Received: 6 July 2020 | Revised: 14 December 2020 | Accepted: 29 January 2021

DOI: 10.1111/jcal.12542

ARTICLE

Journal of Computer Assisted Learning WILEY

Challenges and opportunities of multimodal data in human learning: The computer science students' perspective

Katerina Mangaroska[1] | Roberto Martinez-Maldonado[2] | Boban Vesin[3] | Dragan Gašević[2]

[1]Department of Computer Science, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, Trondheim, Norway

[2]Faculty of Information Technologies, Monash University, Clayton, Victoria, Australia

[3]School of Business, University of South-Eastern Norway, Vestfold, Norway

Correspondence
Boban Vesin, School of Business, University of South-Eastern Norway, Raveien 215, Borre, Vestfold 3184, Norway.
Email: boban.vesin@usn.no

Funding information
Research Council of Norway, Grant/Award Number: 255129/H20

Abstract

Multimodal data have the potential to explore emerging learning practices that extend human cognitive capacities. A critical issue stretching in many multimodal learning analytics (MLA) systems and studies is the current focus aimed at supporting researchers to model learner behaviours, rather than directly supporting learners. Moreover, many MLA systems are designed and deployed without learners' involvement. We argue that in order to create MLA interfaces that directly support learning, we need to gain an expanded understanding of how multimodal data can support learners' authentic needs. We present a qualitative study in which 40 computer science students were tracked in an authentic learning activity using wearable and static sensors. Our findings outline learners' curated representations about multimodal data and the non-technical challenges in using these data in their learning practice. The paper discusses 10 dimensions that can serve as guidelines for researchers and designers to create effective and ethically aware student-facing MLA innovations.

KEYWORDS

ethics, higher education, human-centred analytics, multimodal learning analytics, pervasive surveillance, privacy, thematic analysis

1 | INTRODUCTION

There has been a growing interest in exploiting digital traces that learners leave behind while interacting with educational systems through artificial intelligence (AI), big data (Daniel, 2015), and learning analytics (Siemens, 2013) innovations. These innovations have been designed for multiple educational purposes such as adapting instruction, personalizing feedback, provoking reflection, or generating deeper understanding of learning processes (Lee et al., 2016; Viberg et al., 2018). While important achievements have been obtained in the last decades by mining data collected through online learning activities (e.g., clickstreams and keystrokes; Li et al., 2016; Mousavinasab et al., 2018), learning is ultimately a complex, multimodal process that involves linguistic, gestural, visual, and physical interaction of learners with educational systems, learning artefacts, learning space, peers, and educators (Kress, 2001; Oviatt et al., 2017; Ritella & Hakkarainen, 2012).

It has been proposed that emerging multimodal learning analytics (MLA) have the potential to enable the automated generation of models that account for the complexity of the learning processes with the purpose of providing real-time feedback or developing MLA interfaces (Blikstein, 2013; Blikstein & Worsley, 2016; Drachsler & Schneider, 2018; Mangaroska et al., 2019). For example, some MLA

[Corrections made on 08 March 2021, after first online publication: In-text citations for Figures 1a, 1b, and 2 have been corrected to Figures 1, 2, and 3, respectively, in this version.]

studies have focused on modelling student gaze to identify group synchrony as a proxy of collaboration effectiveness (Schneider, 2020) or to orchestrate basic behaviour regulation (Mangaroska et al., 2018); capturing physiological cues to investigate group regulation strategies (Noroozi et al., 2019) and individual achievement (Pijeira-Díaz et al., 2018); using computer vision systems to identify incorrect postures in healthcare training (Di Mitri, 2019); creating hand tracking algorithms to predict group work quality (Spikol et al., 2018); and using positioning trackers to identify teaching strategies in the physical classroom (Martinez-Maldonado, Schulte, et al., 2020).

Most of current MLA systems and studies are aimed at supporting researchers to model learner behaviours (see reviews by Crescenzi-Lanna, 2020; Di Mitri et al., 2018; Noroozi et al., 2019) or offer technical infrastructures to interconnect sensors and systems (Huertas Celdrán et al., 2020; Shankar et al., 2020, 2018). On the one hand, some researchers attribute this dearth of actual MLA interfaces for learners to the intrinsic complexity of multimodal data (Martinez-Maldonado, Echeverria, et al., 2020; Worsley et al., 2016). Modelling meaningful educational constructs (e.g., mind wandering [Bixler & D'Mello, 2016]) from commonly intertwined data-markers (e.g., heart rate, gaze, arousal) extracted from multiple streams of data is a difficult technical challenge (Blikstein & Worsley, 2016; Oviatt et al., 2018). Even if this modelling challenge gets solved, creating MLA interfaces that could be understood and used to inform actions of learners or educators is an additional human–computer interaction challenge to be addressed (Martinez-Maldonado, Echeverria, et al., 2020). On the other hand, most of the current learning analytics solutions are developed without learner's or educator's involvement in the design process (Buckingham Shum et al., 2019). This is why we commonly witness outcomes from sub-optimal learning analytics dashboard designs that have high potential to result in inadequate tools hard to be adopted (e.g., see reviews by Bodily et al., 2018; Matcha et al., 2019; Schwendimann et al., 2016). We argue that in order to create MLA interfaces that directly support learning, we need to gain an expanded understanding of how multimodal data can support learners' authentic needs.

This paper addresses the lack of research focused on understanding how learners perceive multimodal data and the learning context in which MLA interfaces can be used. In particular, we present perspectives of a computer science (CS) students, most of whom have already incorporated some similar technology (e.g., smart watch) for purposes different than learning into their everyday life. Thus, as emerging learning technologies bring challenges and opportunities for interaction and communication in numerous ways, empowering learners to become more conscious of the impact these technologies have on their learning practices becomes imperative. Moreover, technological innovations cause learners to become more demanding, concerned, and critical (Stephanidis et al., 2019), which is why learning analytics-based systems need to put more focus on human-centred design approaches for building trustful and beneficial relationships with learners and educators (Buckingham Shum et al., 2019).

We present a qualitative study with CS students who wore wearable sensors (e.g., wrist-mounted and head-mounted devices) while engaged in a problem-solving task (i.e., a software programming task). This study also used two sensors that were attached to the computer screen: an eye-tracker and a web camera. We conducted 40 semi-structured pairwise interviews to capture and analyse students' perceptions, attitudes, and expectations regarding the impact and usefulness of multimodal data in teaching and learning. CS students were asked to reflect on their first-hand experience in a multimodal learning activity setup and to envisage potential applications of MLA innovations. The interviews served to investigate the educational and ethical, rather than the technical challenges associated with multimodal data. This idea builds on George Siemens' hypothesis: 'The most significant challenges facing analytics in education are not technical. Concerns about data quality, sufficient scope of the data captured to reflect accurately the learning experience, privacy, and ethics of analytics are among the most significant concerns' (Siemens, 2013, p. 394). Consequently, our work addresses the following research question 'What are the challenges and opportunities of multimodal data in human learning from CS students' perspective?'. The contribution of the paper is two-fold: (1) it reports CS students' perspectives describing new and under-developed ideas about potential uses of multimodal data in educational context; and (2) it advances the discussion on the profound need of human-centred design approaches for educational technologies aimed to embrace the complexity of learning (Buckingham Shum et al., 2019).

2 | RELATED WORK

Our work focuses on the potential application of multimodal data in the learning context, the need for human-centred design approaches in educational technology, and the current open ethical and data privacy issues.

2.1 | Multimodal data and learning

Towards the end of the 20th century, educational research focused on language as a prime medium of communication, and consequently of learning and teaching; while gestures, actions, and images were considered to be an illustrative support (Kress, 2001). However, the multimodal nature of human learning (Jewitt, 2012; Wachsmuth et al., 2008) has led gestures, actions, and visual communication to evolve into articulated semiotic systems (Kress, 2001). This change caused language to become one of the several modes which are used for research and practice of learning and teaching. Multimodality in learning focuses on the multiplicity of modes of communication (e.g., text, image, speech, and haptics) that are active and observable when learners exchange information and create meaning (Kress, 2001). However, when there are hundreds of students in a classroom, not all modes of communication and learner interactions can be easily observable with the naked eye. Therefore, if we aim to gain further understanding how learners create meaning beyond language, we need to extend the current research capacities with new technologies (Blikstein & Worsley, 2016; Ochoa, 2017; Oviatt et al., 2018).

Interaction traces occurring across various modes of communication constitute what we refer to as multimodal data in this paper (Järvelä et al., 2019). Multimodal data capture learning aspects (e.g., mental effort, affective states) not easily observable with the naked eye or through self-reported data, and can be collected in unobtrusive (e.g., gaze data using eye-trackers) and non-invasive (e.g., brain activity data using electroencephalogram) ways using sensor technologies (e.g., eye-trackers) that monitor variations in different modalities (Lazar et al., 2017).

There has been a progress in MLA for research purposes. For example, Echeverria et al. (2019) used multimodal data (e.g., physiological and proximity data) to visualize collaboration aspects as a timeline of events to facilitate reflective learning in healthcare simulation activities. Durall et al. (2015) used EEG data to develop a reflective tool that learners can use to understand what habits and mental states impact their learning performance. Moreover, Hassib et al. (2017) developed a tool that uses EEG data for real-time monitoring of audience engagement that is fed back to learners for real-time reflection. McDuff et al. (2012) presented an interface that acts as a reflection tool for monitoring users' valence, arousal and engagement by combining audio, visual, physiological and contextual data. Finally, Ochoa et al. (2018) proposed an automatic feedback system on learners' presentation skills, by analysing multimodal data, that is, data from posture, gaze, speech and the presentation slides.

Existing studies in MLA show a successful development of digital learning technologies using multimodal data. However, it is still not clear what the main implications of these tools are for directly supporting learning, since (1) most of the tools have been tested in experimental settings with a small number of learners; (2) learners have not been consulted, neither before or after the design of such tools; and (3) making-sense and interpreting the feedback was quite a challenge or not tested. Thus, the work presented in this paper tackles these issues by interviewing learners after having a first-hand experience in a multimodal learning activity to comprehend their views and expectations regarding potential application of multimodal data in education.

