

Review

# Multimodal Data Fusion in Learning Analytics: A Systematic Review

Su Mu <sup>1</sup>, Meng Cui <sup>1,\*</sup>  and Xiaodi Huang <sup>2</sup> 

<sup>1</sup> School of Information Technology in Education, South China Normal University, Guangzhou 510631, China; musu@m.scnu.edu.cn

<sup>2</sup> School of Computing and Mathematics, Charles Sturt University, Albury, NSW 2640, Australia; xhuang@csu.edu.au

\* Correspondence: mengcui@m.scnu.edu.cn

Received: 10 November 2020; Accepted: 28 November 2020; Published: 30 November 2020



**Abstract:** Multimodal learning analytics (MMLA), which has become increasingly popular, can help provide an accurate understanding of learning processes. However, it is still unclear how multimodal data is integrated into MMLA. By following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, this paper systematically surveys 346 articles on MMLA published during the past three years. For this purpose, we first present a conceptual model for reviewing these articles from three dimensions: data types, learning indicators, and data fusion. Based on this model, we then answer the following questions: 1. What types of data and learning indicators are used in MMLA, together with their relationships; and 2. What are the classifications of the data fusion methods in MMLA. Finally, we point out the key stages in data fusion and the future research direction in MMLA. Our main findings from this review are (a) The data in MMLA are classified into digital data, physical data, physiological data, psychometric data, and environment data; (b) The learning indicators are behavior, cognition, emotion, collaboration, and engagement; (c) The relationships between multimodal data and learning indicators are one-to-one, one-to-any, and many-to-one. The complex relationships between multimodal data and learning indicators are the key for data fusion; (d) The main data fusion methods in MMLA are many-to-one, many-to-many and multiple validations among multimodal data; and (e) Multimodal data fusion can be characterized by the multimodality of data, multi-dimension of indicators, and diversity of methods.

**Keywords:** multimodal learning analytics; data fusion; multimodal data; learning indicators; online learning

## 1. Introduction

Learning analytics refers to the measurement, collection, analysis, and reporting of data about learners and their learning contexts, for understanding and optimizing learning and the environment in which it occurs [1]. The data for traditional learning analytics is usually unidimensional [2]. For example, only log data rather than all data generated by a learning management system are commonly used for analyzing the online learning process. Specifically, log data ignore important contextual information about learners [3]. These context data are crucial for understanding students' learning processes. In other words, unidimensional data provide only partial information about the learning process [4,5], which makes it impossible to produce accurate results of learning analytics [6]. The real learning process is complex [7]. To understand a learning process accurately [8], we must collect multimodal data such as learning behavior data, facial expression data, and physiological data as much as possible [7]. In this way, a better, more holistic picture of learning can be revealed.

As a new area of learning analytics [7], multimodal learning analytics (MMLA) [9] captures, integrates, and analyzes learning traces from different sources in a way that enables a holistic understanding of a learning process. By leveraging sophisticated machine learning and artificial intelligence techniques [10], MMLA focuses mainly on its paradigms [11,12], framework [6,13], multimodal data [14,15], system [16,17], multimodal data value chain [18], and case studies [19,20].

Data fusion is a crucial component in MMLA [21]. Different types of data play different roles in their integration. However, data from different sources are often collected at different grain sizes. This makes it difficult to integrate them [22]. Therefore, it is necessary to review the existing MMLA research to understand the ways multimodal data is integrated. This literature review can help researchers to have a deeper understanding of multimodal data integration and promote the development of related research in this area.

The available reviews on MMLA have been conducted from different perspectives, such as its past, present, and potential futures [10,23–25], architectures [26], the multimodal data, and the learning theories in MMLA [7], and MMLA for children [27]. However, to the best of our knowledge, a systematic review with a focus on data fusion in MMLA is not available. To fill this gap, we attempt to present a systematic review of MMLA articles published between 2017–2020, by answering the question of how to integrate and analyze multimodal data. Conducting a detailed review of the research in MMLA, we aim to have a clear understanding of the current research status of multimodal data integration by analyzing the approaches for integrating multimodal data. Through the systematic analysis, we outline the future directions of multimodal data integration. Specifically, the three research questions in this study are as follows:

RQ 1: What is the overall status of MMLA research? (Section 2.4)

RQ 2: What types of multimodal data and learning indicators are used in MMLA? What are the relationships between multimodal data and learning indicators? (Section 5)

RQ 3: How can multimodal data be integrated into MMLA? What are the main methods, key stages, and main features of data fusion in MMLA? (Section 6)

The contributions of this paper are: (1) we propose a novel MMLA framework; (2) according to the proposed framework, this paper summarizes the broad data types and learning indicators in MMLA, proposes a multimodal data classification framework, and characterizes the relationships between multimodal data and learning indicators, and; (3) we review the integration methods and main stages of data integration in MMLA, describing the three-dimensional characteristics of data integration.

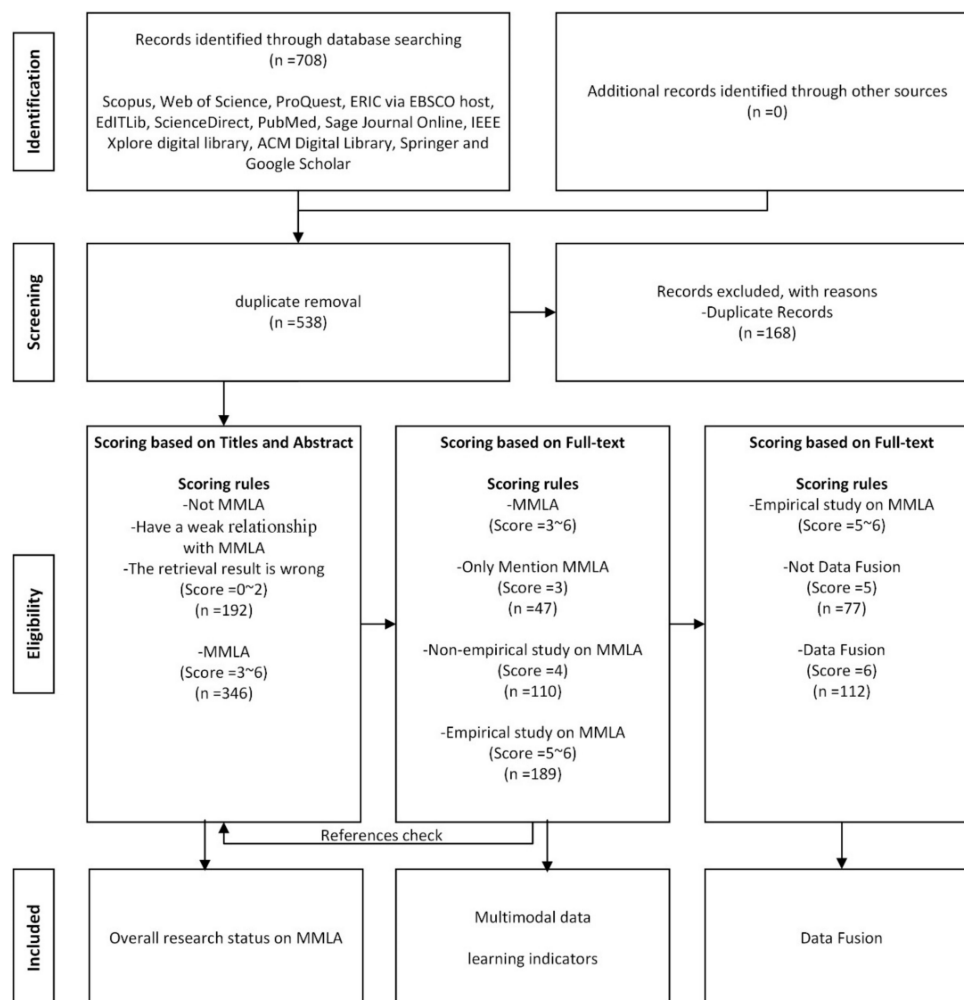
The rest of this paper is organized as follows. Section 2 describes our methods and the detailed process of literature review and summarizes the overall research status of MMLA. Section 3 presents the MMLA model. Section 4 outlines the data types, learning indicators, and their relationships in MMLA. Section 5 reviews the data integration methods, main stages in MMLA and points out the future research directions in MMLA. Section 6 concludes the paper.

## 2. Survey Methods

As a method for systematic review and meta-analysis, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [28] is commonly used for reporting an evidence-based minimum set of items. Primarily in the context of healthcare, PRISMA provides guidelines that consist of a checklist of 27 items on the title, abstract, methods, results, discussion, and funding, as well as on a four-phase flow diagram. The flow diagram illustrates the systematic review and clearly outlines the study identification, screening, eligibility, and inclusion processes, including reasons for study exclusion.

Following the PRISMA guidelines [28], we conducted a systematic review on how to integrate multimodal data in MMLA, by using an explicit and replicable search strategy. In particular, we selected the literature on MMLA based on pre-determined criteria, which have been used for other systematic reviews in education research [29–33]. The procedure of our review is illustrated in the flow diagram

in Figure 1. First, the relevant articles from the databases were retrieved, followed by removing duplicate articles. The articles were then scored and coded according to the inclusion and exclusion criteria. Finally, we conducted a detailed analysis of all the included articles by answering the proposed research questions.



**Figure 1.** Flow diagram based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

### 2.1. Search Method

Using the keywords “Multimodal Learning Analytics”, “MMLA”, “multimodal”, and “Learning Analytics”, we retrieved relevant papers from 12 bibliographic databases. They were Scopus, Web of Science, ProQuest, ERIC via EBSCO host, EdITLib, ScienceDirect, PubMed, Sage Journal Online, IEEE Xplore digital library, ACM Digital Library, Springer, and Google Scholar. Their references to the key retrieved articles were further retrieved for reviewing additional relevant papers. All the articles were limited to publications between January 2017 and June 2020. Additionally, three separate searches were conducted for those published in December 2019, March 2020, and June 2020. In November 2020, we conducted the last round of supplementary searches. As a result, the initial search produced 708 articles.

### 2.2. Inclusion and Exclusion Criteria

Table 1 shows the inclusion and exclusion criteria for this review. All the reviewed articles have met the inclusion criteria. Finally, 538 articles in total were included after removing duplicate articles.

**Table 1.** Inclusion and Exclusion Criteria for Reviewing Papers.

Inclusion Criteria	Exclusion Criteria
The following search keywords are included in the title, abstract, or keywords	<ul style="list-style-type: none"> <li>• Studies published before 2017</li> </ul>
<ul style="list-style-type: none"> <li>• “Multimodal Learning Analytics” OR</li> <li>• “MMLA” OR</li> <li>• “Learning analytics” and “multimodal”</li> </ul>	<ul style="list-style-type: none"> <li>• Duplicate papers (only one paper included)</li> <li>• Articles unrelated to MMLA content</li> <li>• Non-English papers</li> <li>• Not Peer-Reviewed</li> </ul>

### 2.3. Scoring and Encoding

After reading its title, abstract, and full text, we scored each of the 538 articles according to the scoring rules listed in Table 2. Highly similar articles, for example, [4,34,35], were treated as the same category by assigning the same score. We excluded articles with their scores below 3 (not including 3) because they have no or little relation to MMLA. In the end, 346 articles on MMLA were included. Articles with different scores were used to answer different research questions. We conducted a detailed analysis of each included empirical study in the paper, identified its multimodal data and learning indicators, and distinguished them by using short notations. For example, the eye movement data were denoted as EM and the electroencephalogram data as EEG. The detailed notations are given in Table 3.

