Ocumpaugh, 2015). It is clear that this line of work has not yet reached its full capacity, and proper *in situ* setups hold the potential to bridge data-quality and ecological validation. Two important methodological notions that usually antagonise one another (e.g., high data-quality and low ecologically valid lab studies, low data-quality and high ecologically valid in-the-wild studies). To move towards this direction, MMLA community needs to tackle multifaceted challenges of ethical, practical, and methodological nature.

Three main themes of challenges and future research opportunities for MMLA community: Ethical, practical and methodological

During the last few years, we have seen a growing interest and momentum about MMLA. That is also portrayed with the formation of the CrossMMLA SIG and the continuous growth of its respective workshops at LAK conferences (i.e., Spikol et al., 2017; Martinez-Maldonado et al., 2018). The contributions in this special issue cover several dimensions of these MMLA research trends, as also identified from the recent literature review of the field included in this issue (Sharma, & Giannakos, 2020). We expect and welcome further work on these dimensions and would like to conclude this editorial piece reflecting upon three main themes of challenges in MMLA research. As discussed earlier, multimodal data streams are increasingly being deployed in LA research and have a significant potential to advance the field. However, the ethical consequences of MMLA are rarely fully considered. In fact, most MMLA research, development, and deployment currently take place in what is essentially a moral vacuum, with very limited research, no clear guidelines and no regulations to address the specific ethical issues raised by the use of MMLA in real-world practice. For instance, in this special issue, Beardsley et al., (2020), introduced an informed consent comprehension test for MMLA research. This test was put into practice and used to evaluate learners' understanding of informed consent forms and rates of enrolment in an MMLA study. In MMLA research, it is clear that further work is needed to establish adequate protocols and communication approaches that researchers and practitioners can adopt in order to enable students and teachers to be fully aware of the multimodal data they provide while they interact with MMLA systems.

In the same vein, Crescenzi-Lanna (2020) conducted a systematic literature review on MMLA studies with children under six years and analysed the selected contributions based on the degree of performance analytics (students' understanding and engagement), use of machine learning, use of eye-tracking, Kinect, biometrics, human-coded quantitative and qualitative data. The results indicated that, despite its complexity and the various ethical and practical challenges they may lead to (e.g., ethical considerations for the participation of children, parental involvement, potential surveillance issues), multimodal data have valuable contributions on the performance of analytics and face and speech recognition systems in the context of supporting young children's learning. Crescenzi-Lanna (2020) also provides a commentary on some of these issues (e.g., children interacting with a "wizard-of-oz", obtrusive exposure to studies).

We acknowledge that MMLA research and practice can lead to significant concerns among educational stakeholders with regards to the invasion of privacy through various modalities of data collected in learning settings. There might also be concerns with regards to the potential bias built in the computational modelling approaches used in the analysis and interpretation of multimodal data in learning analytics. Furthermore, there might be significant concerns over a potential "surveillance" culture encouraged with the use of MMLA to monitor learner behaviours.

We appreciate the concerns of educational stakeholders and recognise that for MMLA research to make the real-world impact it envisages to make, more research on ethical and practical considerations of MMLA design and implementations is needed.

In addition to the ethical and practical challenges, with increased complexity in data due to its multimodal nature, new methodological challenges also arise. Conducting the data gathering, pre-processing, analysis, annotation, and sense-making, in a way that is meaningful for learning scientists and other stakeholders (e.g., students or teachers), still poses several challenges for the field. Although the equipment is becoming more affordable, there is still sophisticated equipment that is costly and requires specialized technical competence. Moreover, to the best of our knowledge, none of the MMLA studies has yet addressed the key challenge of how to maintain a seamless continuity of the learning experience despite the increasing change in learners' movements within and across various physical, digital and social contexts in their activities. Due to the data-intense nature and focus on interventions in MMLA research, the notions of contextualization and generalizability are of particular importance. We welcome future research investigating different affordances of MMLA research in terms of their potential for a high degree of context-awareness and their generalizability. For instance, recent work from Sharma et al. (2020) provides initial evidence of MMLA's potential to support both context-aware features (e.g. through physiological and facial measures) and a certain degree of generalizability (e.g., through measures with high internal and external validity such as EDA or HRV).