2.2 | Human-centred learning analytics

Despite the many benefits envisioned that multimodal data can bring to education (Blikstein & Worsley, 2016; Järvelä et al., 2019; Oviatt et al., 2018), its effective use by learners and educators requires new competencies (e.g., new knowledge and data/visualization literacy; Reimann et al., 2015). Fundamental to these competences is learners' ability to interpret, comprehend, and generate inference from analytics derived from multimodal data (Mangaroska & Giannakos, 2018; Viberg et al., 2018). Moreover, learner involvement in the design and development of learning technologies through needs analysis, promises to build a sense of ownership that fosters positive user acceptance (Norman & Draper, 1986). Thus, the adoption and effectiveness of novel teaching and learning practices based on multimodal data depend as much on learner involvement as on the need for knowledge and skills.

Related research in learning analytics (LA) demonstrates a small number of studies documenting learner involvement in the design process (Buckingham Shum et al., 2019). In general, the role of learners has been limited to participation in usability evaluations after deployment of tools, such as learner involvement in evaluation of LA solutions (Corrin & De Barba, 2015; Lim et al., 2019) or examination of ethical issues and impact of an early warning dashboard (Sun et al., 2019). Less studies are exploratory, including examination of features learners expect from LA (Schumacher & Ifenthaler, 2018) or understanding the LA potential for learners' success (Knight et al., 2016). Exploratory research is instrumental in understanding the 'why' behind a system/product, and it is meant to collect descriptive information about motivations, expectations, perceptions, or actions that can improve the understanding of something we want to develop, while avoiding the risk of establishing weak concepts or missing to identify important factors. Finally, few studies have reported learner engagement in LA design processes (Chen & Zhu, 2019; Prieto-Alvarez et al., 2018; Rodríguez-Triana et al., 2018). The literature on MLA shows that the field is relatively young, so few studies have been undertaken from a human-centred design perspective (Echeverria Barzola, 2020). Consequently, this paper addresses this gap and builds on the emerging interest in generating deeper understanding of learners' needs and expectations about multimodal data in educational contexts. We consider learners as the major audience targeted to harvest the upcoming MLA benefits; thus, learners should be involved in the design process of new learning technologies to avoid sub-optimal outcomes that can result in inadequate tools used by no one.

2.3 | Ethics, data privacy and pervasive surveillance

Technological advances that made possible humans to learn everywhere and at any time across digital and physical settings, completely changed the educational landscape (Blikstein & Worsley, 2016). This change brought as many challenges as opportunities that continuously raise questions about the identity trade we humans do for the inclusion we aspire to Arora (2016). In this sense, the promise of learning analytics to understand and optimize learning comes along with challenges about stereotyped identities, indirect pressures to perform according artificially set indicators, possible discrimination of data subjects, emergence of system identity (i.e., persons as dynamic clouds of data), privacy and transparency, to name a few (Arora, 2016; Drachsler & Greller, 2016). These issues have become more sensitive with the emergence and collection of multimodal data (Drachsler & Greller, 2016).

Multimodal data can intensify learners' vulnerabilities in the era of pervasive surveillance (Prinsloo & Slade, 2016). Solove's (2005) taxonomy for privacy and vulnerability, should influence the LA research community to define principles and practices for multimodal data collection and processing to decrease the risks associated with misuse, lack of transparency, misinterpretation, and erosion of contextual integrity, that might result in stereotyping and discrimination (Henman, 2004; Prinsloo & Slade, 2016). Moreover, multimodal data

can further deteriorate power relations between learners and teachers (Slade & Prinsloo, 2013). For example, neurologists and neuroscientist have raised concerns about uses of brain recordings and potential novel forms of discrimination that can emerge from the pressures to expand sensory, cognitive and motor capacities (Wilson, 2002). In sum, as with every emergent research area, MLA also requires a community approach into proposing practices for privacy such as data collection, access, manipulation and processing, as well as development of an ethical framework that will regulate the impact of surveillance, power relations, and learners' identity as a temporal and context-bound construct (Slade & Prinsloo, 2013).

3 | METHODS

This section presents the design of the study, the methods used to collect and process the interview data, and the analysis approach to address the issues discussed in Section 2.

3.1 | Design of the study

We designed and implemented a problem-solving learning task to collect a fine-grained multimodal data set. The focus of the study was two-fold: (1) data-driven focus, that is, exploring what MLA can best describe cognitive, affective, and behavioural states of CS students; and (2) learner-driven focus, that is, understanding CS students' perceptions and expectations about the potential application and impact of multimodal data in a learning-context. Before conducting the study, ethics approval was granted from NSD—Norwegian Center for Research Data. The study consisted of three phases, set-up and calibration of sensor devices, a problem-solving task, and a pairwise interview, as shown in Figure 1. Four sensor devices were used during the experiment: a Tobii X3-120 eye-tracker placed on the bottom of the computer screen, a LogiTech web camera placed on the top of the computer screen, an Empatica E4 wrist-mounted sensor, and an ENOBIO head-mounted EEG cap with a 20-channel ENOBIO device. The duration of the study was 1 h for each participant plus 15 min for calibration and set-up of the sensor devices. The set-up of the study is depicted in Figure 2.

The focus of this paper is the interview phase which covers the learner-driven focus of the study. To examine participants understanding in relation to multimodal data and its implicated values (Friedman & Kahn Jr, 2003), we developed a semi-structured interview protocol. This consisted of a brief *retrospective* 'journey' of the overall experience, and a *prospective* evaluation of participants' expectations about multimodal data (Creswell & Creswell, 2017). The generated raw data from the sensor devices was not displayed to the participants during the task, nor were any type of analytics. However, each participant was introduced to the raw data during the calibration and set-up of the sensor devices, following a walk-through of the particular data streams (e.g., electrophysiological activity of the brain, gaze data, and physiological data such as heart rate). The used walk-through protocol is presented in Appendix A of this paper.

3.2 | Participants and procedure

During January 2019, we organized a study at a contrived computer lab at Norwegian University of Science and Technology (NTNU), Trondheim, Norway, with 40 students (8 females and 32 males), age

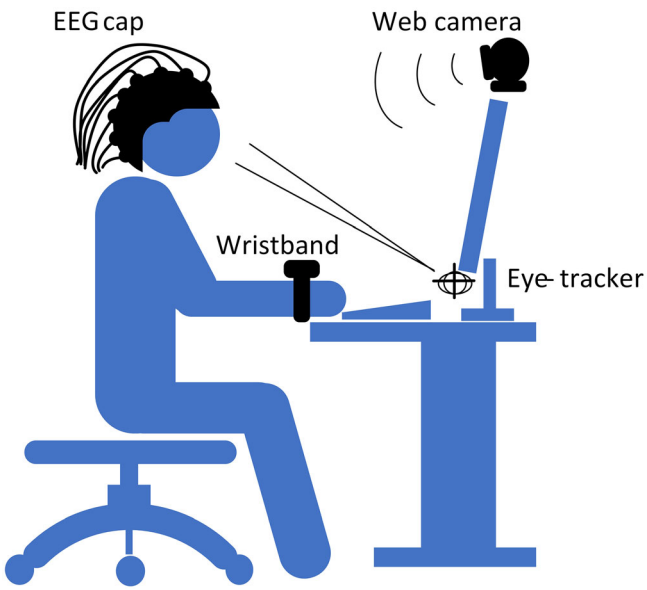

FIGURE 2 Study design: the study set-up [Correction made on 08 March 2021, after first online publication: Caption for Figure 2 has been corrected in this version] [Colour figure can be viewed at wileyonlinelibrary.com]

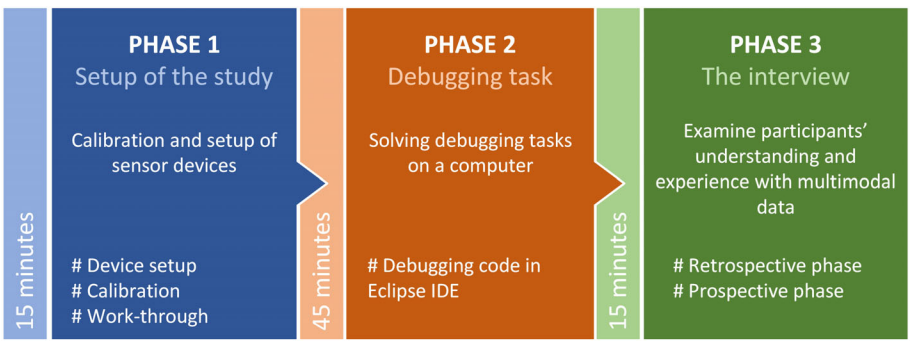

FIGURE 1 Study design: the three phases of the study [Correction made on 08 March 2021, after first online publication: Caption for Figure 1 has been corrected in this version] [Colour figure can be viewed at wileyonlinelibrary.com]

between 20 and 25 (M = 22.1, SD = 1.46), recruited via a mailing list. Students were recruited from all study years from the CS major. CS students are learners that have access to similar technology through science projects at the university, as well as opportunities to incorporate similar technology, such as smart watch, for other purposes than learning into their everyday life. Therefore, these are particular type of students that do not represent the average learner. All participants received a gift voucher for participating in the experiment. The study ran for a week—a total of 20 non-repeat sessions where each session had two students at a time in the computer lab, on two separate computers, for the entire study. The participants were instructed not to talk to each other during phase two, which was verified through the videos recorded with the cameras.

The study consisted of three phases (Figure 1). Phase one lasted for 15 min and included calibration and set-up of sensor devices, including a walk-through of the particular data streams (e.g., electrophysiological activity of the brain, hearth rate, electrodermal activity). Phase two covered the problem-solving task for which we allocated 45 min. The pairwise interview was the last phase with duration of 15 min, and was conducted in person right after the participants finished the task. Each participant was asked to respond to all questions from the interview protocol. Once a question was asked by the principal researcher, the two participants at the time, took turns to answer the same question. We used the same study set-up for the interviews for several reasons. First, pairs could facilitate more discussion at a similar level of interests with less cues from the researcher (Lewis, 1992; McLafferty, 2004). Second, when generating insights for designing learning technologies for collective use, discussing and sharing individual experiences could facilitate collective thinking (Simonsen & Robertson, 2012). Group interviews may have limitation of one participant dominating the interview, indicating potential threat to validity. To avoid this, the researcher was responsible for creating supportive, balanced, and non-threatening atmosphere (Basch, 1987). Finally, after debugging a code for 45 min, this set-up could trigger relaxed and enjoyable experience for our participants.