**Table 2.** Scoring rules.

	Scoring Rules	Score	RQ
Title and Abstract	The topic has nothing with MMLA	(Score = 0)	
	The topic has a little with MMLA	(Score = 1~2)	
	MMLA	(Score = 3~6)	Q1 & Q2 & Q3
Full-text	3.1 Only mentioned MMLA	(Score = 3)	Q1
	3.2 Non-empirical study on MMLA, such as its review and theory	(Score = 4)	Q1
	3.3 An empirical study on MMLA	(Score = 5~6)	Q2
	Without Data Fusion	(Score = 5)	
	With Data Fusion	(Score = 6)	Q3

**Table 3.** Multimodal data classification and case studies.

Type	Multimodal Data and Code	Case Studies	Author
<b>Digital space</b>			
Clickstream	Log data LOG	Log data as a proxy measure of student engagement	[36]
		Interactions in STEAM by a physical computing platform	[37]
	Mouse MO	Behavioral engagement detection of students	[38]
	Keystrokes KS	Surrogate measure for the effort put in by the student	[39]
Qualitative data	Text TE	Learners’ emotions from pedagogical texts	[40]
	Handwriting Dynamic Handwriting Signal	Dynamic handwriting signal to predict domain expertise	[41]
	Features HW	A sensitive measure of handwriting performance	[42]
	Digital footnote DF	Analyzing students’ reviewing behavior	[43,44]

Table 3. Cont.

Type	Multimodal Data and Code	Case Studies	Author
<b>Physical space</b>			
Eye	Eye movement EM	Students/teacher co-attention (i.e., with-me-ness)	[45]
		Improving communication between pair programmers	[46]
	Eye Contact EC	Joint Visual Attention	[47]
		Eye contact in three-party conversations	[48]
Mouth	Audio AU	Exploring collaborative writing of user stories	[49]
		Think-aloud protocols used in cognitive and metacognitive activities	[50]
Face	Facial Expression FE	Investigating emotional variation during interaction	[51]
		Automated detection of engagement behaviors of lecturers and students	[52]
	Facial Region FR	Student behavior monitoring systems	[53]
		Assess the effect of different levels of cognitive load on facial temperature	[54]
Head	facial temperature FT	behavioral engagement detection of students	[55]
		Modeling collaborative problem-solving competence	[56]
Hand	Head Region HER	data glove which captures pressure sensitivity designed to provide feedback for palpation tasks	[57]
		Using hand motion to understand embodied mathematical learning	[58]
Arms	Hand HA	Dynamic adaptive gesturing predicts domain expertise in mathematics	[59]
		Embodied learning behavior in the mathematics curriculum	[60]
Leg	Arms AR	Step counts are used to predicting learning performance in ubiquitous learning	[61]
Body	step count SC	Enhancing multimodal learning through personalized gesture recognition	[62]
		Embodied strategies in the teaching and learning of science	[63]
	Body posture BL	Making spatial pedagogy visible using positioning sensors	[64]
		Tracing students' physical movement during practice-based learning	[65]
	Body Movement and Location MP	Aggregating positioning and orientation in the visualization of classroom proxemics	[66]
Orientation OR	Orientation OR	Detecting cognitive load using EEG during learning	[67]
		Multimodal emotion recognition	[68]
<b>Physiological space</b>			
Brain	Electroencephalogram EEG	Detecting cognitive load using EEG during learning	[67]
		Multimodal emotion recognition	[68]

Table 3. Cont.

Type	Multimodal Data and Code	Case Studies	Author
Skin	electrodermal activity EDA	Profiling sympathetic arousal in a physics course	[69]
	galvanic skin response GSR	The difficulty of learning materials	[70]
	skin temperature ST	Recognition of emotions	[71]
Heart	Electrocardiogram ECG	EDA and ECG study of pair-programming in a classroom environment	[72]
	Photoplethysmography PPG	Recognition of emotions	[73]
	heart rate /variability HR/HRV	Automated detection of engagement	[52]
Blood	blood volume pulse BVP	Recognition of emotions	[71]
Lung	Breathe respiration BR	Recognition of emotions	[71]
<b>Psychometric space</b>			
	Motivation PS	Motivation coming from the questionnaire	[45]
<b>Environmental space</b>			
	Weather condition WC	Predicting performance in self-regulated learning using multimodal data, such as (1) Temperature, (2) Pressure, (3) Precipitation, (4) Weather type	[61]

#### 2.4. Overall Research Status (Q1)

Table 4 reports the scoring results. From the results, it can be seen that MMLA has focused on theoretical and empirical research. In particular, empirical research on multimodal data fusion accounts for a relatively large proportion of the overall research (37.90%). This indicated that data integration is an important part of MMLA research. There are many existing pieces of research on multimodal data integration, and most of them are empirical studies, aiming to solve a specific problem of multimodal data integration. However, there is currently a lack of a theoretical and overall review of the current research status of multimodal data integration. So, it is necessary to conduct a systematic review of how to integrate the data in MMLA.

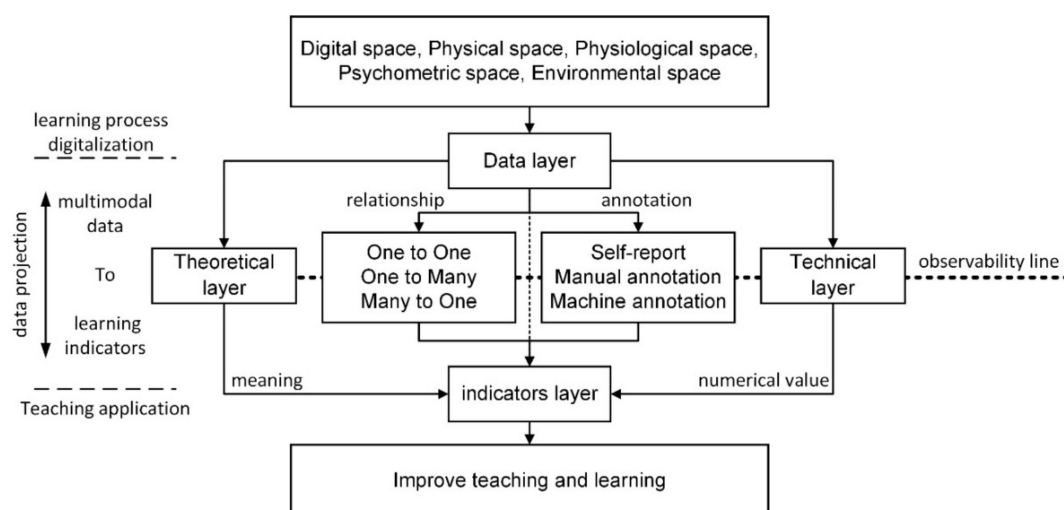
Table 4. Scoring results.

Score	Num. of Articles	Percentage	Remarks
3	47	3.36%	Only Mention MMLA
4	110	35.26%	Non-empirical study on MMLA
5	77	24.68%	An empirical study on MMLA BUT without Data Fusion
6	112	37.90%	An empirical study on MMLA AND Data Fusion

### 3. MMLA Conceptual Model

Understanding the relationships between multimodal data and learning indicators is essential for MMLA [7]. As shown in Figure 2, we proposed a conceptual model for multimodal data analysis. The purpose of the proposed conceptual model was to better understand the relationships between multimodal data and learning indicators. The conceptual model consisted mainly of three stages and four layers. The three stages were: (a) acquisition of data on the learning process, (b) mapping multimodal data into learning indicators measured, and (c) improvement of students' learning performance. The three stages focused on external learning behavior, the internal psychological

mechanism, as well as practical teaching and learning, respectively. The four layers referred to the data layer, indicator layer, theory layer, and technology layer. The data layer was about data on visible and directly measurable learning behavior, such as eye movement data. The indicator layer represented the invisible learning indicators relating to the sense-making process that cannot be directly measured, such as learning performance, behavior, and emotion. Although the analysis of multimodal data offered a holistic picture of learning, its inherent complexity made it difficult to understand and interpret [74]. Current digital systems are largely blind to users' cognitive states [55]. There is a conceptual line of demarcation between the data layer and the indicator layer. All observable evidence was above the line, with all the possible interpretations below the line. The semantic interpretation of the data layer was weak in that it could not be used to directly explain the learning process [75]. However, the data layer could be converted into the indicator layer that directly explained the learning process through psychological and educational theories (theoretical layer) and methods (technical layer). The theoretical layer was about psychological and educational theories which tell us how the relationships are drawn between multimodal data and learning indicators [7]. The technical layer was about the methods of how to transform multimodal data into learning indicators. This process is also called "data projection".



**Figure 2.** A conceptual model of multimodal data analysis.

The three types of annotation methods that transform multimodal data into learning indicators are manual annotation, self-report annotation, and machine annotation [7]. Manual annotation and self-report annotation are commonly used. However, manual annotation is time-consuming and laborious, and self-report annotation is too subjective. Therefore, these two methods are not suitable for large-scale automatic analysis. With the advance of intelligent techniques, automatic machine annotation [35] has received more and more attention. By comparing the accuracies of manual and machine annotations, some studies concluded that a combination of the two methods performs better, producing more accurate results [76–78].

The ultimate goal of MMLA is to improve the quality of teaching and learning. The applications of MMLA in teaching and learning mainly include (1) real-time visual feedback of the learning process [79–81]; (2) real-time monitoring of the learning process, such as the real-time assessment of attention in the classroom [82], and real-time analysis of teacher-student interactions in a classroom [83], and; (3) teaching design supported by multimodal data, which promotes students' cognitive development [84].

## 4. Multimodal Data, Learning Indicators and Their Relationships (Q2)

### 4.1. Multimodal Data

Most of the existing MMLA studies recognized the importance of multimodal data. However, few studies systematically classified multimodal data types. As shown in Figure 3, we grouped data used in the existing literature on MMLA [7] into different types in our multimodal data classification framework, together with the typical examples as given in Table 3.

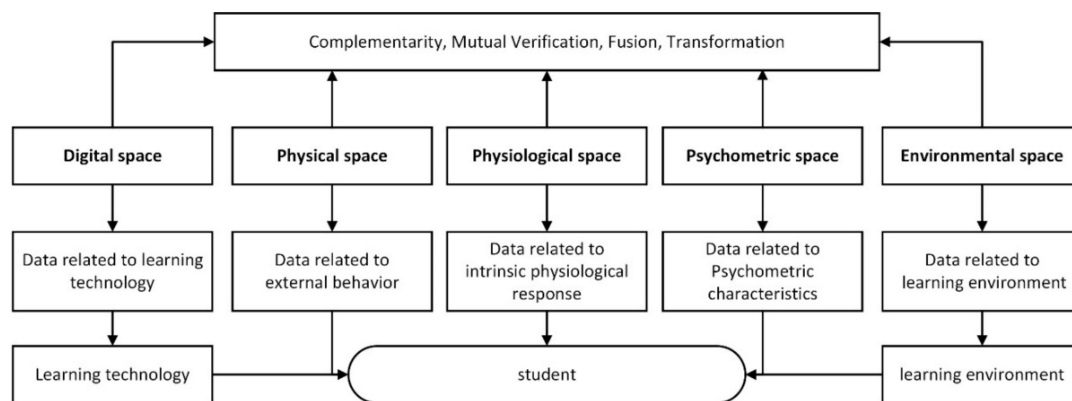


Figure 3. The classification framework of the multimodal data.

Specifically, our classification framework consisted of digital space [7], physical space [85], physiological space [71], psychometric space, and environmental space [61]. Digital space referred to various digital traces generated on the system platform during the learning process, such as an online learning platform [52], virtual experiment platform [22], or STEAM educational software [86]. Physical space was about the data obtained by various sensors, such as gesture, posture, and body movement. With the development of sensors, the physical data obtained was more refined at the micro-level, such as the angle of head movement [56] and finger movement on a screen [87]. The perception and analysis of physical data was significant for interpreting the learning process. Physiological space referred to the data related to human internal physiological reflection, including EEG and ECG, which objectively reflected students' learning status. In contrast, psychometric space, a relatively common source of learning data, referred to various self-report questionnaires that subjectively reflected the learner's mental state. Environmental space referred to the data about a learning environment where a learner was physically located, such as temperature and weather. Studies have shown that a learning environment has some influence on learning [61]. The increasing analysis of environmental data is a trend in MMLA. Based on this framework, three problems that researchers in MMLA have faced are (1) How to obtain multimodal data, (2) How to use multimodal data to infer students' learning status (emotions, cognition, attention, etc.), and (3) What learning services can be provided to students based on MMLA?