We celebrate the success of MMLA within recent years, but recognise that if we are to stay relevant in the upcoming years, we must find ways to generate considerable real-world impact with our research. This requires significant advancements on various ethical, practical, and methodological challenges we are currently facing in the field, some of which are highlighted in this editorial and the articles contained in this special issue.

Acknowledgements

We would like to thank the reviewers of this special issue for their timely and thorough reviews; and Manolis Mavrikis, the editor-in-chief of BJET for this special issue, for his ongoing support, constructive critique and guidance during the preparation of this special issue.

References

- Ahn, T. B. and Harley, J. (2020). Exploring Emotions and Multimodal Learning Analytics: Eye-tracking and Facial Recognition. *British Journal of Educational Technology*.
- Bardram, J. 2005. Activity-based computing: support for mobility and collaboration in ubiquitous computing. *Personal Ubiquitous Computing*, 9, 5 (September 2005), 312-322.
- Baltrušaitis, T., Ahuja, C., & Morency, L. P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), 423-443.

- Beardsley, M., Martínez-Moreno, J., Vujovic, M., Santos, P. and Hernández-Leo, D. (2020). Enhancing consent forms to support participant decision making in multimodal learning research. *British Journal of Educational Technology*.
- 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.
- Crescenzi-Lanna, L. (2020). Multimodal Learning Analytics research with young children: a systematic review. *British Journal of Educational Technology*.
- Cukurova, M. (2018). A syllogism for designing collaborative learning technologies in the age of AI and multimodal data. In *European Conference on Technology Enhanced Learning* (pp. 291-296). Springer, Cham. Chicago
- Cukurova, M. (2019). Learning analytics as AI extenders in education: Multimodal machine learning versus multimodal learning analytics. In *Proceedings of the Artificial Intelligence and Adaptive Education Conference* (pp. 1-3).
- 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.
- DeVane, B., & Squire, K. D. (2012). Activity Theory in Learning Technologies. *Theoretical foundations of learning environments*, 242.
- Dindar, M., Oulun Y., Järvelä, S., Haataja, E. and Oulun Y. (2020). What does physiological synchrony reveal about metacognitive experiences and group performance? *British Journal of Educational Technology*.
- Drachsler, H., & Schneider, J. (2018). JCAL Special Issue on Multimodal Learning Analytics. *J. Comp. Assisted Learning*, 34(4), 335-337.
- Dumas, D., McNeish, D., & Greene, J. A. (2020). Dynamic measurement: A theoretical–psychometric paradigm for modern educational psychology. *Educational Psychologist*, 55(2), 88-105.
- Echeverria, Vanessa, Roberto Martinez-Maldonado, and Simon Buckingham Shum. (2019). Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems, CHI'19*, (pp. 39, 31-16). ACM.
- Eisenberg, M. (2017). The binding of Fenrir: Children in an emerging age of transhumanist technology. In *Proceedings of the Sixteenth International Conference of Interaction Design and Children, IDC '17* (pp. 329–333). New York, NY: ACM.
- Emerson, A. Azevedo, R. and Lester, J. (2020). Multimodal Learning Analytics for Game-Based Learning. In *British Journal of Educational Technology*
- Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management*, 48, 108-119.