3.3 | Interview protocol

Table 1 shows the outline of the interview protocol, including the context from which the interview questions were generated. The interview questions frame of reference comes from the existing literature in LA and MLA (Azevedo, 2015; Beattie et al., 2014; Blikstein & Worsley, 2016; Järvelä et al., 2019; Pardo & Siemens, 2014; Slade & Prinsloo, 2013), and the forthcoming challenges of technological evolution addressed by Stephanidis et al. (2019). Finally, to support replication, scientific comparison, and alternative contextualisation, we developed the interview protocol along the lines of Greller and Drachsler (2012)'s LA framework.

TABLE 1 Overview of the interview protocol

Interview phase	Retrospective	Prospective
Guiding question	How was it?	How should it be?
Elicitation mode	Reflective observation	Abstract conceptualisation
Context	(1) Comfort of wearable sensors; (2) the learning activity (individual attitudes); (3) multimodal data (access, users)	(1) Learners' expectations in learning situations; (2) multimodal data and activity data; (3) potential limitations of multimodal data

We used the retrospective interview technique to empower participants to objectify their experience in a larger context of teaching and learning, underlying their needs and values as learners (Hassenzahl & Sandweg, 2004; Yue et al., 2014). We started by asking the participants how comfortable they felt with the wearable devices. Then, we asked them to reflect on their attitudes about multimodal data in the particular learning activity, to understand how they perceive what constitutes multimodal data. Finally, we asked them who should have access to the data, and whether and with whom they are willing to share their own data. The interview continued with questions asking participants to conceptualize learning situations where multimodal data can support them, either in collocated or distributed learning settings. We wanted to examine if any tensions exist between participants' expectations, and the potential challenges associated with multimodal data (Davis & Nathan, 2015). Next, we asked participants how a combination of multimodal data and log data (i.e., traces of their online behaviour) can be used to develop multimodal innovations (e.g., multimodal interfaces, cognition-aware systems). Finally, we wanted to know if participants saw any limitations regarding implementation or functionality of multimodal data.

3.4 | Data analysis

Interviews were audio-recorded, fully transcribed, and imported into NVivo 12. We applied an elicitation technique in a systematic manner and conducted a *thematic analysis* on the interview recordings following Braun and Clarke's (2006) six-step framework. Unlike many qualitative methodologies, this analysis is not tied to a particular epistemological or theoretical perspective (Maguire & Delahunt, 2017). Because the nature of the study was exploratory, so was the analysis; thus, we did not test any framework or impose existing scheme, but used an inductive approach and looked for themes that emerged from the text (Braun & Clarke, 2006). The interviews were divided between two researchers, who independently reviewed each transcript. After the individual coding process, the researchers met and settled any differences in the coded transcripts to achieve reliability and trustworthiness (Morse et al., 2002). We achieved high inter-rater reliability (Cohen's k = 0.74). In total, 174 coded nodes were generated, which led to 6 themes and 28 unique codes. Most coded nodes are associated with one theme, but some are associated with more than one theme. Table 2 displays the identified 28 unique codes.

In the second part of the analysis, we used a paper-based *affinity diagramming* (Beyer & Holtzblatt, 1997) on the coded nodes to arrange the insights about multimodal data. The initial 174 coded nodes were clustered in 28 unique codes, which represent the most

TABLE 2 The 28 unique codes under the six themes

Theme 1	Theme 2	Theme 3
Level of comfort wearing sensor	Type of sensors and data	Wrong conduct and data-profiting
Flexibility when wearing sensors	Increase motivation and engagement	Unequal power relations
How wearables affect learning	Capturing pitfalls, misconceptions, and tracking performance	Access to data—what, when, who
	Data literacy skills	Regulations for storing/using data
	Incentives to use MLA	Constantly and pervasively surveilled
Theme 4	**Theme 5**	**Theme 6**
Reflection and monitoring of progress	Reasons to use MLA in education	Merging learning technologies
Cognition, affect, and behaviour of individuals	Users and uses of MLA	Adaptation to individuality
Identify areas for improvement	Plan, design, and adapt curricula, assignments and exams	Organize and optimize considering cognition, affect and behaviour
Support learning not just passing a course/ getting a good grade	Effect of individual differences on learning	Personalized smart apps
Personalize feedback, instruction and scaffolds	Set the right feedback loop	External factors that affectlearning

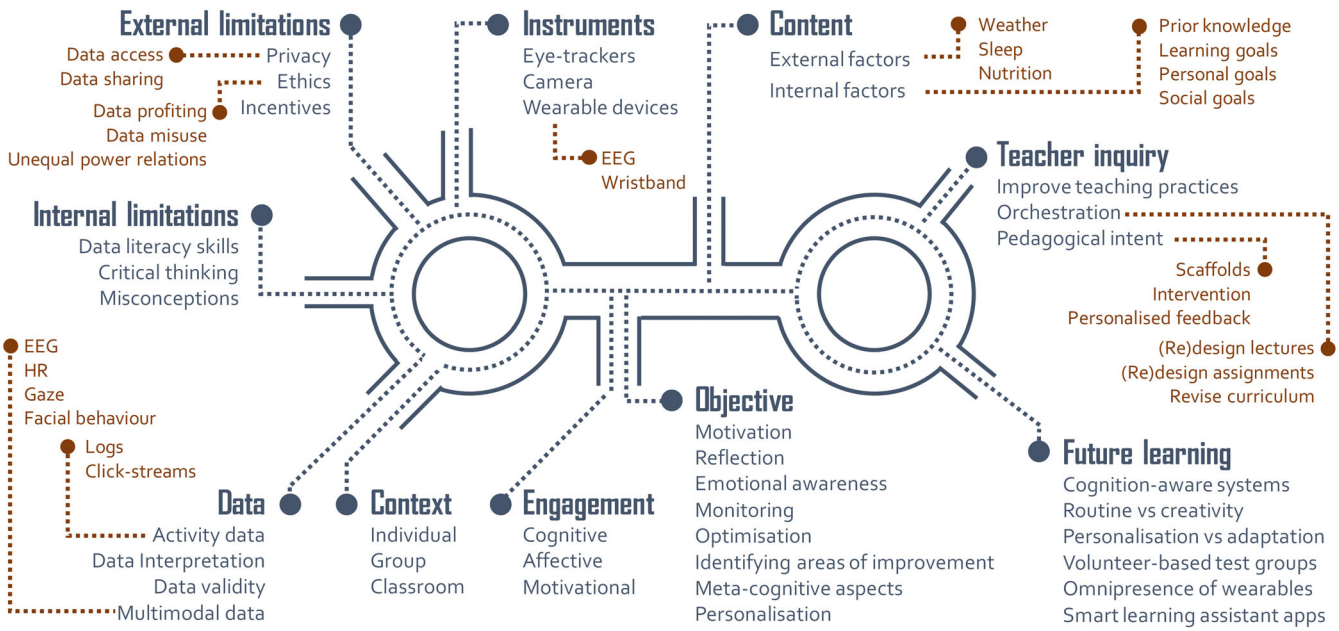

FIGURE 3 Visualization of the CS students' ideas elaborated throughout the six themes and framed in 10 dimensions [Correction made on 08 March 2021, after first online publication: Figure 3 was omitted and has been reinstated in this version.]

important concepts in the data set. The same 174 coded nodes were written on paper notes and clustered according to the similarities they shared on a topic. Affinity diagramming was applied to systematize and group learners' perceptions and expectations under practical design implications, discussed through 10 dimensions (shown in Figure 3) that extend the current MLA design knowledge.

4 | RESULTS

The first three themes originate from learners' reflections on their overall experience, while the second set of three themes originate from learners' vision about potential applications of multimodal data in education. At the end of this section, we present a visualization that illustrates how CS students' ideas can be framed in 10 dimensions that were explored throughout the six themes.

4.1 | Theme 1: Degree of intrusiveness of wearable technology

When it comes to *wearable devices*, the state of physical ease and freedom from unpleasant sensations or constraints, were reported by all 40 participants as important features that need to be considered in

educational research and practice. Following this, none of the participants expressed discomfort with the *wristband sensor*.

However, the majority of the participants (30 in total) expressed discomfort wearing the EEG cap. Their experience ranged from feeling slightly distracted (–P23: 'At the start was OK, but later on, when you start to move your head you kind of begin to feel all those tiny needles'), to irritation (–P12: 'It was kind of annoying, I started to feel like scratching my back head, but then I thought "I cannot move, it will get sloppy data", so I tried not to think about it'), headache, and even pain for some of the participants (–P01: 'It was painful, I could not stop thinking about it! I am not used to be in pain while coding, so I wanted to complete the assignment as fast as possible'). From those 10 participants that reported wearing the EEG cap to cause no unpleasant sensations, seven reported that they would not mind wearing it once a week for an hour to generate data if it is required from the instructor, while three reported that the ear electrode was the only part from the *EEG equipment* that caused them discomfort.

The lack of freedom and flexibility to move during the computer-mediated learning activity was a big issue (–P15: 'The only thing that irritated me was not moving when I'm coding. This affected my attention a bit'). The participants expressed that the rigid feeling they experienced when coding is far from their natural setting that they are used to when learning programming, and this change is annoying. However, beyond the intrusion of the learning space CS students are used to, some participants (10 in total) expressed willingness to wear the equipment if they can help instructors to design better assignments or a study curriculum.

4.2 | Theme 2: Opportunities and challenges for computing education

All participants were familiarized with and understood the physiological data, as most of them own a smartwatch and monitor their heart rate during physical activities. However, most of the participants had never engaged with gaze data, nor had they been familiar with *EEG data* (i.e., the brain wave frequencies), with the exception of three master students who had experience working on Virtual Reality projects. Adding to this, some participants expressed that researchers use EEG the same way as *eye tracking devices*, for example, –P10: 'capturing and measuring attention, but instead of observing your eyes, the researcher observes your brain activity'.

In general, almost all participants expressed positive attitudes towards the opportunities multimodal data might enable for computing education –P03: 'It would be useful if there is some way of categorising the types of debugging problems which cause more frustration and especially the type of problems that actually cause you to give up'; –P39: 'It is a good idea if you could use it for self-diagnosis when you are struggling and you cannot understand why you are not able to solve the task'; –P11: '*Tracking your errors*, *performance*, and skills development in different programming environments, and *reflecting* on the same patterns along those environments can help you improve and become more efficient'. Another interesting suggestion was using multimodal data for *motivational support*. For example –P09 expressed this as follows: 'If I feel like I am gonna give up, I would like to see my data, like how far I can go before I give up, so then maybe I will find a way to push myself and not give up quickly due *frustration or confusion*'.