Due to technological advances such as the Internet of Things, wearable devices, and cloud data storage, learning data at the high-frequency, fine-grained, and micro-level can be collected conveniently and accurately. From multiple dimensions, MMLA reflects learners' real learning state better [7], especially in some courses [6]. Students interact with learning content, peers, and teachers in a variety of ways, such as facial expressions, audio, and body movements. It is essential that the learning processes are analyzed by using these multimodal data.

The multimodal data are complementary, mutual verification, fusional, and transformed. (a) Complementarity is an important characteristic of multimodal data. Any data types provide a partial explanation about a certain learning phenomenon or process. (b) Mutual verification—the same results are verified by different types of learning data [7]. (c) Fusion—some data integration systems store data in physical spaces such as body movements and gestures in synchronization with



log data in digital platforms [7]. (d) Transformation—physical data are transformed into digital data. Two examples are given below: digitizing students' handwriting processes through a smartpen and then predicting learning performance through dynamic writing features [41]; digitizing the traces and footnotes of students reviewing a paper test, and then analyzing students' review behavior [44,88]. The advantage of these studies is that they break the limitations of recording data methods through the use of a mouse and keyboard by retaining the information about students' authentic learning behavior and learning states as much as possible.

#### 4.2. Learning Indicators

The common learning indicators used in the MMLA literature are behavior, attention, cognition [89], metacognition [90], emotion [91], collaboration, interaction [47], engagement, and learning performance. Some of them can be further classified. In particular, learning behavior is divided into three categories—online learning behavior [88], learning behavior in the classroom [53], and embodied learning behavior [92]. Attention includes personal attention [93] and joint attention [45]. Emotions refer to those in autonomous learning [94] and collaborative learning [51]. The collaboration consists of face-to-face collaboration [48] and remote collaboration [95]. Engagement refers to engagement in autonomous learning [52] and the face-to-face classroom [96]. As a summary evaluation, the examination score [59,97], the score of game learning [98], is the common learning performance indicator. Some studies propose complex performance calculation methods to improve the accuracy of learning performance evaluation [99]. Some use formative assessment methods to evaluate learning performance, such as collaborative problem-solving ability [37,56,86,91,100]. Some studies focus on various aspects of learning performance, such as collaboration quality, task performance, and learning [101]. Skills include oral presentation skills [102] and medical operation skills [103].

By examining learning indicators, we found that: (1) There are different kinds of learning indicators, which reflect the complexity of the real learning process; and (2) The meaning of some learning indicators overlapped—the indicators related to learning scene, learning activities, and learning theory. For example, some studies conducted a separate analysis of behavior [44], cognitive engagement [104], and emotion [68] in the learning process. In contrast, some studies combined the three factors to measure learning engagement. Relying on the theory of engagement, Kim et al. [105] observed engagement by using different modalities of linguistic alignment as an indicator of cognitive engagement, kinesics as bodily engagement, and vocal cues as emotional engagement. As another example, collaborations can be analyzed separately [106] and learning engagement in collaborative learning can also be analyzed [105]. (3) There are some rules for selecting learning indicators. Collaborative learning focuses on collaborative features [56] and collaborative interaction [107], while autonomous learning focuses on attention [108], cognition [55], and engagement [39]. There are more learning indicators of face-to-face collaboration [48], with relatively few of remote collaboration [46]. (4) With a more in-depth examination of the learning process, learning indicators will be more diverse. For example, the researchers first paid attention to the learning path in the whole learning process and then focused on the learning path in each webpage from a micro perspective [109].

#### 4.3. The Relationships between Multimodal Data and Learning Indicators

MMLA creates a multi-dimensional exploration space to complicate the relationship between data and indicators [24]. The relationships between multimodal data and learning indicators are shown in Table 5. This study found that there were three types of corresponding relationships between multimodal data and learning indicators (multimodal data vs. learning indicators): one-to-one, many-to-one, and one-to-many. "One-to-one" meant that a type of data was suitable only for measuring one learning indicator. This was the most common type in MMLA literature. With the development of technology, the measurement potential of each type of data is gradually tapped. The type of the corresponding relationship, "one-to-one", has become increasingly rare. For example, the most common methods to measure cognition are interviews and self-reported questionnaires [89].

By means of the new sound-thinking method, cognition is measured by using audio data in the method of thinking aloud [50]. As the physiological measurement is available, physiological data such as EEG are also used to measure cognition [67]. We regarded these new methods as the second type of the corresponding relationship: many-to-one. Precisely, “many-to-one” referred to the fact that multiple types of data measure the same learning indicators. For example, EM, EEG, and EDA measure the degree of learning engagement of learners [110]. Finally, “one-to-many” is the third type of the corresponding relationship; that is, one type of data measures several types of learning indicators. For example, eye movement data measures attention [93], cognition [84], emotion [111], collaboration [46], and engagement [83].

**Table 5.** The relationships between multimodal data and learning indicators.

Multimodal Data	Learning Indicator						
	Behavior	Attention	Cognition Metacognition	Emotion	Collaboration	Engagement	Learning Performance
Digital space	[41] [43,44]		[112]	[40]	[95] [113]	[36] [38] [39]	[61] [98] [114] [37] [41]
Physical space	[53] [115] [116] [92] [42] [62] [63] [103] [117] [53] [64]	[93] [118] [108] [119] [82] [45] [120]	[84] [112] [50] [90]	[73] [111] [121] [51] [79] [68]	[122] [47] [48] [123] [49] [124] [125] [126] [50] [91] [127] [107] [65] [128] [95] [113] [46]	[52] [110] [38] [129] [130] [83] [38] [60]	[70] [98] [37] [98] [114] [37] [59] [61]
Physiological space			[67] [69] [131] [89]	[111] [68] [71] [73] [68]	[72] [107] [128] [122] [131] [72]	[52] [110] [130]	[132] [98] [70] [61]
Physiological space		[99] [70]	[89]	[121] [79]		[70] [129] [52]	[61]

The underlying reason why there are diverse corresponding relationships between learning data and learning indicators is that the range of valid measurement and quality of learning data vary with technical and theoretical conditions. In general, the measurement range of a particular type of data is limited, with obvious advantages in terms of its measurement. There are one or several learning indicators with better measurement effects. For example, online learning data (e.g., logs) are often used to characterize learning behavior [88], while eye movement data are often used to analyze a learner’s cognitive state, attention level, and information processing process about learning content [133]. The facial expression has a better measurement effect on emotion [68] and engagement [83]. Facial expression is a good measure of strong emotion (joy and anger), and physiological data on subtle emotion [134]. Studies have shown that a learning indicator can be measured using either single-dimensional or multi-dimensional data. The measurement of learning indicators must consider not only the optimal data but also the supplement of other types of data, which is of significance to data fusion.

## 5. Data Fusion (Q3)

We analyzed the empirical studies on multimodal data fusion from three aspects: integration methods, data type, and learning indicators. The results are reported in Table 6.

**Table 6.** Data integration in multimodal learning analytics (MMLA).

Integration Methods	Data Type	Learning Indicators	Author
Many-to-One	FE, PPG	Emotion	[73,134,135]
	AU, FA, LOG, HA	Learning performance	[37,86,100]
	LOG, AU, BL, SR	Collaboration	[114,136]
	PS, AU	Emotion	[121,137]
	AU, FE, BL, EDA, VO	Collaboration, engagement, learning performance	[78,130,138]
	EM, AU, VB, MP	Teaching behavior	[139,140]
	FE, HER, EM	Engagement	[83,141]
	FR, HER, BL	Engagement	[99,142]
	AR, HER, FR	Collaboration	[56,76,91]
	FE, EM, EEG, EDA, BVP, HR, TEMP	Engagement	[110,143]
	AU, LOG	Collaboration	[113,144]
	AU, LOG	Emotion	[145]
	AU, VB	Engagement	[105]
	FR, MO, LOG	Engagement	[146]
	FE, HR, LBP-TOP	Engagement	[52]
	AU, LOG, BL	Oral presentations	[147]
	PS, AU, FE	Emotion	[79]
	EM, EEG	Emotion	[111]
	AU, FE, ECG, EDA	Emotion	[68]
	VB, LOG	Cognition	[74]
	FE, HER, LOG	Engagement	[96]
	SC, LOG, HR, EN	Learning performance	[61]
	HER, LOG	Engagement	[148]
	PE, PS, AU, FE, BL, EM, EEG, BVP, GSA	Learning performance	[98]
	GSR, ST, HR, HRV, PD	Cognitive load	[149]
	AU, EM, LOG	Dialogue failure in human-computer interaction	[150]
	AU, HAR, FR	Collaboration	[151]
	HAR, EC, FR	Engagement	[152]
	BL, MP, LOG	Attention	[119]
	AU, FE, EM, LOG	Collaboration	[95]
EEG, EOG, ST, GSR, BVP	Emotion	[71]	
AU, EC, AR, MP	Oral presentations	[153]	
AU, BL, LOG	Embodied learning behavior	[154]	
Many-to-Many	FE, BL, AU, EC	Oral presentations	[155–157]
	BL, AU	Collaboration	[81,158–160]
	MP, AU, LOG, EDA, PS	Medical operation skills	[80,161–166]
	BL, EMG, LOG	Medical operation skills	[103,167]
	AU, EM, MP, BL	Embodied learning behavior	[168,169]
	FA, EC, MP	Face-to-face classroom	[54]
	AU, HER, HA, AR, MP	Oral presentations	[102]
	FE, HER, AR, LE, MP	Dancing skills	[115]
	FA, EDA, HR	-	[170]
	AU, MP, BL, LOG	Oral presentations	[171]
	EM, EEG	Attention, cognition	[172]
	-	Dancing skills	[173]
	AU, BL, MP, LOG	-	[174]
	EM, EEG	Adaptive self-assessment activity	[175]
	AU, VB, LOG	-	[176]
	EM, LOG	Open-ended learning environments	[94]
	BL, EC, AU, LOG	Oral presentations	[177]
MP, FE, AU	Oral presentations	[178]	
Mutual Verification between multimodal data	VO, FE, EDA	Collaboration, emotion	[107,128]
	BL, EDA, EM, AU, BVP, IBI, EDA, HR	Collaboration	[101,106,122,179–182]
	LOG, SR, AU	Online learning	[3,22]
	FR, EC	Embodied learning behavior	[116,183]
	PS, EM, LOG	Calligraphy training	[129,184]
	PS, GSR, ST, LOG	Online learning problem solving	[89,185]
	BL, MP, AU	Collaboration	[127]
	HER, AR	Language learning	[92]
	EDA, ECG	Collaboration	[72]
	EEG, LOG	Cognition	[186]
	-	Collaboration	[187]
	MP, OR	Teaching behavior	[66]
	VB, ONLINE	Emotion	[188]
	EM, BL	Collaboration	[189]
	EDA, PS	-	[97]
	EC, MP	Collaboration	[123]
FE, EM, GSR	Learning performance	[70]	
EM, FA, LOG	Learning difficulties	[190]	
EM, LOG	Cognition	[112]	
EM, AU, LOG	Engagement, collaboration, learning performance	[191]	

According to the four types of multimodal data proposed in this paper, methods for data integration included cross-type data, such as the integration of digital data and physical data [146] and the integration of psychometric data and physiological data [97]. There was also a non-cross-type, such as internal data integration of physiological data types [71,104]. In terms of learning indicators, the current literature on data integration focused on a single indicator, such as learning

engagement [83,141], as well as on multiple indicators, such as collaboration, engagement, and learning performance [78,130,138]. From the perspective of the relationships between data integration and learning indicators, data integration can be divided roughly into three categories, as shown in Figure 4: (1) “many to one” (multimodal data vs. learning indicator, for improving the accuracy of measurement), (2) “many to many” (multimodal data vs. multiple learning indicators, for improving information richness), and (3) mutual verification between multimodal data (providing empirical evidence for data fusion and integration). Further, the meaning of data integration in the literature had a broad sense and narrow sense. In a broad sense, the results of experiments on multimodal data were better than on single-mode data. The added value of data integration lies in improving measurement accuracy and information richness, or bringing more meaningful conclusions. In a narrow sense, only “many to one” can achieve data integration.