- Goodyear, P. and Carvalho, L. 2014. Framing the analysis of learning network architectures. In *The architecture of productive learning networks*, Lucila Carvalho & Peter Goodyear (Ed.), Routledge, New York, NY, 48-70.
- Jetter, Hans-Christian, Harald Reiterer and Florian Geyer. 2014. Blended Interaction: understanding natural human-computer interaction in post-WIMP interactive spaces. *Personal and Ubiquitous Computing*, 18, 5 (June 2014), 1139-1158.
- Larmuseau, C., Cornelis, J., Lancieri, L., Desmet, P. and Depaepe, F. (2020). Multimodal learning analytics to investigate cognitive load during online problem-solving. *British Journal of Educational Technology*.
- Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, 50(6), 2824-2838.
- Martinez-Maldonado, R., Goodyear, P., J.Kay, Thompson, K., and Carvalho, L. (2016) An Actionable Approach to Understand Group Experience in Complex, Multi-surface Spaces. *SIGCHI Conference: Human Factors in Computing Systems, CHI 2016*, 2062-2074.
- Martinez-Maldonado, R., Echeverria, V., Prieto, L. P., Rodriguez-Triana, M. J., Spikol, D., Cukurova, M., ... & Worsley, M. (2018). 2nd Crossmmla: Multimodal learning analytics across physical and digital spaces. In *CEUR Workshop Proceedings*.
- Nigay, L., & Coutaz, J. (1993). A design space for multimodal systems: Concurrent processing and data fusion. *Proceedings of the SIGCHI Conference on Human Factors in Computing*
- Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. Routledge. *Systems – CHI'93*, (January 1993), 172–178
- Ochoa, X., Domínguez, F., Guamán, B., Maya, R., Falcones, G., & Castells, J. (2018). The RAP system: Automatic feedback of oral presentation skills using multimodal analysis and low-cost sensors. In *Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 360-364).
- Ochoa, X. and Domínguez, F., (2020) Controlled Evaluation of a Multimodal System to Improve Oral Presentation Skills in a Real Learning Setting. *British Journal of Educational Technology*.
- Olsen, J., Sharma, K., Rummel, N. and Alevin, V. (2020). Using Multimodal Data to Temporally Analyze Collaborative Learning Outcomes: Benefits and Challenges. *British Journal of Educational Technology*.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review*, 18(4), 315-341.
- Picard, R. W. (1997): *Affective computing*. Ma, USA, The MIT Press
- Rogers, Y. (2012). *HCI theory: classical, modern, and contemporary*. Morgan & Claypool Publishers.
- Sharma, K. and Giannakos, M. (2020). Multimodal data capabilities for learning: What multimodal data can tell us about learning?. *British Journal of Educational Technology*.
- Sharma, K., Niforatos, E., Giannakos, M., Kostakos, E. (2020). *Assessing Cognitive*

- Performance Using Physiological and Facial Features: Generalizing Across Contexts. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (forthcoming), 31 pages.
- Sharples, M., & Roschelle, J. (2010). Guest editorial: Special section on mobile and ubiquitous technologies for learning. *IEEE Transactions on Learning Technologies*, (1), 4-6.
- Spikol, D., Prieto, L. P., Rodríguez-Triana, M. J., Worsley, M., Ochoa, X., Cukurova, M., ... & Ringtved, U. L. (2017). Current and future multimodal learning analytics data challenges. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 518-519).
- Spikol, D., Ruffaldi, E., Dabisias, G., & Cukurova, M. (2018). Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *Journal of Computer Assisted Learning*, 34(4), 366-377.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4), 295-312.
- Vinciarelli, A., Pantic, M., & Bourlard, H. (2009). Social signal processing: survey of an emerging domain. *Image and vision computing*, 27(12), 1743-1759.
- Vujovic, M.,^[1]Hernández-Leo, D.,^[1]Tassani, S. and Spikol, D., (2020). Studying collaborative learning and space design with multimodal learning analytics. *British Journal of Educational Technology*.
- Worsley, M., Abrahamson, D., Blikstein, P., Grover, G., Schneider, B., & Tissenbaum, M. (2016). Situating multimodal learning analytics. In *Proceedings of International Conference of the Learning Sciences, ICLS 2016*, (pp. 1346-1349). ISLS.