Interestingly, awareness and recognition of *cognitive and affective states* during debugging were highlighted by learners as important constructs that multimodal data can help educators with, to design *interventions* and improvements that can increase *engagement* in computer-mediated learning activities. For example, P01 explained: 'If you can collect data on how I debug and you can observe when I have entered the unproductive phase, then you can give me a hint if I should take a break or change the task'. –P03 also mentioned the following: 'I think that by measuring the *[cognitive] load*, you can observe what is overloading your brain and then find ways to use it as an input for feedback. It is easy to be overloaded or feel lost if you are thinking in a single way but with the *right feedback* you can change the way you work through the problems and that can really help'.

From all data streams that we collected with the various sensor devices, *gaze data* were recognized by almost all participants, as the most applicable data that could help students to learn programming (–P37: 'Gaze data can show me if I look 60% of the time into the wrong parts of the code which might be a reason for my under-performance'). Gaze data have been used in the literature to observe trajectories when students write a code or debug a code and what transitions mark a successful behaviour (–P12: 'I think eye tracking can clearly show how we are thinking based on how we are approaching the problems and from where on the screen to where our gaze is moving, which we may not consider it when we are doing it ourselves'). However, the use of gaze data to generate and present automatic and actionable feedback in real time is still an open research challenge. Related to this opportunity, –P01 requested the following: 'I want automatic *[personalised] feedback* from tracking my eyes, such as hints on where I haven't looked, because you usually get stuck in looking at the same things making the same patterns, and this is difficult to notice it on your own'.

Fifteen participants pointed out to the usage of physiological data, in particular, monitoring stress from *heart rate*, as a possible solution that could help them alleviate frustration by 'escaping' earlier when they were stuck in coding without progress. Participants also highlighted the possible usage of brain activity data, but only for *classroom lectures* (e.g., –P18: 'In my opinion, these caps make sense for the whole class to put it on, and the teachers can see what kind of lectures are the best for the whole class. This might help teachers to do the most efficient kind of teaching. I am not sure if using it to personalise learning is the way to go'). Importantly, none of the participants mentioned any possibilities they could anticipate from using the *facial expressions data* from the videos.

Finally, few participants expressed some concerns regarding multimodal data. Five of the participants raised a concern that multimodal data collection is time consuming and expensive, and instructors might lack *incentives* to see the value of using it (e.g., –P22: 'Teachers

won't earn anything from this; this type of data cost extra time and money, and universities will need extra teachers and extra research on how to use it'). Moreover, one concern that was raised by some of the participants (seven in total) regarding the multimodal data streams is the applicability of the average, that is, 'for most of the students', as we all approach to problem solving differently, so what works for some might not work for others. At the end, some participants (eight in total) expressed doubts regarding *data validity* (e.g., –P30: 'If you move, the data will get blurred and it won't have the value you want to have, so the analytics will be biased; students might be difficult to deal with if they have to learn in restrained conditions').

4.3 | Theme 3: Ethics, data privacy and pervasive surveillance

This theme elicited from the participants' strong views on who should have *access* and who should use their data. Most of the participants (28 in total) conveyed positive statements towards sharing their data with educators, only if it is anonymous and aggregated. Moreover, some participants (15 in total) expressed their concerns regarding potential power issues (–P17: 'I feel if my teacher knows too much, he has the power to use it against me'), and diminished agency (–P02: 'If the teacher has all the data for every student individual, at every moment, it would be quite easy to manipulate people with that information'). This suggests that learners perceived multimodal data to be a potential source of *unequal power relations* between educators and learners.

Other participants (10 in total) were not comfortable to share their data at all (e.g., –P19: 'I wouldn't like the professor to know that I have done nothing for his assignment. It makes me feel uncomfortable'). For example, P36 explicitly expressed concerns regarding the kind of *data that can be captured* via *wearable sensors compared to clickstreams*, as follows: 'I feel like it is a sort of a violation someone to have all of my data. It is very invasive. I know I cannot stop Google to gather my data, but for sure I will have a say whether the school can have my biodata'. Thus, some participants pointed out that they would prefer to manage and curate their own data (–P45: 'If I have the tools that collect and analyse the data, I would collect everything myself, and make sure that *everything I do is optimal*'). One participant even suggested a volunteer-based test group consisting of students who like to *share their data*, so educators can put in a group to test the planned course assignments and exams (–P25: 'I would say that it would be nice to have a control group of volunteers that actually want to do this and look through the exercises before the beginning of a semester, so that the professor can identify and fix the problems instead of asking for my data'). Almost all participants indicated that *individual physiological data* should be used only by the individual from whom the data were collected for *monitoring their own emotions, performance, and progress*. This suggests that each data stream needs to be treated differently from a privacy point of view, since some can point at constructs that may be unrelated to the learning task at hand.

Strong arguments were raised by several participants (eight in total) regarding potential *data-profiting* given the amount of fine-grained physiological markers being pervasively captured via sensors. For example, P38 stated the following: 'The bad thing about selling your data is that the more people have your data the more you are exposed and unsafe. And this data you collect is very personal'. Related to this, P36 said: 'If someone is profiting off, I am not okay with it. But if they do it, I would rather them doing it safely'. Ten participants made an interesting note regarding possible *misuse of data*, as well as being able to have a wider overview of the learning experience (e.g., –P13: 'As long as the professor has the data from everyone in the course, and not just from individuals who can have a bad day or be amenable to horrible weather conditions or lectures early in the morning, can use it as aggregated data to make changes in the course; otherwise, he will probably use non-representative data to make changes').

Finally, a small number of participants (five in total) expressed their concern about being 'watched' (e.g., –P01: 'I am getting a little bit paranoid thoughts about it, because there is a lot of personal data that can be tracked and misused' and –P02: 'I don't like someone to know absolutely everything about me. I am not comfortable being watched. I want my own privacy'. These statements revealed a very important ethical implication about the assumption that all data captured from learners is relevant for learning, which is not true for multimodal data.

4.4 | Theme 4: Learning how to improve your learning using analytics

Participants communicated a wide range of ideas how analytics from multimodal data can assist them to improve the way they learn; for example, by *identifying areas of improvement* (–P07: 'I think having analytics can help me improve the way I learn because I can use it as a troubleshooting process'); by focusing on *meta-cognitive aspects* (–P22: 'These data might help students to focus on learning not just passing the subjects'), and by helping them to *reflect* on their learning strategies (–P37 stated: 'If [analytics] it contributes to realise what you are doing wrong or what you are doing right, maybe you can change the way you learn and apply it to other subjects'). Moreover, learners identified potential solutions based on multimodal data for detecting mind wandering (e.g., –P14: 'I like to know where my mind wanders, like when my eyes are not really focusing on the letters or if I am looking somewhere else, and how to get my attention back on the task at hand'), and data triangulation (e.g., –P30: 'Analytics can be a combination of all these data streams, so that teachers can avoid fake effects or assumptions they have from exams').

It is no surprise that the learners (22 in total) were *aware of their emotions* when learning, and quite often they pointed out that multimodal data could help them manage and control their own emotions (e.g., –P05: 'If I know what is stressing me up when learning (e.g., lack of sleep, poor organisation, bad planning) I can turn around the way I

think and approach learning'), or stress levels (e.g., –P15: 'A stress development over a semester could be useful to observe my spikes—whether they are towards the end of all courses or my work load was balanced over the semester. Then, I can base the rhythm in which I study the best for the next semester'). Other learners believed that analytics could help learners and their instructors monitor and manage emotions, so that learners could *stay engaged* (e.g., –P26: 'Most of the time I don't realise when I am irritated and I suddenly give up. Maybe my phone can send me notifications to cool down').

However, some participants (11 in total) expressed worries about the pressure that comes with being able to know everything about their performance, but not being able to *interpret the collected data*. For example, P06 suggested that she would need scaffolding to interpret her data: 'I don't know what to do with those numbers. What if I am confused?! Someone needs to tell me what to do when I am confused, not just that I am confused'. Moreover, these students also were concerned about how they could use the analytics to take an action and proceed further (e.g., –P29: 'It is more about politics right now. Schools support the ones who are not that great, and they don't care helping better students to become even better'). All these statements showed that learners find analytics from multimodal data to be a valid indicator to inform reflection, awareness, and monitoring progress, but data-driven actions via multimodal data were not identified to lead to straightforward benefits.

4.5 | Theme 5: Academic improvement actions

The majority of the participants (38 in total) envisaged that a key use of multimodal data is for educators to *improve their teaching practices* (e.g., lecturing, designing assignments, planning the course, and designing the classroom lectures); in other words, to improve the orchestration of learning activities. Hence, this theme relates to the educator's use of multimodal data as a support tool (1) to assist learners to improve their learning practices; (2) to identify content and course problems for *design refinements*; and (3) to reflect on their own teaching choices and fine-tune them accordingly. Data from *eye-tracking*, for example, can help educators to understand where students are looking when working on assignments and how much time they spend on particular concepts (e.g., –P11: 'Eye-tracking can be beneficial because it can help teachers to study which parts of the curriculum are troubling and use that data to design lectures that focus on those aspects'), in order to implement data-driven changes and *optimize the level of difficulty*, or design assignments in a different way.

Some participants (12 in total) suggested that educators could use the insights from learner *engagement metrics to improve teaching* (e.g., –P03: 'Teachers can categorise the types of problems which many students find complex and difficult to understand, and use this data to better scale teaching concepts, avoiding potential gaps in knowledge or leading students to easily give up'); while others highlighted the *effect of individual differences* on performance (e.g., –P17: 'The teacher should use my data when I am struggling; to observe the way I am trying to find a solution, because trying to solve a problem is personal and depends on experience, and is the most difficult part to teach. It is highly contextual and individualized'). The 'one-size-fits-all' teaching approach is not alien to learners, nor are learners completely detached from the idea about personalisation in a learning context.