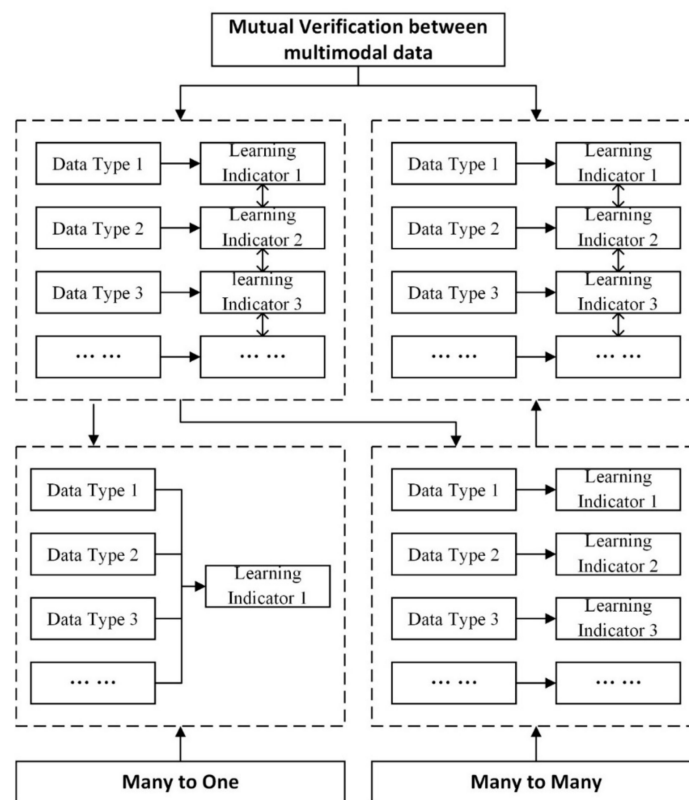


Figure 4. Data integration methods.

## 5.1. Integration Methods

### 5.1.1. “Many-to-One” (Improving Measurement Accuracy)

The characteristics of this category are as follows: (1) There is a clear data integration algorithm model. Multimodal data is usually used as the model input, while one learning indicator is the model output; (2) data integration improves the accuracy of learning indicator measurement. For example, audio data measure emotions [121], and facial expression data also measure emotions [51]. Audio and facial expression data was integrated by [79] to measure emotions and improve the accuracy of emotion measurement. In this line of research, the increase of data mode, the selection of data features, the division of data integration proportion, and the selection of the algorithm will affect the accuracy of the measurement. Some studies have compared single-mode data with multimodal data. The results showed that the measurement from multimodal data integration is more accurate than a single type of data [121,137]. Selecting the features from the raw data that are relevant to learning can increase their interpretability. Some studies just make use of the raw data [68]. In most studies, the data integration

ratio is 1:1. As mentioned before, different types of data have different accuracies in measuring the same learning indicator. For example, the use of EM and EEG results in different accuracies in predicting emotion [111]. Therefore, data integration is not as simple as one-to-one mapping. Based on the possible measurement accuracy of various types of data, and the correlations between data and learning indicators, the weights of data types used in the experiments should be allocated accordingly. Finding an efficient algorithm model is key [68]. Machine learning is widely used as algorithm models. Most studies compare the performance of several different algorithm models to determine the optimal algorithm. For example, deep learning methods are compared with traditional machine learning in terms of their performance [37].

#### 5.1.2. “Many-to-Many” (Improving Information Richness)

The characteristics of this type are described as follows: (1) There are more than two multi-dimensional learning indicators; (2) data and learning indicators are the one-to-one mapping; (3) there are no data integration algorithms, and; (4) data integration improves information richness. For example, EM for measuring attention and EEG for measuring cognition are used simultaneously [172]. Multi-dimensional learning indicators can accurately reflect the learning process. So, this line of research needs multiple learning indicators and obtains multimodal data that are suitable for measuring these indicators with the help of data integration systems. Integrated systems include the Oral Presentation Training System [155–157], the Sensor-Based Calligraphy Trainer [129], the Medical Operation Training System [103,167], the Ubiquitous Learning Analysis System [168,169], the Classroom Behavior Monitoring System [54,192], and the Dance Training System [115]. Some studies also use one type of data to measure several learning indicators simultaneously. For example, three learning indicators are measured simultaneously by EM (attention, anticipation, fatigue). EEG data are used to measure three learning indicators (load, mental workload, load on memory) [175]. However, this study does not advocate using only multimodal data to measure multiple indicators simultaneously. The overuse of one type of data will also reduce the accuracy of the measurement result to a certain extent. It is necessary to use the most suitable type of data to measure the most suitable learning indicators.

#### 5.1.3. Multimodal Data Validation (Provides Empirical Evidence for Data Fusion)

The objective of this type of MMLA is to increase confidence in the findings by using multimodal data validation, which is also called triangulation. In other words, this type of MMLA produces reliable conclusions through the triangular evidence of multimodal data analysis. Different types of data in an experiment are independent and parallel. Each measures the same learning indicator with different measurement accuracies. Through comparative analysis, we can take the measurement advantages of single-mode data and provide multiple validations for data integration of “many-to-one” and “many-to-many”. For example, [122] first collected multimodal data for collaborative learning analytics. Each type of data measurement collaboration is then analyzed separately, such as audio data [180], body posture [106], movement data [181], and physiological data [101,182]. As another example, self-report data and eye movement data on learning engagement are analyzed [129]. Additionally, some research focuses on the analysis of the relationships between various types of data [128]. The typical questions are what is the relationship between physiological arousal and learning interaction in collaborative learning, when does physiological arousal occur, and how do students’ emotions change [107,128]. Studies by [116,183] used gestures to analyze movement patterns, eye movements to analyze attention patterns, and they analyzed the correlation between movement patterns and attention patterns. Self-report and physiological data are used to measure the cognitive load by calculating the correlation between the two measurements [89,185].

#### 5.1.4. Other Integration Methods

The above three are common data integration methods. With more research on MMLA, methods of data integration will be more diverse in the future. For example, from the perspective of learning

process analysis, according to different research questions, selecting different types of data for analysis at different analysis stages is also an idea of data integration. For instance, this multi-step approach uses coarse-grain temporality—learning trajectories across knowledge components—to identify and further explore “focal” moments worthy of more fine-grain, context-rich analysis [22]. As another example, the log data is first used to analyze the overall learning path, and the macro learning path rules are found. Then, the eye movement data is used to analyze the two key learning stages of watching the video and doing testing, to deeply mine the learners’ cognitive preferences.

5.2. Summary of the Key Stages of Data Integration and Research Directions

The collection of different fine-grained and multimodal data is the premise of data integration. Within a data synchronous acquisition system, multimodal data on the learning process can be collected at the same time. A data integration system often consists of multiple modules, such as an expression analysis module [83,116,141,183], a VR module [156], a body posture module [142,193], and a self-reflection module [157]. The multimodal data is collected separately first and then co-located by using their timestamps. As a set of tools, STREAMS (Structured TRansactional Event Analysis of Multimodal Streams) integrates log data into multimodal data streams for analyses [22], for example. Therefore, temporal alignment is one of the key steps in data integration.

Data integration analysis is a crucial step in MMLA. Data from different sources are often collected at different times with varying grain sizes. It is highly time-consuming [22] to integrate them. For example, some studies have used data integration acquisition systems, such as presentation training systems, named the Presentation Trainer [155–157], but only select a single type of data for analysis. Therefore, these studies are not involved in data integration.

We summarize MMLA in Figure 5, in which the X-axis represents the multimodality of data, the Y-axis the methods used, and the Z-axis the multidimensional indicators. The existing methods for data integration improve either the accuracy of measure (point A) or the richness of learning information (point B). Ideal data integration should consider the intersection of the X-axis, Y-axis, and Z-axis, such as the C-point. That is, data integration should improve measurement accuracy and information richness, capture the states of learners over their learning, and characterize all aspects of the learning process. For example, the use of eye movement data and log data is measured for cognition, facial expression data for emotions, interview data for metacognition, and self-report data for motivation [94]. In other words, we should make the best use of multimodal data by taking advantage of the individual strengths of its components.

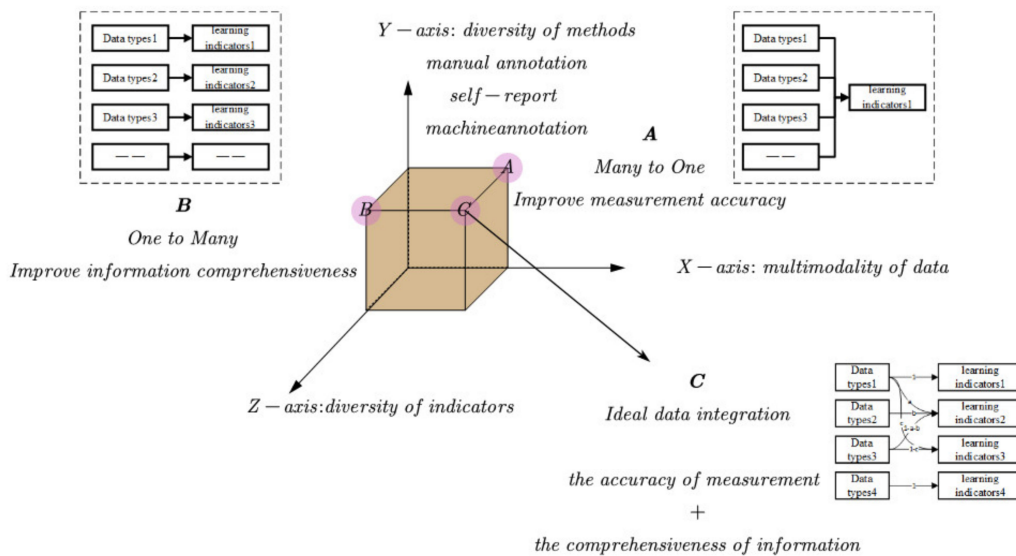


Figure 5. Three-dimensional features of data integration in MMLA.

MMLA focuses on what types of data are collected and how to integrate them in a way that the learning process can be characterized accurately. Three factors have contributed to the rapid development of MMLA. First, from the data perspective, the availability of different kinds of perceptual devices that are capable of collecting rich learning data promotes MMLA. Second, from the indicator perspective, the educational inquiry into the mechanism of learning and psychological factors in learning motivates MMLA. Finally, from the method level, the recent advancement of artificial intelligence enables MMLA.

The use of multimodal data does not mean data integration. Seamless, effective integration of multimodal data for accurately measuring the effectiveness of teaching and learning is an important future research direction of MMLA. Specifically, we believe there are two directions at methodological and practical levels.

At the methodological level, the research indirections in MMLA may lie in answering the following question. (1) For measuring a given learning indicator, which type of data is well suitable? Some findings on this have already been reported in the literature. However, there is no comprehensive research on answering this question by comparing different types of data against the measurement of learning indicators. (2) How to align multimodal data in a way that the learning indicators can be well-reflected. By experiments, different approaches should be compared in terms of their capacities of capturing hidden correlation information among the data. The complementary information from the different types of data should be exploited by using the similarity-based alignment, for example. (3) For qualifying a given learning indicator, how to fuse the multimodal data so that the complementary correlations from both the intermodality and cross-modality are effectively integrated. A combination of different ways of fusions and the degree of contributions from different types of data to the final performance should be examined.

At the practical level, we should consider how to select the use of multimodal data, data integration degree, and learning indicators, based on the research results. This will involve multi-disciplines such as data science, computer science, and educational technology. The collection and storage of high-frequency, fine-grained, and micro-level multimodal data should be part of a multimodal data education system. Guidance on how to effectively use multimodal data in learning and teaching for educators is a research direction.