Next, eight participants proposed educators to use multimodal data as a *motivation tool* (e.g., –P09: 'When I don't know how to solve a problem, sometimes I loose interest and I give up. Maybe the teacher can motivate me to push myself further'). Moreover, the majority of participants (29 in total) envisaged that multimodal data can best help educators *inside the classroom*, to map the general awareness of the audience, especially when there are hundreds of students attending a lecture (e.g., –P10: 'The teacher can see how people respond to the explanations he/she gives in class'; and –P13: 'Data from the brain activity can tell the teacher when students are tired or how concentrated they are on particular parts of the lecture'). Participants also shared an interesting observation about lecture attendance, as expressed by P09: 'although lecture halls are full, half of the students are usually sleeping, and the other half is not receptive because they are either bored or their mind wanders'. Hence, participants suggested educators to use multimodal data to explore attention. For example, P20 explained: 'when students' brain is active or not so that the teacher is aware what kind of lectures are good for morning and what are better for the afternoon classes'. P38 also explained the following: 'If every student is phasing out after 30 minutes, then the teacher should try to get the attention with something else, maybe a quiz or switch to an easy topic'. One student (P33) even associated the potential use of multimodal data with her teaching experience: 'When I was a student assistant, I wished I could have a button, so that students could click when they understand or not, and when numbers spike I might be able to do something about it in real-time'. These ideas can be linked back to theme 2—that data from EEG device make most sense to be applied for classroom activities, such as lectures; although it is contradictory with the statements in theme 1, about the level of intrusiveness of an EEG cap. This also suggests that some participants might have overlooked to understand the potential application of EEG for a single individual.

4.6 | Theme 6: Re-thinking learning

Radical ideas and change require rethinking of goals and strategies. Thus, this theme depicts the envisioned future where technology aligns with human cognitive architecture (i.e., our limited working memory that deals with conscious activities and the unlimited long-term memory that stores our knowledge) to optimize the quality and quantity of knowledge transfer and retention. A very common answer among participants was identification of *unproductive phases* and frustration in real time; thus, getting alert notifications on your phone to take a break or engage in a different activity. Not so

common idea was that of a system based on physiological data metrics that notices when a learner loses attention, and triggers a pop-up message to get the learner's attention back by offering a *personalized feedback*.

Some participants (18 in total) pointed out to a failure in current learning tools/systems to consider the influence of *external factors* on student learning, such as light, temperature, weather, sleep and nutrition. For example, –P37 stated the following: 'when I go to a lecture, it is more about my day-to-day things rather than just a very specific information presented on that lecture that affect my concentration. So, my pulse and my brain activity is affected by private issues or lack of sleep, which is more important in my opinion'. Similarly, P33 explained: 'Sleep and social skills are just a few of the things I am struggling with. I would like data to tell me what I am bad at, like scheduling, organising, eating healthy, or lack of sleep'. Learners' expectations for acknowledging the influence of external factors on their well-being and learning, suggest a new opportunity for MLA, one that can lead towards development of more comprehensive learner models.

Fifteen participants let their imagination and needs to reach a whole new level, by envisioning development of *cognition-aware systems*. These systems should follow learners' physiological data, learning progress, and habits, and align all of that with their *goals and needs*, in order to support them in organizing and optimizing their day-to-day activities (e.g., –P24: 'It's like a *guided assistant*; it knows your workflow and it can suggest new things to do, new methods to try, and you can experience more based on your alertness', and –P38: 'Look what data say about you—if you are better spending time on homework or going to the gym. After all, it is as much about the quality of the work as it is about your own happiness'). Finally, five participants raised a concern that cognition-aware systems might promote *routine and reduce creativity* (e.g., –P24: 'Envisioning future with multimodal data might lead to more routines in life, and I won't use it to create a routine for myself', and –P02: 'You will know everything and you might achieve a more fulfilling life because everything will fit perfectly, but I think we will get bored very fast'). Routine and reduced creativity is something that has been barely discussed in MLA; thus, hearing it from learners suggests that MLA community needs to consider it when designing emerging learning technologies.

4.7 | Emerging dimensions

CS students' perceptions and expectations have been summarized in 10 dimensions (please see Figure 3), that were identified using affinity diagram technique on the coded nodes from the interview data. During the thematic analysis, we used open coding (we did not have any pre-set codes) and generated 174 coded nodes. These initial 174 coded nodes were later clustered into 28 unique codes (shown in Table 2), which represent the most important concepts in the data set. The 28 unique codes (which contain all 174 initial coded nodes) helped us to identify the six themes, that is, the patterns that express what is the most significant about the data. Considering the six themes, the unique codes, and the frequency of the coded nodes in the interview data, we visually present 10 dimensions (shown in Figure 3) that can support researchers to examine the aspects which CS students deem as important for designing human-centered MLA tools, interfaces, systems and methods. The mapping between the six themes and the 10 dimensions is presented in Appendix B. Most of the unique codes are associated with one dimension, although some are associated with two dimensions.

These dimensions can act as a springboard that can lead the design of future learning technologies to undertake a human-centred design perspective. The 10 dimensions are: (1) *objective*—reasons why we need to design a specific learning technology; (2) *context*—whether the learning technology is intended for individuals, group, or as a classroom tool, in a physical or digital settings; (3) *Instruments* what tools and sensor devices are appropriate for: individual activities, collaborative tasks, or lectures; (4) *content*—what factors affect learning, performance, and outcomes that we should account for; (5) *engagement*—what causes a learner to be engaged or disengaged in a learning task; (6) *internal limitations*—knowledge and skills that learners and educators have or lack; (7) *teacher inquiry* educators' competences to use data to improve teaching, orchestration, and pedagogical intent; (8) *external limitations*—ethical and data privacy issues; (9) *data*—type of data to be collected, data quality and relevance, access to and usage of data; and (10) *future learning*—latent factors and learners' expectations for emerging learning technologies. These 10 dimensions should welcome more productive reflections and actions when exploring learners' needs and expectations in the design process, and serve as guidelines for researchers and designers to create effective and ethically aware student-facing MLA innovations.

5 | DISCUSSION

Our work highlights the profound need for a human-centred perspective in the design of learning technologies, because technology exists to support human skills and ingenuity, and not the other way around (Gill, 2012). On the one hand, the motivation for our work emerged from the increased interest in multimodal data as a powerful source of real-time information that links cognitive, affective, motivational and metacognitive states of learners (Azevedo, 2015), and as such, have a potential to augment human cognition (Schmidt et al., 2011). On the other hand, new technologies bring more opportunities and challenges, driving learners to be more demanding, concerned and critical, which is why learning analytics researchers need human-centred design approaches to build trustful and beneficial relationships (Stephanidis et al., 2019). Building the discussion around Figure 3 aims to emphasize the connections between the dimensions, which are represented as if they are all joining the streets of a city, establishing relevant pointers for what is appropriate and ethically necessary when designing learning technologies, rather than efficient and profitable, no matter how

attractive the economic argument might be. These dimensions represent a driving force of the creative thought itself, with a strong emphasis on interaction, communication, and meaning, to model human natural behaviour and communication within contexts, so that we design MLA tools and interfaces that are more intuitive (physically, perceptually, cognitively and emotionally) and freer of technology-induced distractions (Oviatt, 2006). Moreover, these dimensions also control how human values will be implicated in the design of learning technologies, to create learning conditions that support psychological well-being, autonomy, learner identity, diversity and universal usability (Friedman & Kahn Jr, 2003).

It is not surprising that CS students understand the holistic and complex nature of learning, and seek for acknowledgement of factors that directly and indirectly affect their well-being, progress, performance and learning. As shown in Figure 3, *internal factors* of the *content dimension* represent students' prior knowledge and their kaleidoscope of goals, such as learning goals, personal goals, social goals and so on. According to Roschelle (1997), 'learning proceeds primarily from prior knowledge, and only secondarily from the presented materials' (p. 1). Therefore, neglecting prior knowledge can cause students to lose interest (Tobias, 1994), develop misconceptions (Mladenovic et al., 2016), or learn something opposed to the educator's intentions (Roschelle, 1997). Moreover, the criticism of past research on goal-directed behaviour in educational psychology (Boekaerts et al., 2006) raised the awareness that achievement goals are only a fraction of the goals students bring and seek in the learning settings. Thus, examining the patterns that students have established between achievement and non-achievement goals, can increase the understanding of what goals give meaning, purpose, and direction to students' actions in diverse learning settings (Boekaerts et al., 2006).

CS students also emphasized that *external factors* such as weather, lack of sleep, or poor nutrition, are some of the factors that affect their learning and performance by influencing their well-being. For a moment, taking into account external factors might seem unimportant in a learning context, but living in a world where technology is omnipresent, students feel empowered to demand new technologies that can ensure and enhance the human well-being. Human well-being is one of the seven forthcoming human–computer interaction Grand Challenges for living and interacting in technology-augmented environments (Stephanidis et al., 2019); thus, learners are entitled to ask for learning systems (Bosch et al., 2015) that are beneficial to humans, endorsing their values and expectations, and facilitating their well-being. Therefore, during the interviews, many CS students shared their expectations for *smart learning assistant applications* and *cognition-aware systems* of the *future learning dimension*, that can assist them to organize and align their day-to-day-activities with their cognitive capacities and emotional states. CS students believe that research should consider the connection between the increased availability of affordable and accurate *wearable sensor devices* with the need for *personalized and adaptive* learning technologies, to further develop our educational systems by improving quality and outcomes, as well as eliminating structural barriers and inequalities (Banathy, 1991). However, our participants expressed concerns that future learning technologies, which are heavily dependent on *multimodal data*, might bring *routine* and diminish the *creativity* in finding solutions on their own, particularly when an individual does know everything about oneself at every moment.

The need for recognition and inclusion of external factors in future learning technologies plays an important role in the connection between the content and the instrument dimensions. The *instrument dimension* represents different types of tools that can measure cognitive, behavioural, and emotional aspects of individuals (such as attention, stress, cognitive load, frustration, boredom, and confusion) caused by a combination of factors, including internal and external factors as presented in the content dimension. Moreover, a very important consideration when it comes to an instrument selection is the requirement for avoiding unpleasant sensations, as reported by the students in the interviews. This is aligned with the prospective usefulness of eye-trackers and wrist-mounted sensors in educational technology research (Calvo & D'Mello, 2010; Was et al., 2016). This also suggests the need for more research focused on affective wearables (Picard & Healey, 1997) that are flexible, easy to manipulate, and painless (Bonato, 2003; Olguín Olguín, 2011; Poh et al., 2010). Furthermore, while different methods/tools (e.g., think-alouds, pre-tests/post-tests, and self-reported questionnaires) have been practised in the past to measure cognition, affect, motivation and metacognition (Azevedo, 2015), multimodal analytics researchers are demonstrating that sensor devices can measure in real time, more accurately and in an objective way, various cognitive, behavioural and emotional states of learners (Blikstein & Worsley, 2016; Ochoa, 2017; Oviatt et al., 2018).