In the future, we believe that MMLA will be available in classrooms in real-time.

## 6. Conclusions

As more and more data on learning processes become available, MMLA is becoming increasingly important. This paper has conducted a systematic review of the literature on MMLA published in the past three years. Specifically, we have presented a novel conceptual model for better understanding and classifying multimodal data, learning indicators, and their relationships. We classified the types of multimodal data in MMLA into digital data, physical data, physiological data, psychometric data, and environment data. The learning indicators were grouped as behavior, cognition, emotion, collaboration, and engagement. The relationships between multimodal data and learning indicators were one-to-one, one-to-many, and many-to-one. The complex relationships between multimodal data and learning indicators were the key to data fusion. We summarized the integration methods of multimodal data as many-to-one (improving measurement accuracy), many-to-many (improving information richness), and multimodal data validation (providing empirical evidence for data fusion and integration). Data integration in MMLA has been characterized by three aspects: multimodality of data, multi-dimension of indicators, and diversity of methods. This review highlights that the temporal alignment of multimodal data is the key step in data fusion. A description of three-dimensional characteristics in MMLA was presented and we pointed out the future direction of data fusion in MMLA.

**Author Contributions:** S.M. proposed the research topic; S.M. and X.H. discussed the research contents and method; M.C. completed the draft of the article; S.M. and X.H. revised the draft of the article; M.C. edited the final article according to the journal submission requirements. S.M. provided funding support. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China named Research on Eye Movement Big Data Based Students' Learning State Profiling and Its Applications, grant number 61907009.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Siemens, G.; Baker, R.S.J.d. Learning analytics and educational data mining: Towards communication and collaboration. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK 2012), New York, NY, USA, 29 April–2 May 2012; pp. 252–254.
2. Schwendimann, B.A.; Rodríguez-Triana, M.J.; Vozniuk, A.; Prieto, L.P.; Boroujeni, M.S.; Holzer, A.; Gillet, D.; Dillenbourg, P. Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. *IEEE Trans. Learn. Technol.* **2017**, *10*, 30–41. [\[CrossRef\]](#)
3. Liu, R.; Stamper, J.; Davenport, J.; Crossley, S.; McNamara, D.; Nzinga, K.; Sherin, B. Learning linkages: Integrating data streams of multiple modalities and timescales. *J. Comput. Assist. Learn.* **2019**, *35*, 99–109. [\[CrossRef\]](#)
4. Eradze, M.; Laanpere, M. Lesson Observation Data in Learning Analytics Datasets: Observata. In Proceedings of the 12th European Conference on Technology-Enhanced Learning (EC-TEL 2017), Tallinn, Estonia, 12–15 September 2017; pp. 504–508.
5. Rodríguez-Triana, M.J.; Prieto, L.P.; Vozniuk, A.; Boroujeni, M.S.; Schwendimann, B.A.; Holzer, A.; Gillet, D. Monitoring, awareness and reflection in blended technology enhanced learning: A systematic review. *Int. J. Technol. Enhanc. Learn.* **2017**, *9*, 126–150. [\[CrossRef\]](#)
6. Di Mitri, D. Digital Learning Projection. In *Artificial Intelligence in Education*; André, E., Baker, R., Hu, X., Rodrigo, M.M.T., du Boulay, B., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2017; Volume 10331, pp. 609–612.
7. Di Mitri, D.; Schneider, J.; Specht, M.; Drachler, H. From signals to knowledge: A conceptual model for multimodal learning analytics. *J. Comput. Assist. Learn.* **2018**, *34*, 338–349. [\[CrossRef\]](#)
8. Ochoa, X.; Worsley, M. Editorial: Augmenting Learning Analytics with Multimodal Sensory Data. *J. Learn. Anal.* **2016**, *3*, 213–219. [\[CrossRef\]](#)
9. Blikstein, P. Multimodal learning analytics. In Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium, 8–12 April 2013; pp. 102–106.
10. Spikol, D.; Prieto, L.P.; Rodríguez-Triana, M.J.; Worsley, M.; Ochoa, X.; Cukurova, M. Current and future multimodal learning analytics data challenges. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17), Vancouver, BC, Canada, 13–17 March 2017; pp. 518–519.
11. Cukurova, M. A Syllogism for Designing Collaborative Learning Technologies in the Age of AI and Multimodal Data. In *Proceedings of the Lifelong Technology-Enhanced Learning, Leeds, UK, 3–5 September 2018*; Pammer-Schindler, V., Pérez-Sanagustín, M., Drachler, H., Elferink, R., Scheffel, M., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 291–296.
12. Peffer, M.E. Combining Multimodal Learning Analytics with Backward Design to Assess Learning. In Proceedings of the 8th International Conference on Learning Analytics & Knowledge (LAK18), Sydney, Australia, 5–9 March 2018; pp. 1–5.
13. Prieto, L.P.; Rodríguez-Triana, M.J.; Martínez-Maldonado, R.; Dimitriadis, Y.; Gašević, D. Orchestrating learning analytics (OrLA): Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level. *Australas. J. Educ. Technol.* **2019**, *35*, 14–33. [\[CrossRef\]](#)
14. Haider, F.; Luz, S.; Campbell, N. Data Collection and Synchronisation: Towards a Multiperspective Multimodal Dialogue System with Metacognitive Abilities. In *Dialogues with Social Robots: Enablements, Analyses, and Evaluation*; Jokinen, K., Wilcock, G., Eds.; Lecture Notes in Electrical Engineering; Springer: Singapore, 2017; pp. 245–256. ISBN 978-981-10-2585-3.



15. Turker, A.; Dalsen, J.; Berland, M.; Steinkuehler, C. Challenges to Multimodal Data Set Collection in Games-based Learning Environments. In Proceedings of the Sixth Multimodal Learning Analytics (MMLA) Workshop, Vancouver, BC, Canada, 13–17 March 2017; pp. 1–7.
16. Chua, Y.H.V.; Rajalingam, P.; Tan, S.C.; Dauwels, J. EduBrowser: A Multimodal Automated Monitoring System for Co-located Collaborative Learning. In Proceedings of the Learning Technology for Education Challenges, Zamora, Spain, 15–18 July 2019; pp. 125–138.
17. Lahbi, Z.; Sabbane, M. U-Edu: Multimodal learning activities analytics model for learner feedback in ubiquitous education system. *Int. J. Adv. Trends Comput. Sci. Eng.* **2019**, *8*, 2551–2555. [[CrossRef](#)]
18. Shankar, S.K.; Rodríguez-Triana, M.J.; Ruiz-Calleja, A.; Prieto, L.P.; Chejara, P.; Martínez-Monés, A. Multimodal Data Value Chain (M-DVC): A Conceptual Tool to Support the Development of Multimodal Learning Analytics Solutions. *Revista Iberoamericana de Tecnologías del Aprendizaje* **2020**, *15*, 113–122. [[CrossRef](#)]
19. Bannert, M.; Molenar, I.; Azevedo, R.; Järvelä, S.; Gašević, D. Relevance of learning analytics to measure and support students' learning in adaptive educational technologies. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17), Vancouver, BC, Canada, 13–17 March 2017; pp. 568–569.
20. Martinez-Maldonado, R.; Kay, J.; Buckingham Shum, S.; Yacef, K. Collocated Collaboration Analytics: Principles and Dilemmas for Mining Multimodal Interaction Data. *Hum. Comput. Interact.* **2019**, *34*, 1–50. [[CrossRef](#)]
21. Samuelsen, J.; Chen, W.; Wasson, B. Integrating multiple data sources for learning analytics—Review of literature. *Res. Pract. Technol. Enhanc. Learn.* **2019**, *14*, 11. [[CrossRef](#)]
22. Liu, R.; Stamper, J.; Davenport, J. A novel method for the in-depth multimodal analysis of student learning trajectories in intelligent tutoring systems. *J. Learn. Anal.* **2018**, *5*, 41–54. [[CrossRef](#)]
23. Mitri, D.D.; Schneider, J.; Specht, M.; Drachsler, H. The Big Five: Addressing Recurrent Multimodal Learning Data Challenges. In Proceedings of the Companion Proceedings of the 8th International Conference on Learning Analytics and Knowledge: Towards User-Centred Learning Analytics, Sydney, Australia, 5–9 March 2018; pp. 420–424.
24. Oviatt, S. Ten Opportunities and Challenges for Advancing Student-Centered Multimodal Learning Analytics. In Proceedings of the 20th ACM International Conference on Multimodal Interaction, Boulder, CO, USA, 16–20 October 2018; pp. 87–94.
25. Worsley, M. Multimodal learning analytics' past, present, and, potential futures. In Proceedings of the 8th International Conference on Learning Analytics & Knowledge (LAK18), Sydney, Australia, 5–9 March 2018.
26. Shankar, S.K.; Prieto, L.P.; Rodríguez-Triana, M.J.; Ruiz-Calleja, A. A Review of Multimodal Learning Analytics Architectures. In Proceedings of the 2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT), Mumbai, India, 9–13 July 2018; pp. 212–214.
27. Crescenzi-Lanna, L. Multimodal Learning Analytics research with young children: A systematic review. *Br. J. Educ. Technol.* **2020**, *51*, 1485–1504. [[CrossRef](#)]
28. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Group, T.P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med.* **2009**, *6*, e1000097. [[CrossRef](#)] [[PubMed](#)]
29. Bond, M. Facilitating student engagement through the flipped learning approach in K-12: A systematic review. *Comput. Educ.* **2020**, *151*, 103819. [[CrossRef](#)]
30. Crompton, H.; Burke, D. Mobile learning and pedagogical opportunities: A configurative systematic review of PreK-12 research using the SAMR framework. *Comput. Educ.* **2020**, *156*, 103945. [[CrossRef](#)]
31. Diacopoulos, M.M.; Crompton, H. A systematic review of mobile learning in social studies. *Comput. Educ.* **2020**, *154*, 103911. [[CrossRef](#)]
32. Hooshyar, D.; Pedaste, M.; Saks, K.; Leijen, Ä.; Bardone, E.; Wang, M. Open learner models in supporting self-regulated learning in higher education: A systematic literature review. *Comput. Educ.* **2020**, *154*, 103878. [[CrossRef](#)]
33. Papadopoulou, I.; Lazzarino, R.; Miah, S.; Weaver, T.; Thomas, B.; Koulouglioti, C. A systematic review of the literature regarding socially assistive robots in pre-tertiary education. *Comput. Educ.* **2020**, *155*, 103924. [[CrossRef](#)]
34. Eradze, M.; Rodríguez Triana, M.J.; Laanpere, M. How to Aggregate Lesson Observation Data into Learning Analytics Datasets? Available online: <https://infoscience.epfl.ch/record/229372> (accessed on 17 March 2020).
35. Eradze, M.; Rodríguez-Triana, M.J.; Laanpere, M. Semantically Annotated Lesson observation Data in Learning Analytics Datasets: A Reference Model. *Interact. Des. Archit. J.* **2017**, *33*, 91–95.