As learning is a complex process (Van Merriënboer & Kirschner, 2017) and a process that 'emerges' (Jacobson et al., 2016); we need sophisticated instruments to generate more nuanced understanding of its complexities. At present, researchers have been utilizing *eye-trackers*, *cameras*, *and wearable devices* presented in the instrument dimension to generate *multimodal data* highlighted in the *data dimension*. Using these devices, researchers can collect data such as *heart rate*, *gaze*, *electrophysiological activity of the brain*, *or facial expressions data*, to study and model learning strategies (Mangaroska et al., 2018; Worsley & Blikstein, 2015), to predict high-level constructs such as learner attention and engagement (Chan et al., 2020), to design multimodal learning interfaces (Echeverria et al., 2019), or to generate insights about teaching at a more fine-grained levels (Martinez-Maldonado, Echeverria, et al., 2020; Martinez-Maldonado, Schulte, et al., 2020; Prieto et al., 2018).

The connection between the data, the context, and the instrument dimension emphasizes that choosing the instruments to collect data from learners in various learning settings (e.g., physical and digital environments) requires a careful consideration of the *context dimension*, the degree of intrusiveness and invasiveness of the selected instruments, and the data dimension. Such connection increases the understanding of the collected data, in particular, the *interpretation of the data*, as well as the *data validity* presented in the

data dimension. Echeverria et al.'s (2019) study is an example of a pervasive learning activity where *activity data*, such as logs, heart rate and proximity data, can support nursing students to reflect on their collaborative learning experiences more holistically, on *individual* and *group* levels, augmenting their perception and cognition. In our study, most of our participants recognized the potential applications of *EEG data* at a *classroom level*, amidst the discomfort it causes, and contrary to the evidence in the literature which verifies the value of *EEG data* in measuring cognitive functions and dimensions of learning at an *individual level* (Antonenko et al., 2010; Galán & Beal, 2012; Klimesch, 1999). Moreover, although *facial expressions* were not mentioned by the participants, possibly due to the risks of pervasive surveillance (Ogan, 2019), *wristband sensors* and *cameras* from the instrument dimension, might become an unobtrusive and a powerful instrumentation that researchers can easily scale up at a *classroom level* (McDuff et al., 2012; Ogan, 2019). However, for now, a good alternative towards which the MLA community can focus, as it might be more acceptable and decrease concerns of pervasive surveillance, is the approach proposed by Stone et al. (2019), in which they use natural language processing to provide teachers with meaningful automated feedback about the quality of their classroom discourse.

The connection between the data dimension and the *engagement dimension* opens many opportunities for researchers and educators to distinguish between *cognitive*, *affective*, and *motivational engagement*, and understand learning in a variety of contexts. At present, engagement is one of the most misused and overgeneralised constructs in educational research (Azevedo, 2015), which can benefit from examination on a grain-sized continuum utilizing *multimodal data*, from an individual in the moment to a group of learners in a class (Sinatra et al., 2015). To that end, the participants in our study highlighted their expectations for a change in perspective among educators, on how teachers address progress and engagement in learning. Students, in our study, believed that their learning progress could and should be explained as a combination of mental, behavioural and motivational aspects, rather than inferred from a unimodal perspective of performance metrics. The cognitive-affective significance in learning prompted our participants to require educators to differentiate between cognitive and affective states, because being frustrated or cognitively challenged are two different things (D'Mello & Graesser, 2011, 2012). Consequently, advances in modelling and measuring various dimensions of engagement (e.g., *affective*, *cognitive*, *motivational*) can lead to instructional and design recommendations for learning environments, tools, and interfaces that effectively engage students (Azevedo, 2015).

Different dimensions of engagement can play an important role in the connection between the data dimension and the *teacher inquiry dimension*. As shown in Figure 3, CS students believe that *multimodal data* can improve *teaching practices*, optimize *orchestration*, and lead to learning design recommendations for *pedagogical intent*. Prior research has already shown the value of *multimodal data* for educators' role in orchestrating (e.g., designing, managing, adapting) learning activities at multiple social levels (Aslan et al., 2019; Dillenbourg, 2013; Prieto et al., 2018). In our study, the participants expressed that EEG data can assist educators in instructional decision-making, as educators can gain insights into how learners understand the lecture material, how concentrated they are, what parts of the lecture are difficult or boring, so that they can make changes or interventions accordingly (e.g., *(re)design lectures*), and even find the optimal engagement level that fits the audience as well as the instructor (Hassib et al., 2017). The participants also expressed their positive attitudes regarding gaze data in *(re)designing assignments* or optimizing the instructional *scaffolding*, as shown in the teacher inquiry dimension. In fact, today's technological advances can support for shifts of interaction and attention accommodating a wide variety of users. However, we need to scale up and transform the existing methods to comprehend and address the evolving human needs, without being steered by the current technological capabilities (Stephanidis et al., 2019).

Fostering and supporting healthy behaviours with the help of technology advocates for positive human development and learning progress that may prevent, reduce, and manage stress (Coventry, 2012). Therefore, a careful consideration behind the reasons why we need to design a specific learning technology should produce outcomes that surpass the ever-increasing availability of data about students in learning environments. In our study, CS students expressed several ideas presented in the *objective dimension*, ranging from increase in *motivation* and *emotional awareness*, *identification of areas of improvement*, *monitoring and reflection*, to detecting unproductive phases for *optimisation* and *personalisation*. The objective dimension in this paper touches upon the gap in the current literature why researchers and designers do not frequently engage with students in the design process (Buckingham Shum et al., 2019). In general, the role of students has been limited to participation in usability evaluations after deployment of tools (Corrin & De Barba, 2015; Lim et al., 2019; Sun et al., 2019). On the one hand, researchers and designers cannot consider students to be experts about the possibilities of emerging technologies. While this argument carries some weight, students certainly have expertise in certain other aspects, such as their own learning experiences, the challenges they face, and their preferences. This opposing argument acknowledges that students are not experts in pedagogy, nor have formal education training, but their voices regarding their needs and preferences can be valuable assets for designing effective learning tools, interfaces, and technologies that will ultimately be used by them or for their own benefit.

Finally, the connection between the internal and the *external dimensions* with the data dimension, emphasizes the importance of ethics and privacy. These are two themes that have always been significant in all technological domains (Stephanidis et al., 2019), and discussed for a long time in the learning analytics community (Drachsler & Greller, 2016; Slade & Prinsloo, 2013). While privacy relates to the 'right to freedom from surveillance or unauthorized disclosure of one's personal information' (Corrin et al., 2019, p. 10), ethics refers to social and cultural conventions, doing the right thing with data considering human values. As shown in Figure 3,

privacy in the external limitations dimension refers to *data access* and *data sharing*. CS students in our study, expressed concerns and lack of willingness to give access and share data from their biomarkers.

In comparison, Tsai et al. (2020) found that students trust universities in handling their data, although they are aware that they cannot prevent universities to use their data. This is known as privacy paradox, because although students are protective over their data, they are also indifferent in taking appropriate actions to ensure data security (Tsai et al., 2020). However, Tsai et al. (2020) did not analyse *multimodal data* but rather asked students about other more typically types of data commonly used in learning analytics, such as clickstreams, academic, and socio-demographic-economic data. In our study, multimodal data includes physiological data, electrophysiological activity of the brain, and body-related data, which are perceived as more invasive (*under-the-skin*) types of data that can impose an additional moral dilemma—one's right to control the use and disposition of one's body (Alterman, 2003). Therefore, because multimodal data can include more intimate types of data that might be affected by various factors unrelated to learning (Järvelä et al., 2019), not all data captured about learners is relevant to the learning activity. Thereby, access and use of multimodal data require a stronger focus on privacy protection.

Ethics in the external limitations dimension refers to *data profiting*, *data misuse* and *unequal power relations*. CS students expressed their concerns regarding these issues and suggested that researchers and educators need to consider it when designing learning technologies, due to the need for quality, trust and sense of agency (Scheffel et al., 2014). They also believe that one way to avoid these issues and employ analytics to pursue positive outcomes and benefits in education, requires educators to be *incentivized* and supported in interpreting and acting upon the outcomes of the data analysis, because data could never provide educators with a full understanding of students learning experiences. Moreover, educators' knowledge (e.g., *data literacy*, *misconceptions*) and abilities (e.g., *critical thinking*) presented in the *internal limitations dimension* play an important role in the connection between the data dimension, and the external limitations dimension. Educators are required to apply analytics output as an input in subsequent decision-making, such as, *interventions*, *scaffolds*, and *(re)design of lectures or assignments*, presented in the teacher inquiry dimension. Hence, if educators are not able to make sense of the data, and link that data with a pedagogical intent, their responsive actions towards learners may be vulnerable to misinterpretation and bias (Wise & Jung, 2019). These insights highlight that the value of technology is not only with the functionality of a tool, but also with the subsequent importance to the people who use it (identification of the meaning which the technology offer, i.e., the why).

5.1 | Limitations

In addition to reporting the contributions from our study, we also report some limitations to which our contributions are subjected to. First, we acknowledge the potential bias in the sampling. The participants in our study were CS students that already had some experience with sensor devices, such as a smart watch. Therefore, the findings do not represent perspectives of students who are not familiar with sensor devices as consumers. This sampling bias also implies that such students may be more familiarized with multimodal data that are captured with their smartwatches, which they tend to use it as interactive tools rather than learning tools. In addition, most of our participants in the sample were overwhelmingly male, suggesting that female CS students might have different attitudes. Second, our study privileged computer-mediated learning, where our students were engaged individually with a computing device for the entire activity, reflecting a certain learning context, that is, computer-supported individual learning. Finally, the study was performed in a controlled environment that might affect the ecology of the study, because the participants were aware of the physiological data collection, which may cause increase in desire 'to perform' and generate good biomarker metrics. Also, because the participants were subjected to a walk-through in the sensor devices before the start of the learning activity, the wording of the information might have influenced some of the participants' perceptions.

6 | CONCLUDING REMARKS

Based on the findings of our analysis, we argue that incorporating human values and needs in the design of learning technologies requires multi-disciplinary approaches, new methods, skills and knowledge, to ensure that technology will serve users' needs in the most ethical and practical way. To that end, our paper aims to provoke deeper, more productive reflections and conversations about multimodal data and its value for learners in higher education. Like in every scientific discipline, there is no single, unambiguous truth to discern; thus, moving the field of educational technology forward is fraught with risks and obstacles, that can be overcome with a consolidated model of theory, human-centred design and data science.