36. Henrie, C.R.; Bodily, R.; Larsen, R.; Graham, C.R. Exploring the potential of LMS log data as a proxy measure of student engagement. *J. Comput. High. Educ.* **2018**, *30*, 344–362. [[CrossRef](#)]
37. Spikol, D.; Ruffaldi, E.; Dabisias, G.; Cukurova, M. Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *J. Comput. Assist. Learn.* **2018**, *34*, 366–377. [[CrossRef](#)]
38. Okur, E.; Alyuz, N.; Aslan, S.; Genc, U.; Tanriover, C.; Arslan Esme, A. Behavioral Engagement Detection of Students in the Wild. In *Proceedings of the Artificial Intelligence in Education, Wuhan, China, 28 June–1 July 2017*; André, E., Baker, R., Hu, X., Rodrigo, M.M.T., du Boulay, B., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 250–261.
39. Su, Y.-S.; Ding, T.-J.; Lai, C.-F. Analysis of Students Engagement and Learning Performance in a Social Community Supported Computer Programming Course. *Eurasia J. Math. Sci. Technol. Ed.* **2017**, *13*, 6189–6201. [[CrossRef](#)]
40. Suero Montero, C.; Suhonen, J. Emotion analysis meets learning analytics: Online learner profiling beyond numerical data. In *Proceedings of the 14th Koli Calling International Conference on Computing Education Research, Koli, Finland, 20–23 November 2014*; pp. 165–169.
41. Oviatt, S.; Hang, K.; Zhou, J.; Yu, K.; Chen, F. Dynamic Handwriting Signal Features Predict Domain Expertise. *ACM Trans. Interact. Intell. Syst.* **2018**, *8*, 1–21. [[CrossRef](#)]
42. Loup-Escande, E.; Frenoy, R.; Poplimont, G.; Thouvenin, I.; Gapenne, O.; Megalakaki, O. Contributions of mixed reality in a calligraphy learning task: Effects of supplementary visual feedback and expertise on cognitive load, user experience and gestural performance. *Comput. Hum. Behav.* **2017**, *75*, 42–49. [[CrossRef](#)]
43. Hsiao, I.-H.; Huang, P.-K.; Murphy, H. Integrating Programming Learning Analytics Across Physical and Digital Space. *IEEE Trans. Emerg. Top. Comput.* **2020**, *8*, 206–217. [[CrossRef](#)]
44. Paredes, Y.V.; Azcona, D.; Hsiao, I.-H.; Smeaton, A. Learning by Reviewing Paper-Based Programming Assessments. In *Proceedings of the Lifelong Technology-Enhanced Learning*; Pammer-Schindler, V., Pérez-Sanagustín, M., Drachslar, H., Elferink, R., Scheffel, M., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 510–523.
45. Sharma, K.; Dillenbourg, P.; Giannakos, M. Stimuli-Based Gaze Analytics to Enhance Motivation and Learning in MOOCs. In *Proceedings of the 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT), Macei, Brazil, 15–18 July 2019*; pp. 199–203.
46. D’Angelo, S.; Begel, A. Improving Communication Between Pair Programmers Using Shared Gaze Awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017*; pp. 6245–6290.
47. Schneider, B.; Sharma, K.; Cuendet, S.; Zufferey, G.; Dillenbourg, P.; Pea, R. Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *Int. J. Comput.-Support. Collab. Learn.* **2018**, *13*, 241–261. [[CrossRef](#)]
48. Ding, Y.; Zhang, Y.; Xiao, M.; Deng, Z. A Multifaceted Study on Eye Contact based Speaker Identification in Three-party Conversations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017*; pp. 3011–3021.
49. Noel, R.; Riquelme, F.; Lean, R.M.; Merino, E.; Cechinel, C.; Barcelos, T.S.; Villarroel, R.; Muñoz, R. Exploring Collaborative Writing of User Stories with Multimodal Learning Analytics: A Case Study on a Software Engineering Course. *IEEE Access* **2018**. [[CrossRef](#)]
50. Paans, C.; Molenaar, I.; Segers, E.; Verhoeven, L. Temporal variation in children’s self-regulated hypermedia learning. *Comput. Hum. Behav.* **2019**, *96*, 246–258. [[CrossRef](#)]
51. Martin, K.; Wang, E.Q.; Bain, C.; Worsley, M. Computationally Augmented Ethnography: Emotion Tracking and Learning in Museum Games. In *Proceedings of the Advances in Quantitative Ethnography, Madison, WI, USA, 20–22 October 2019*; pp. 141–153.
52. Monkaresi, H.; Bosch, N.; Calvo, R.A.; D’Mello, S.K. Automated Detection of Engagement Using Video-Based Estimation of Facial Expressions and Heart Rate. *IEEE Trans. Affect. Comput.* **2017**, *8*, 15–28. [[CrossRef](#)]
53. Watanabe, E.; Ozeki, T.; Kohama, T. Analysis of interactions between lecturers and students. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge-LAK ’18, Sydney, Australia, 5–9 March 2018*; pp. 370–374.
54. Ngoc Anh, B.; Tung Son, N.; Truong Lam, P.; Phuong Chi, L.; Huu Tuan, N.; Cong Dat, N.; Huu Trung, N.; Umar Aftab, M.; Van Dinh, T. A Computer-Vision Based Application for Student Behavior Monitoring in Classroom. *Appl. Sci.* **2019**, *9*, 4729. [[CrossRef](#)]

55. Abdelrahman, Y.; Velloso, E.; Dingler, T.; Schmidt, A.; Vetere, F. Cognitive Heat: Exploring the Usage of Thermal Imaging to Unobtrusively Estimate Cognitive Load. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2017**, *1*, 33:1–33:20. [[CrossRef](#)]
56. Cukurova, M.; Zhou, Q.; Spikol, D.; Landolfi, L. Modelling collaborative problem-solving competence with transparent learning analytics: Is video data enough? In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, Frankfurt, Germany, 23–27 March 2020; pp. 270–275.
57. Asadipour, A.; Debattista, K.; Chalmers, A. Visuohaptic augmented feedback for enhancing motor skills acquisition. *Vis. Comput.* **2017**, *33*, 401–411. [[CrossRef](#)]
58. Ou, L.; Andrade, A.; Alberto, R.; van Helden, G.; Bakker, A. Using a cluster-based regime-switching dynamic model to understand embodied mathematical learning. In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, Frankfurt, Germany, 23–27 March 2020; pp. 496–501.
59. Sriramulu, A.; Lin, J.; Oviatt, S. Dynamic Adaptive Gesturing Predicts Domain Expertise in Mathematics. In Proceedings of the 2019 International Conference on Multimodal Interaction (ICMI' 19), Suzhou, China, 14–18 October 2019; pp. 105–113.
60. Rosen, D.; Palatnik, A.; Abrahamson, D. A Better Story: An Embodied-Design Argument for Generic Manipulatives. In *Using Mobile Technologies in the Teaching and Learning of Mathematics*; Calder, N., Larkin, K., Sinclair, N., Eds.; Mathematics Education in the Digital Era; Springer International Publishing: Cham, Switzerland, 2018; pp. 189–211.
61. Di Mitri, D.; Scheffel, M.; Drachsler, H.; Börner, D.; Ternier, S.; Specht, M. Learning pulse: A machine learning approach for predicting performance in self-regulated learning using multimodal data. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17), Vancouver, BC, Canada, 13–17 March 2017; pp. 188–197.
62. Junokas, M.J.; Lindgren, R.; Kang, J.; Morphew, J.W. Enhancing multimodal learning through personalized gesture recognition. *J. Comput. Assist. Learn.* **2018**, *34*, 350–357. [[CrossRef](#)]
63. Ibrahim-Didi, K.; Hackling, M.W.; Ramseger, J.; Sherriff, B. Embodied Strategies in the Teaching and Learning of Science. In *Quality Teaching in Primary Science Education: Cross-Cultural Perspectives*; Hackling, M.W., Ramseger, J., Chen, H.-L.S., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 181–221. ISBN 978-3-319-44383-6.
64. Martinez-Maldonado, R. "I Spent More Time with that Team": Making Spatial Pedagogy Visible Using Positioning Sensors. In Proceedings of the 9th International Conference on Learning Analytics & Knowledge-LAK19, Tempe, AZ, USA, 4–8 March 2019; pp. 21–25.
65. Healion, D.; Russell, S.; Cukurova, M.; Spikol, D. Tracing physical movement during practice-based learning through multimodal learning analytics. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17), Vancouver, BC, Canada, 13–17 March 2017; pp. 588–589.
66. An, P.; Bakker, S.; Ordanovski, S.; Paffen, C.L.E.; Taconis, R.; Eggen, B. Dandelion Diagram: Aggregating Positioning and Orientation Data in the Visualization of Classroom Proxemics. In Proceedings of the Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–8.
67. Mills, C.; Fridman, I.; Soussou, W.; Waghray, D.; Olney, A.M.; D'Mello, S.K. Put your thinking cap on: Detecting cognitive load using EEG during learning. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17), Vancouver, BC, Canada, 13–17 March 2017; pp. 80–89.
68. Tzirakis, P.; Trigeorgis, G.; Nicolaou, M.A.; Schuller, B.W.; Zafeiriou, S. End-to-End Multimodal Emotion Recognition Using Deep Neural Networks. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 1301–1309. [[CrossRef](#)]
69. Pijeira-Díaz, H.J.; Drachsler, H.; Kirschner, P.A.; Järvelä, S. Profiling sympathetic arousal in a physics course: How active are students? *J. Comput. Assist. Learn.* **2018**, *34*, 397–408. [[CrossRef](#)]
70. Edwards, A.A.; Massicci, A.; Sridharan, S.; Geigel, J.; Wang, L.; Bailey, R.; Alm, C.O. Sensor-based Methodological Observations for Studying Online Learning. In Proceedings of the 2017 ACM Workshop on Intelligent Interfaces for Ubiquitous and Smart Learning, Limassol, Cyprus, 13 March 2017; pp. 25–30.
71. Yin, Z.; Zhao, M.; Wang, Y.; Yang, J.; Zhang, J. Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. *Comput. Methods Programs Biomed.* **2017**, *140*, 93–110. [[CrossRef](#)]
72. Ahonen, L.; Cowley, B.U.; Hellas, A.; Puolamäki, K. Biosignals reflect pair-dynamics in collaborative work: EDA and ECG study of pair-programming in a classroom environment. *Sci. Rep.* **2018**, *8*, 1–16. [[CrossRef](#)]

73. Pham, P.; Wang, J. AttentiveLearner2: A Multimodal Approach for Improving MOOC Learning on Mobile Devices. In *Artificial Intelligence in Education*; André, E., Baker, R., Hu, X., Rodrigo, M.M.T., du Boulay, B., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2017; Volume 10331, pp. 561–564.
74. Chejara, P.; Prieto, L.P.; Ruiz-Calleja, A.; Rodríguez-Triana, M.J.; Shankar, S.K. Exploring the Triangulation of Dimensionality Reduction When Interpreting Multimodal Learning Data from Authentic Settings. In *Proceedings of the Transforming Learning with Meaningful Technologies*; Scheffel, M., Broisin, J., Pammer-Schindler, V., Ioannou, A., Schneider, J., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 664–667.
75. Kim, J.; Meltzer, C.; Salehi, S.; Blikstein, P. Process Pad: A multimedia multi-touch learning platform. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '11)*, Kobe, Japan, 13–16 November 2011; pp. 272–273.
76. Cukurova, M.; Luckin, R.; Mavrikis, M.; Millán, E. Machine and Human Observable Differences in Groups' Collaborative Problem-Solving Behaviours. In *Proceedings of the Data Driven Approaches in Digital Education*, Tallinn, Estonia, 12–15 September 2017; pp. 17–29.
77. Spikol, D.; Avramides, K.; Cukurova, M.; Vogel, B.; Luckin, R.; Ruffaldi, E.; Mavrikis, M. Exploring the interplay between human and machine annotated multimodal learning analytics in hands-on STEM activities. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, Edinburgh, UK, 25–29 April 2016; pp. 522–523.
78. Worsley, M.A.B. Multimodal Learning Analytics for the Qualitative Researcher. In *Proceedings of the 2018 International Conference of the Learning Sciences*, London, UK, 23–27 June 2018; pp. 1109–1112.
79. Ez-zaouia, M.; Lavoué, E. EMODA: A tutor oriented multimodal and contextual emotional dashboard. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17)*, Vancouver, BC, Canada, 13–17 March 2017; pp. 429–438.
80. Martinez-Maldonado, R.; Echeverria, V.; Fernandez Nieto, G.; Buckingham Shum, S. 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*, Honolulu, HI, USA, 25–30 April 2020; pp. 1–15.
81. Praharaaj, S.; Scheffel, M.; Drachsler, H.; Specht, M. Multimodal Analytics for Real-Time Feedback in Co-located Collaboration. In *Proceedings of the Lifelong Technology-Enhanced Learning*, Leeds, UK, 3–5 September 2018; pp. 187–201.
82. Zaletelj, J.; Košir, A. Predicting students' attention in the classroom from Kinect facial and body features. *EURASIP J. Image Video Process.* **2017**, *2017*, 80. [[CrossRef](#)]
83. Thomas, C. Multimodal Teaching and Learning Analytics for Classroom and Online Educational Settings. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, Boulder, CO, USA, 20–26 October 2018; pp. 542–545.
84. Sommer, S.; Hinojosa, L.; Traut, H.; Polman, J.; Weidler-Lewis, J. Integrating Eye-Tracking Activities Into a Learning Environment to Promote Collaborative Meta-Semiotic Reflection and Discourse. In *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning (CSCL) 2017*, Philadelphia, PA, USA, 11–12 February 2017; pp. 1–4.
85. Martinez-Maldonado, R.; Echeverria, V.; Santos, O.C.; Santos, A.D.P.D.; Yacef, K. Physical learning analytics: A multimodal perspective. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK'18)*, Sydney, Australia, 5–9 March 2018; pp. 375–379.
86. Spikol, D.; Ruffaldi, E.; Cukurova, M. Using Multimodal Learning Analytics to Identify Aspects of Collaboration in Project-Based Learning. In *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning*, Philadelphia, PA, USA, 18–22 June 2017; pp. 263–270.
87. Duijzer, C.A.C.G.; Shayan, S.; Bakker, A.; Van der Schaaf, M.F.; Abrahamson, D. Touchscreen Tablets: Coordinating Action and Perception for Mathematical Cognition. *Front. Psychol.* **2017**, *8*. [[CrossRef](#)]
88. Paredes, Y.V.; Hsiao, I.; Lin, Y. Personalized guidance on how to review paper-based assessments. In *Proceedings of the 26th International Conference on Computers in Education*, Main Conference Proceedings, Manila, Philippines, 26–30 November 2018; pp. 257–265.
89. Larmuseau, C.; Vanneste, P.; Desmet, P.; Depaepe, F. Multichannel data for understanding cognitive affordances during complex problem solving. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK '19)*, Tempe, AZ, USA, 4–8 March 2019; pp. 61–70.

90. Sonnenberg, C.; Bannert, M. Using Process Mining to examine the sustainability of instructional support: How stable are the effects of metacognitive prompting on self-regulatory behavior? *Comput. Hum. Behav.* **2019**, *96*, 259–272. [[CrossRef](#)]
91. Cukurova, M.; Luckin, R.; Millán, E.; Mavrikis, M. The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. *Comput. Educ.* **2018**, *116*, 93–109. [[CrossRef](#)]
92. Gorham, T.; Jubaed, S.; Sanyal, T.; Starr, E.L. Assessing the efficacy of VR for foreign language learning using multimodal learning analytics. In *Professional Development in CALL: A Selection of Papers*; Research-Publishing.Net: Voillans, France, 2019; pp. 101–116.
93. Sun, B.; Lai, S.; Xu, C.; Xiao, R.; Wei, Y.; Xiao, Y. Differences of online learning behaviors and eye-movement between students having different personality traits. In Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education, Glasgow, UK, 13 November 2017; pp. 71–75.
94. Munshi, A.; Biswas, G. Personalization in OELs: Developing a Data-Driven Framework to Model and Scaffold SRL Processes. In *Artificial Intelligence in Education*; Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., Luckin, R., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2019; Volume 11626, pp. 354–358.
95. Andrade, A.; Maddox, B.; Edwards, D.; Chopade, P.; Khan, S. Quantitative Multimodal Interaction Analysis for the Assessment of Problem-Solving Skills in a Collaborative Online Game. In Proceedings of the Advances in Quantitative Ethnography, Madison, WI, USA, 20–22 October 2019; pp. 281–290.
96. Aslan, S.; Alyuz, N.; Tanriover, C.; Mete, S.E.; Okur, E.; D'Mello, S.K.; Arslan Esme, A. 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 (CHI '19), Glasgow, UK, 4–9 May 2019; pp. 1–12.
97. Dindar, M.; Malmberg, J.; Järvelä, S.; Haataja, E.; Kirschner, P.A. Matching self-reports with electrodermal activity data: Investigating temporal changes in self-regulated learning. *Educ. Inf. Technol.* **2020**, *25*, 1785–1802. [[CrossRef](#)]
98. Giannakos, M.N.; Sharma, K.; Pappas, I.O.; Kostakos, V.; Velloso, E. Multimodal data as a means to understand the learning experience. *Int. J. Inf. Manag.* **2019**, *48*, 108–119. [[CrossRef](#)]
99. Burnik, U.; Zaletelj, J.; Košir, A. Video-based learners' observed attention estimates for lecture learning gain evaluation. *Multimed. Tools Appl.* **2018**, *77*, 16903–16926. [[CrossRef](#)]
100. Spikol, D.; Ruffaldi, E.; Landolfi, L.; Cukurova, M. Estimation of Success in Collaborative Learning Based on Multimodal Learning Analytics Features. In Proceedings of the 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), Timisoara, Romania, 3–7 July 2017; pp. 269–273.
101. Dich, Y.; Reilly, J.; Schneider, B. Using Physiological Synchrony as an Indicator of Collaboration Quality, Task Performance and Learning. In *Artificial Intelligence in Education*; Penstein Rosé, C., Martínez-Maldonado, R., Hoppe, H.U., Luckin, R., Mavrikis, M., Porayska-Pomsta, K., McLaren, B., du Boulay, B., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2018; Volume 10947, pp. 98–110.
102. Gan, T.; Li, J.; Wong, Y.; Kankanhalli, M.S. A Multi-sensor Framework for Personal Presentation Analytics. *ACM Trans. Multimed. Comput. Commun. Appl.* **2019**, *15*, 1–21. [[CrossRef](#)]
103. Di Mitri, D.; Schneider, J.; Specht, M.; Drachslar, H. Detecting Mistakes in CPR Training with Multimodal Data and Neural Networks. *Sensors* **2019**, *19*, 3099. [[CrossRef](#)] [[PubMed](#)]
104. Nourbakhsh, N.; Chen, F.; Wang, Y.; Calvo, R.A. Detecting Users' Cognitive Load by Galvanic Skin Response with Affective Interference. *ACM Trans. Interact. Intell. Syst.* **2017**, *7*, 12:1–12:20. [[CrossRef](#)]
105. Kim, Y.; Butail, S.; Tscholl, M.; Liu, L.; Wang, Y. An exploratory approach to measuring collaborative engagement in child robot interaction. In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, Frankfurt, Germany, 23–27 March 2020; pp. 209–217.
106. Reilly, J.M.; Ravenell, M.; Schneider, B. Exploring Collaboration Using Motion Sensors and Multi-Modal Learning Analytics. In Proceedings of the International Educational Data Mining (EDM), Raleigh, NC, USA, 16–20 July 2018; pp. 1–7.
107. Malmberg, J.; Järvelä, S.; Holappa, J.; Haataja, E.; Huang, X.; Siipo, A. Going beyond what is visible: What multichannel data can reveal about interaction in the context of collaborative learning? *Comput. Hum. Behav.* **2019**, *96*, 235–245. [[CrossRef](#)]

108. Hutt, S.; Mills, C.; Bosch, N.; Krasich, K.; Brockmole, J.; D’Mello, S. “Out of the Fr-Eye-ing Pan”: Towards Gaze-Based Models of Attention during Learning with Technology in the Classroom. In Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization, Bratislava, Slovakia, 9–12 July 2017; pp. 94–103.
109. Mu, S.; Cui, M.; Wang, X.J.; Qiao, J.X.; Tang, D.M. Learners’ attention preferences of information in online learning: An empirical study based on eye-tracking. *Interact. Technol. Smart Educ.* **2019**, *16*, 186–203. [[CrossRef](#)]
110. Sharma, K.; Papamitsiou, Z.; Giannakos, M. Building pipelines for educational data using AI and multimodal analytics: A “grey-box” approach. *Br. J. Educ. Technol.* **2019**, *50*, 3004–3031. [[CrossRef](#)]
111. Zheng, W.-L.; Liu, W.; Lu, Y.; Lu, B.-L.; Cichocki, A. EmotionMeter: A Multimodal Framework for Recognizing Human Emotions. *IEEE Trans. Cybern.* **2019**, *49*, 1110–1122. [[CrossRef](#)] [[PubMed](#)]
112. Taub, M.; Mudrick, N.V.; Azevedo, R.; Millar, G.C.; Rowe, J.; Lester, J. Using multi-channel data with multi-level modeling to assess in-game performance during gameplay with Crystal Island. *Comput. Hum. Behav.* **2017**, *76*, 641–655. [[CrossRef](#)]
113. Viswanathan, S.A.; Van Lehn, K. High Accuracy Detection of Collaboration from Log Data and Superficial Speech Features. In Proceedings of the 12th International Conference on Computer Supported Collaborative Learning (CSCCL) 2017, Philadelphia, PA, USA, 18–22 June 2017; pp. 1–8.
114. Vrzakova, H.; Amon, M.J.; Stewart, A.; Duran, N.D.; D’Mello, S.K. Focused or stuck together: Multimodal patterns reveal triads’ performance in collaborative problem solving. In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, Frankfurt, Germany, 23–27 March 2020; pp. 295–304.
115. Romano, G.; Schneider, J.; Drachsler, H. Dancing Salsa with Machines—Filling the Gap of Dancing Learning Solutions. *Sensors* **2019**, *19*, 3661. [[CrossRef](#)] [[PubMed](#)]
116. Andrade, A.; Danish, J.; Maltest, A. A Measurement Model of Gestures in an Embodied Learning Environment: Accounting for Temporal Dependencies. *J. Learn. Anal.* **2017**, *4*, 18–45. [[CrossRef](#)]
117. Donnelly, P.J.; Blanchard, N.; Olney, A.M.; Kelly, S.; Nystrand, M.; D’Mello, S.K. Words matter: Automatic detection of teacher questions in live classroom discourse using linguistics, acoustics, and context. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Vancouver, BC, Canada, 13–17 March 2017; pp. 218–227.
118. Mudrick, N.V.; Azevedo, R.; Taub, M. Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning. *Comput. Hum. Behav.* **2019**, *96*, 223–234. [[CrossRef](#)]
119. Bosch, N.; Mills, C.; Wammes, J.D.; Smilek, D. Quantifying Classroom Instructor Dynamics with Computer Vision. In Proceedings of the Artificial Intelligence in Education; Penstein-Rosé, C., Martínez-Maldonado, R., Hoppe, H.U., Luckin, R., Mavrikis, M., Porayska-Pomsta, K., McLaren, B., du Boulay, B., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 30–42.
120. Schneider, B. A Methodology for Capturing Joint Visual Attention Using Mobile Eye-Trackers. *J. Vis. Exp. JoVE* **2020**. [[CrossRef](#)]
121. Cukurova, M.; Kent, C.; Luckin, R. Artificial intelligence and multimodal data in the service of human decision-making: A case study in debate tutoring. *Br. J. Educ. Technol.* **2019**, *50*, 3032–3046. [[CrossRef](#)]
122. Starr, E.L.; Reilly, J.M.; Schneider, B. Toward Using Multi-Modal Learning Analytics to Support and Measure Collaboration in Co-Located Dyads. In Proceedings of the 13th International Conference of the Learning Sciences (ICLS) 2018, London, UK, 23–27 June 2018; pp. 1–8.
123. Vujovic, M.; Tassani, S.; Hernández-Leo, D. Motion Capture as an Instrument in Multimodal Collaborative Learning Analytics. In Proceedings of the Transforming Learning with Meaningful Technologies, Delft, The Netherlands, 16–19 September 2019; pp. 604–608.
124. Cornide-Reyes, H.; Noël, R.; Riquelme, F.; Gajardo, M.; Cechinel, C.; Mac Lean, R.; Becerra, C.; Villarroel, R.; Munoz, R. Introducing Low-Cost Sensors into the Classroom Settings: Improving the Assessment in Agile Practices with Multimodal Learning Analytics. *Sensors* **2019**, *19*, 3291. [[CrossRef](#)]
125. Riquelme, F.; Munoz, R.; Mac Lean, R.; Villarroel, R.; Barcelos, T.S.; de Albuquerque, V.H.C. Using multimodal learning analytics to study collaboration on discussion groups. *Univers. Access Inf. Soc.* **2019**, *18*, 633–643. [[CrossRef](#)]
126. Sullivan, F.R.; Keith, P.K. Exploring the potential of natural language processing to support microgenetic analysis of collaborative learning discussions. *Br. J. Educ. Technol.* **2019**, *50*, 3047–3063. [[CrossRef](#)]