ACKNOWLEDGEMENTS

This work was supported by the Research Council of Norway under the project FUTURE LEARNING (255129/H20). In addition, the authors are extremely grateful to the associate editor and the reviewers for their constructive comments and useful insights, which significantly improved the paper.

CONFLICT OF INTEREST

We have no conflicts of interest to disclose.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jcal.12542.

DATA AVAILABILITY STATEMENT

Due to the nature of this research, participants of this study did not agree for their data to be shared publicly, so supporting data are not available.

ORCID

Katerina Mangaroska https://orcid.org/0000-0002-7853-0429
Roberto Martinez-Maldonado https://orcid.org/0000-0002-8375-1816
Boban Vesin https://orcid.org/0000-0002-6490-4311
Dragan Gašević https://orcid.org/0000-0001-9265-1908

REFERENCES

Alterman, A. (2003). "A piece of yourself": Ethical issues in biometric identification. *Ethics and Information Technology*, *5*(3), 139–150.

Antonenko, P., Paas, F., Grabner, R., & Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*(4), 425–438.

Arora, P. (2016). Bottom of the data pyramid: Big data and the global south. *International Journal of Communication*, *10*, 19.

Aslan, S., Alyuz, N., Tanriover, C., Mete, S. E., Okur, E., D'Mello, S. K., & Arslan Esme, A. (2019). Investigating the impact of a real-time, multimodal student engagement analytics technology in authentic classrooms. In *Proceedings of the 2019 chi conference on human factors in computing systems* (p. 304).

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, *50*(1), 84–94.

Banathy, B. H. (1991). *Systems design of education: A journey to create the future*. Educational Technology.

Basch, C. E. (1987). Focus group interview: An underutilized research technique for improving theory and practice in health education. *Health Education Quarterly*, *14*(4), 411–448.

Beattie, S., Woodley, C., & Souter, K. (2014). Creepy analytics and learner data rights. *Rhetoric and reality: Critical perspectives on educational technology*. *Proceedings ascilite* (pp. 421–425).

Beyer, H., & Holtzblatt, K. (1997). *Contextual design: Defining customer-centered systems*. Elsevier.

Bixler, R., & D'Mello, S. (2016). Automatic gaze-based user-independent detection of mind wandering during computerized reading. *User Modeling and User-Adapted Interaction*, *26*(1), 33–68.

Blikstein, P. (2013). Multimodal learning analytics. In *Proceedings of the third international conference on learning analytics and knowledge* (pp. 102–106).

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.

Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., & Verbert, K. (2018). Open learner models and learning analytics dashboards: a systematic review. In *Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 41–50).

Boekaerts, M., De Koning, E., & Vedder, P. (2006). Goal-directed behavior and contextual factors in the classroom: An innovative approach to the study of multiple goals. *Educational Psychologist*, *41*(1), 33–51.

Bonato, P. (2003). Wearable sensors/systems and their impact on biomedical engineering. *IEEE Engineering in Medicine and Biology Magazine*, *22*(3), 18–20.

Bosch, N., D'Mello, S., Baker, R., Ocumpaugh, J., Shute, V., Ventura, M., … Zhao, W. (2015). Automatic detection of learning-centered affective states in the wild. In *Proceedings of the 20th international conference on intelligent user interfaces* (pp. 379–388).

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, *6*(2), 1–9.

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, *1*(1), 18–37.

Chan, M. C. E., Ochoa, X., & Clarke, D. (2020). Multimodal learning analytics in a laboratory classroom. In *Machine learning paradigms* (pp. 131–156). Springer.

Chen, B., & Zhu, H. (2019). Towards value-sensitive learning analytics design. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 343–352).

Corrin, L., & De Barba, P. (2015). How do students interpret feedback delivered via dashboards? In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 430–431).

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. *Accessed online*, 26.

Coventry, L. (2012). Interfaces magazine, special issue on "slow hci-designing to promote well-being for individuals, society and nature", vol. 92.

Crescenzi-Lanna, L. (2020). Multimodal learning analytics research with young children: A systematic review. *British Journal of Educational Technology* (in press), *51*(5), 1485–1504.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.

Daniel, B. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, *46*(5), 904–920.

Davis, J., & Nathan, L. P. (2015). *Value sensitive design: Applications, adaptations, and critiques. Handbook of ethics, values, and technological design: Sources, theory, values and application domains* (pp. 11–40), Springer.

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education*, *69*, 485–492.

Di Mitri, D. (2019). Detecting medical simulation errors with machine learning and multimodal data. In *17th conference on artificial intelligence in medicine* (pp. 1–6).

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.

D'Mello, S., & Graesser, A. (2011). The half-life of cognitive-affective states during complex learning. *Cognition & Emotion*, *25*(7), 1299–1308.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157.

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 & knowledge* (pp. 89–98).

Drachsler, H., & Schneider, J. (2018). JCAL special issue on multimodal learning analytics. *Journal of Computer Assisted Learning*, *34*(4), 335–337.

Durall, E., Leinonen, T., et al. (2015). Feeler: Supporting awareness and reflection about learning through eeg data. In *Artel@ ec-tel* (pp. 67–73).

Echeverria, V., Martinez-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* (p. 39).

Echeverria Barzola, V. I. (2020). *Designing feedback for collocated teams using multimodal learning analytics* (Unpublished doctoral dissertation).

Friedman, B., & Kahn Jr, P. H. (2003). Human values, ethics, and design. The human-computer interaction handbook, 1177–1201.

Galán, F. C., & Beal, C. R. (2012). Eeg estimates of engagement and cognitive workload predict math problem solving outcomes. In *International conference on user modeling, adaptation, and personalization* (pp. 51–62).

Gill, K. S. (2012). *Human machine symbiosis: The foundations of human-centred systems design*. Springer Science & Business Media.

Greller, W., & Drachsler, H. (2012). *Translating learning into numbers: A generic framework for learning analytics*. International Forum of Educational Technology & Society.

Hassenzahl, M., & Sandweg, N. (2004). From mental effort to perceived usability: Transforming experiences into summary assessments. In Chi'04 extended abstracts on human factors in computing systems (pp. 1283–1286).

Hassib, M., Schneegass, S., Eiglsperger, P., Henze, N., Schmidt, A., & Alt, F. (2017). Engagemeter: A system for implicit audience engagement sensing using electroencephalography. In *Proceedings of the 2017 chi conference on human factors in computing systems* (pp. 5114–5119).

Henman, P. (2004). Targeted! Population segmentation, electronic surveillance and governing the unemployed in Australia. *International Sociology*, *19*(2), 173–191.

Huertas Celdrán, A., Ruipérez-Valiente, J. A., García Clemente, F. J., Rodríguez-Triana, M. J., Shankar, S. K., & Martínez Pérez, G. (2020). A scalable architecture for the dynamic deployment of multimodal learning analytics applications in smart classrooms. *Sensors*, *20*(10), 2923.

Jacobson, M. J., Kapur, M., & Reimann, P. (2016). Conceptualizing debates in learning and educational research: Toward a complex systems conceptual framework of learning. *Educational Psychologist*, *51*(2), 210–218. https://doi.org/10.1080/00461520.2016.1166963

Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2019). What multimodal data can tell us about the students' regulation of their learning process? *Learning and Instruction*, *72*, 101203. https://www.sciencedirect.com/science/article/pii/S095947521830416X?via%3Dihub.

Jewitt, C. (2012). *Technology, literacy, learning: A multimodal approach*. Routledge.

Klimesch, W. (1999). Eeg alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2–3), 169–195.

Knight, D. B., Brozina, C., & Novoselich, B. (2016). An investigation of first-year engineering student and instructor perspectives of learning analytics approaches. *Journal of Learning Analytics*, *3*(3), 215–238.

Kress, G. (2001). *Multimodal teaching and learning: The rhetorics of the science classroom*. A&C Black.

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann.

Lee, M., Kirschner, P. A., & Kester, L. (2016). Learning analytics in massively multi-user virtual environments and courses. *Journal of Computer Assisted Learning*, *32*(3), 187–189.

Lewis, A. (1992). Group child interviews as a research tool. *British Educational Research Journal*, *18*(4), 413–421.

Li, Y., Chang, M., Kravcik, M., Popescu, E., Huang, R., Chen, N.-S., et al. (2016). *State-of-the-art and future directions of smart learning*. Springer.

Lim, L., Dawson, S., Joksimovic, S., & Gašević, D. (2019). Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference? In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 250–259).

Maguire, M., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *AISHE-J: The All Ireland Journal of Teaching and Learning in Higher Education*, *9*(3), 335-1–335-14. https://ojs.aishe.org/index.php/aishe-j/article/view/335.

Mangaroska, K., & Giannakos, M. N. (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, *12*(4), 516–534.

Mangaroska, K., Sharma, K., Giannakos, M., Trætteberg, H., & Dillenbourg, P. (2018). Gaze insights into debugging behavior using learner-centred analysis. In *Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 350–359).

Mangaroska, K., Vesin, B., & Giannakos, M. (2019). Cross-platform analytics: A step towards personalization and adaptation in education. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 71–75).

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* (pp. 1–15).

Martinez-Maldonado, R., Schulte, J., Echeverria, V., Gopalan, Y., & Shum, S. B. (2020). Where is the teacher? Digital analytics for classroom proxemics. *Journal of Computer Assisted Learning*, *36*, 741–762.

Matcha, W., Gasevic, D., Pardo, A., et al. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226–245.

McDuff, D., Karlson, A., Kapoor, A., Roseway, A., & Czerwinski, M. (2012). Affectaura: an intelligent system for emotional memory. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 849–858).

McLafferty, I. (2004). Focus group interviews as a data collecting strategy. *Journal of Advanced Nursing*, *48*(2), 187–194.

Mladenovic, M. N., Mangaroska, K., & Abbas, M. M. (2016). Assessment of students' preconceptions in an introductory transportation engineering course: Case study at Virginia tech. *Journal of Professional Issues in Engineering Education and Practice*, *142*(3), 05016002.

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International Journal of Qualitative Methods*, *1*(2), 13–22.

Mousavinasab, E., Zarifsanaiey, N., Niakan Kalhori, R., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M. (2018). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, *29*(1), 142–163.

Norman, D. A., & Draper, S. W. (1986). *User centered system design: New perspectives on human-computer interaction*. CRC Press.

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

Ochoa, X. (2017). Multimodal learning analytics. *Handbook of Learning Analytics*, *1*, 129–141.

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).

Ogan, A. (2019). Reframing classroom sensing: Promise and peril. *Interactions*, *26*(6), 26–32.

Olguín Olguín, D. (2011). *Sensor-based organizational design and engineering* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Oviatt, S. (2006). Human-centered design meets cognitive load theory: designing interfaces that help people think. In *Proceedings of the 14th acm international conference on multimedia* (pp. 871–880).

Oviatt, S., Grafsgaard, J., Chen, L., & Ochoa, X. (2018). Multimodal learning analytics: Assessing learners' mental state during the process of learning. In *The handbook of multimodal-multisensor interfaces* (pp. 331–374). Association for Computing Machinery (ACM).

Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., & Potamianos, G. (2017). *The handbook of multimodal-multisensor interfaces, volume 1: Foundations, user modeling, and common modality combinations*. Morgan & Claypool.

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, *45*(3), 438–450.

Picard, R. W., & Healey, J. (1997). Affective wearables. *Personal Technologies*, *1*(4), 231–240.

Pijeira-Díaz, H. J., Drachsler, H., Kirschner, P. A., & Järvelä, S. (2018). Profiling sympathetic arousal in a physics course: How active are students? *Journal of Computer Assisted Learning*, *34*(4), 397–408.

Poh, M.-Z., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Transactions on Biomedical Engineering*, *57*(5), 1243–1252.

Prieto, L. P., Sharma, K., Kidzinski, Ł., Rodríguez-Triana, M. J., & Dillenbourg, P. (2018). Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. *Journal of Computer Assisted Learning*, *34*(2), 193–203.

Prieto-Alvarez, C. G., Martinez-Maldonado, R., & Shum, S. B. (2018). Mapping learner-data journeys: evolution of a visual co-design tool. In *Proceedings of the 30th australian conference on computer-human interactions* (pp. 205–214).

Prinsloo, P., & Slade, S. (2016). Student vulnerability, agency, and learning analytics: An exploration. *Journal of Learning Analytics*, *3*(1), 159–182.

Reimann, P., Bull, S., Kickmeier-Rust, M., Vatrapu, R., & Wasson, B. (2015). *Measuring and visualizing learning in the information-rich classroom*. Routledge.

Ritella, G., & Hakkarainen, K. (2012). Instrumental genesis in technology-mediated learning: From double stimulation to expansive knowledge practices. *International Journal of Computer-Supported Collaborative Learning*, *7*(2), 239–258.

Rodríguez-Triana, M. J., Prieto, L. P., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2018). The teacher in the loop: Customizing multimodal learning analytics for blended learning. In *Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 417–426).

Roschelle, J. (1997). *Learning in interactive environments: Prior knowledge and new experience*. Citeseer.

Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Journal of Educational Technology & Society*, *17*(4), 117–132.

Schmidt, A., Langheinrich, M., & Kersting, K. (2011). Perception beyond the here and now. *IEEE Computer*, *44*(2), 86–88.

Schneider, B. (2020). A methodology for capturing joint visual attention using mobile eye-trackers. *JoVE (Journal of Visualized Experiments)*, *155*, e60670.

Schumacher, C., & Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in Human Behavior*, *78*, 397–407.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., & Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, *10*(1), 30–41.

Shankar, S. K., Prieto, L. P., Rodríguez-Triana, M. J., & Ruiz-Calleja, A. (2018). A review of multimodal learning analytics architectures. In *2018 ieee 18th international conference on advanced learning technologies (icalt)* (pp. 212–214).

Shankar, S. K., Rodríguez-Triana, M. J., Ruiz-Calleja, A., Prieto, L. P., Chejara, P., & Martínez-Monés, A. (2020). Multimodal data value chain (m-dvc): A conceptual tool to support the development of multimodal learning analytics solutions. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, *15*(2), 113–122.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, *57*(10), 1380–1400.

Simonsen, J., & Robertson, T. (2012). *Routledge international handbook of participatory design*. Routledge.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). *The challenges of defining and measuring student engagement in science*. Taylor & Francis.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510–1529.

Solove, D. J. (2005). A taxonomy of privacy. *University of Pennsylvania Law Review*, *154*, 477.

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.

Stephanidis, C. C., Salvendy, G., Members of the Group Margherita Antona, Chen, J. Y. C., Dong, J., Duffy, G., … Zhou, J. (2019). Seven HCI grand challenges. *International Journal of Human–Computer Interaction*, *35*(14), 1229–1269. https://doi.org/10.1080/10447318.2019.1619259

Stone, C., Donnelly, P. J., Dale, M., Capello, S., Kelly, S., Godley, A., & D'Mello, S. K. (2019). Utterance-level modeling of indicators of engaging classroom discourse. In *Proceedings of the 12th international conference on educational data mining (edm 2019)* (pp. 420–425).

Sun, K., Mhaidli, A. H., Watel, S., Brooks, C. A., & Schaub, F. (2019). It's my data! tensions among stakeholders of a learning analytics dashboard. In *Proceedings of the 2019 chi conference on human factors in computing systems* (p. 594).

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research*, *64*(1), 37–54.

Tsai, Y.-S., Whitelock-Wainwright, A., & Gašević, D. (2020). The privacy paradox and its implications for learning analytics. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 230–239).

Van Merriënboer, J. J., & Kirschner, P. A. (2017). *Ten steps to complex learning: A systematic approach to four-component instructional design*. Routledge.

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, *89*, 98–110.

Wachsmuth, I., Lenzen, M., & Knoblich, G. (2008). Introduction to embodied communication: Why communication needs the body. In *Embodied communication in humans and machines* (pp. 1–28). Oxford University Press.

Was, C., Sansosti, F., & Morris, B. (2016). *Eye-tracking technology applications in educational research*. IGI Global.

Wilson, G. F. (2002). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *The International Journal of Aviation Psychology*, *12*(1), 3–18.

Wise, A. F., & Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics*, *6*(2), 53–69.

Worsley, M., Abrahamson, D., Blikstein, P., Grover, S., Schneider, B., & Tissenbaum, M. (2016). Situating multimodal learning analytics. In *Proceedings of international conference of the learning sciences, icls* (Vol. 2, pp. 1346–1349).

Worsley, M., & Blikstein, P. (2015). Leveraging multimodal learning analytics to differentiate student learning strategies. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 360–367).

Yue, Z., Litt, E., Cai, C. J., Stern, J., Baxter, K. K., Guan, Z., … Zhang, G. G. (2014). Photographing information needs: the role of photos in experience sampling method-style research. In *Proceedings of the sigchi conference on human factors in Computing Systems* (pp. 1545–1554).

How to cite this article: Mangaroska K, Martinez-Maldonado R, Vesin B, Gašević D. Challenges and opportunities of multimodal data in human learning: The computer science students' perspective. *J Comput Assist Learn*. 2021;1–18. https://doi.org/10.1111/jcal.12542

APPENDIX A.

PARTICIPANTS' WALK-THROUGH OF THE DATA STREAMS (I.E., BRAIN ACTIVITY, GAZE DATA, FACIAL DATA AND PHYSIOLOGICAL DATA SUCH AS HEART RATE) USED IN THE STUDY

During the learning activity, which will last for 45 min, we will be collecting data from you using the sensors that we just placed on you (i.e., on your head and on your wrist) and the sensors that are attached on the computer screen. We will not show you any measures from your data during the learning activity, nor we will show you the raw collected data. However, because this is your data you have the right to know what type of data we actually will collect. If you do not feel comfortable after hearing the type of data we will be collecting and after reading the consent form, you are free to leave the study, now or at any moment you would like to.

This is a quick walk-through the types of data that we will be collecting:

1. Let us start with the camera. The camera is focused on your face and will record your expressions throughout the activity. Whenever you are smiling or you are bored we will be able to see that once we analyse the video recordings.
2. The eye tracker that is placed at the bottom of the computer screen will record your gaze. With this device, we will be able to see where you were looking at on the screen during the activity.
3. The sensor device that you have on your wrist is similar to a smart watch. Using this device, we will follow and record your heart rate (i.e., whether your heart beats faster or slower and when during the learning activity); the temperature of your skin (because, e.g., due to various physical and emotional stimuli [cold weather, stress] your blood vessels can increase or decrease the flow of the blood and with that cause decrease or increase in the temperature of your skin); and the electrical activity of your skin which changes when your glands are active (e.g., you are having 'wet' palms when you are worried or nervous). As you can see on the phone screen, the data are continuous and presented as signal waves along the *x*- and *y*-axes.
4. Finally, on your head you have an electroencephalogram cap that will collect data from the activity of your brain. There are 20 electrodes (one is attached on your ear) that will record different brain waves from different regions on your head during the activity. As you can see now on the interface, where I am checking to see if all of the electrodes are green and picking up signal, the data are continuous and presented as signal waves along the *x*- and *y*-axes. The signal line can have a different shape (depending on frequency and amplitude) which represents different states, depending if you are expressing excitement or relaxation.

APPENDIX B.

MAPPING BETWEEN THE THEMES AND THE DIMENSIONS

Theme 1	
level of comfort wearing sensors	instruments
flexibility when wearing sensors	instruments
how wearables affect learning	instruments
Theme 2	
type of sensors and data	instruments, data
increase motivation and engagement	objective, engagement
capturing pitfalls, misconceptions, and tracking performance	content
data literacy skills	internal limitations
incentives to use MLA	internal limitations, external limitations
Theme 3	
wrong conduct and data-profiting	external limitations
more power for teachers	external limitations
access to data – what, when, who	external limitations
regulations for storing/using data	external limitations
constantly and pervasively surveilled	external limitations
Theme 4	
reflection and monitoring of progress	objective
cognition, affect, and behavior of individuals	engagement
identify areas for improvement	objective
support learning not just passing a course/getting a good grade	objective, teacher inquiry
personalize feedback, instruction, and scaffolds	teacher inquiry, future learning
Theme 5	
reasons to use MLA in education	objective
users and uses of MLA	objective, context
plan, design, and adapt curricula, assignments, and exams	teacher inquiry
effect of individual differences on learning	content
set the right feedback loop	teacher inquiry, internal limitations
Theme 6	
emerging learning technologies	future learning
adaptation to individuality	internal limitations
organization and optimization considering cognition, affect, and behavior	objective
personalized smart apps	future learning
external factors that affect learning	future learning, content