127. Davidsen, J.; Ryberg, T. “This is the size of one meter”: Children’s bodily-material collaboration. *Int. J. Comput.-Support. Collab. Learn.* **2017**, *12*, 65–90. [\[CrossRef\]](#)
128. Järvelä, S.; Malmberg, J.; Haataja, E.; Sobocinski, M.; Kirschner, P.A. What multimodal data can tell us about the students’ regulation of their learning process? *Learn. Instr.* **2019**, 101203. [\[CrossRef\]](#)
129. Limbu, B.H.; Jarodzka, H.; Klemke, R.; Specht, M. Can You Ink While You Blink? Assessing Mental Effort in a Sensor-Based Calligraphy Trainer. *Sensors* **2019**, *19*, 3244. [\[CrossRef\]](#) [\[PubMed\]](#)
130. Worsley, M. (Dis)engagement matters: Identifying efficacious learning practices with multimodal learning analytics. In Proceedings of the 8th International Conference on Learning Analytics and Knowledge-LAK’ 18, Sydney, Australia, 5–9 March 2018; pp. 365–369.
131. Furuichi, K.; Worsley, M. Using Physiological Responses To Capture Unique Idea Creation In Team Collaborations. In Proceedings of the Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW ’18), Jersey City, NJ, USA, 3–7 November 2018; pp. 369–372.
132. Beardsley, M.; Hernández-Leo, D.; Ramirez-Melendez, R. Seeking reproducibility: Assessing a multimodal study of the testing effect. *J. Comput. Assist. Learn.* **2018**, *34*, 378–386. [\[CrossRef\]](#)
133. Minematsu, T.; Tamura, K.; Shimada, A.; Konomi, S.; Taniguchi, R. Analytics of Reading Patterns Based on Eye Movements in an e-Learning System. In Proceedings of the Society for Information Technology & Teacher Education International Conference, Waynesville, NC, USA, 18 March 2019; pp. 1054–1059.
134. Pham, P.; Wang, J. Understanding Emotional Responses to Mobile Video Advertisements via Physiological Signal Sensing and Facial Expression Analysis. In Proceedings of the 22nd International Conference on Intelligent User Interfaces, Limassol, Cyprus, 13–16 March 2017; pp. 67–78.
135. Pham, P.; Wang, J. Predicting Learners’ Emotions in Mobile MOOC Learning via a Multimodal Intelligent Tutor. In Proceedings of the Intelligent Tutoring Systems; Nkambou, R., Azevedo, R., Vassileva, J., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 150–159.
136. Amon, M.J.; Vrzakova, H.; D’Mello, S.K. Beyond Dyadic Coordination: Multimodal Behavioral Irregularity in Triads Predicts Facets of Collaborative Problem Solving. *Cogn. Sci.* **2019**, *43*, e12787. [\[CrossRef\]](#)
137. Cukurova, M.; Kent, C.; Luckin, R. The Value of Multimodal Data in Classification of Social and Emotional Aspects of Tutoring. In *Artificial Intelligence in Education*; Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., Luckin, R., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2019; Volume 11626, pp. 46–51. ISBN 978-3-030-23206-1.
138. Worsley, M.; Blikstein, P. A Multimodal Analysis of Making. *Int. J. Artif. Intell. Educ.* **2018**, *28*, 385–419. [\[CrossRef\]](#)
139. Prieto, L.P.; Sharma, K.; Kidzinski, Ł.; Rodríguez-Triana, M.J.; Dillenbourg, P. Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. *J. Comput. Assist. Learn.* **2018**, *34*, 193–203. [\[CrossRef\]](#)
140. Prieto, L.P.; Sharma, K.; Kidzinski, Ł.; Dillenbourg, P. Orchestration Load Indicators and Patterns: In-the-Wild Studies Using Mobile Eye-Tracking. *IEEE Trans. Learn. Technol.* **2018**, *11*, 216–229. [\[CrossRef\]](#)
141. Thomas, C.; Jayagopi, D.B. Predicting student engagement in classrooms using facial behavioral cues. In Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education, Glasgow, UK, 13 November 2017; pp. 33–40.
142. Ashwin, T.S.; Guddeti, R.M.R. Unobtrusive Students’ Engagement Analysis in Computer Science Laboratory Using Deep Learning Techniques. In Proceedings of the 2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT), Mumbai, India, 9–13 July 2018; pp. 436–440.
143. Sharma, K.; Papamitsiou, Z.; Olsen, J.K.; Giannakos, M. Predicting learners’ effortful behaviour in adaptive assessment using multimodal data. In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, Frankfurt, Germany, 23–27 March 2020; pp. 480–489.
144. Viswanathan, S.A.; VanLehn, K. Using the Tablet Gestures and Speech of Pairs of Students to Classify Their Collaboration. *IEEE Trans. Learn. Technol.* **2018**, *11*, 230–242. [\[CrossRef\]](#)
145. Grawemeyer, B.; Mavrikis, M.; Holmes, W.; Gutiérrez-Santos, S.; Wiedmann, M.; Rummel, N. Affective learning: Improving engagement and enhancing learning with affect-aware feedback. *User Model. User-Adapt. Interact.* **2017**, *27*, 119–158. [\[CrossRef\]](#)
146. Alyuz, N.; Okur, E.; Genc, U.; Aslan, S.; Tanriover, C.; Esme, A.A. An unobtrusive and multimodal approach for behavioral engagement detection of students. In Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education, Glasgow, UK, 13 November 2017; pp. 26–32.

147. Hanani, A.; Al-Amleh, M.; Bazbus, W.; Salameh, S. Automatic Estimation of Presentation Skills Using Speech, Slides and Gestures. In Proceedings of the Speech and Computer, Hatfield, UK, 12–16 September 2017; pp. 182–191.
148. Fwa, H.L.; Marshall, L. Modeling engagement of programming students using unsupervised machine learning technique. *GSTF J. Comput.* **2018**. [[CrossRef](#)]
149. Larmuseau, C.; Cornelis, J.; Lancieri, L.; Desmet, P.; Depaepe, F. Multimodal learning analytics to investigate cognitive load during online problem solving. *Br. J. Educ. Technol.* **2020**, *51*, 1548–1562. [[CrossRef](#)]
150. Min, W.; Park, K.; Wiggins, J.; Mott, B.; Wiebe, E.; Boyer, K.E.; Lester, J. Predicting Dialogue Breakdown in Conversational Pedagogical Agents with Multimodal LSTMs. In *Artificial Intelligence in Education*; Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., Luckin, R., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2019; Volume 11626, pp. 195–200.
151. Nihei, F.; Nakano, Y.I.; Takase, Y. Predicting meeting extracts in group discussions using multimodal convolutional neural networks. In Proceedings of the 19th ACM International Conference on Multimodal Interaction, Glasgow, UK, 13–17 November 2017; pp. 421–425.
152. Kaur, A.; Mustafa, A.; Mehta, L.; Dhall, A. Prediction and Localization of Student Engagement in the Wild. In Proceedings of the 2018 Digital Image Computing: Techniques and Applications (DICTA), Canberra, Australia, 10–13 December 2018; pp. 1–8.
153. Mihoub, A.; Lefebvre, G. Social Intelligence Modeling using Wearable Devices. In Proceedings of the 22nd International Conference on Intelligent User Interfaces, Limassol, Cyprus, 13–16 March 2017; pp. 331–341.
154. Smith, C.; King, B.; Gonzalez, D. Using Multimodal Learning Analytics to Identify Patterns of Interactions in a Body-Based Mathematics Activity. *J. Interact. Learn. Res.* **2016**, *27*, 355–379.
155. Schneider, J.; Börner, D.; van Rosmalen, P.; Specht, M. Presentation Trainer: What experts and computers can tell about your nonverbal communication. *J. Comput. Assist. Learn.* **2017**, *33*, 164–177. [[CrossRef](#)]
156. Schneider, J.; Romano, G.; Drachsler, H. Beyond Reality—Extending a Presentation Trainer with an Immersive VR Module. *Sensors* **2019**, *19*, 3457. [[CrossRef](#)]
157. Schneider, J.; Börner, D.; van Rosmalen, P.; Specht, M. Do You Know What Your Nonverbal Behavior Communicates?—Studying a Self-reflection Module for the Presentation Trainer. In Proceedings of the Immersive Learning Research Network, Coimbra, Portugal, 26–29 June 2017; pp. 93–106.
158. Praharaaj, S. Co-located Collaboration Analytics. In Proceedings of the 2019 International Conference on Multimodal Interaction, Suzhou, China, 14–18 October 2019; pp. 473–476.
159. Praharaaj, S.; Scheffel, M.; Drachsler, H.; Specht, M. MULTIFOCUS: MULTImodal Learning Analytics For Co-located Collaboration Understanding and Support. In Proceedings of the European Conference on Technology Enhanced Learning, Leeds, UK, 3–6 September 2018; pp. 1–6.
160. Praharaaj, S.; Scheffel, M.; Drachsler, H.; Specht, M. Group Coach for Co-located Collaboration. In Proceedings of the Transforming Learning with Meaningful Technologies, Delft, The Netherlands, 16–19 September 2019; pp. 732–736.
161. Buckingham Shum, S.; Echeverria, V.; Martinez-Maldonado, R. The Multimodal Matrix as a Quantitative Ethnography Methodology. In Proceedings of the Advances in Quantitative Ethnography, Madison, WI, USA, 20–22 October 2019; pp. 26–40.
162. Echeverria, V.; Martinez-Maldonado, R.; Buckingham Shum, S. Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), Glasgow, Scotland, 4–9 May 2019; pp. 1–16.
163. Martinez-Maldonado, R.; Elliott, D.; Axisa, C.; Power, T.; Echeverria, V.; Shum, S.B. Designing translucent learning analytics with teachers: An elicitation process. *Interact. Learn. Environ.* **2020**, *36*, 1–15. [[CrossRef](#)]
164. Martinez-Maldonado, R.; Echeverria, V.; Elliott, D.; Axisa, C.; Power, T.; Shum, B. Making the Design of CSCL Analytics Interfaces a Co-design Process: The Case of Multimodal Teamwork in Healthcare. In Proceedings of the 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, Lyon, France, 15–19 July 2019; pp. 859–860.
165. Martinez-Maldonado, R.; Pechenizkiy, M.; Buckingham Shum, S.; Power, T.; Hayes, C.; Axisa, C. Modelling Embodied Mobility Teamwork Strategies in a Simulation-Based Healthcare Classroom. In Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization, Bratislava, Slovakia, 9–12 July 2017; pp. 308–312.



166. Martinez-Maldonado, R.; Power, T.; Hayes, C.; Abdiprano, A.; Vo, T.; Buckingham Shum, S. Analytics meet patient manikins: Challenges in an authentic small-group healthcare simulation classroom. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17), Vancouver, BC, Canada, 13–17 March 2017; pp. 90–94.
167. Di Mitri, D. Multimodal Tutor for CPR. In *Artificial Intelligence in Education*; Penstein Rosé, C., Martínez-Maldonado, R., Hoppe, H.U., Luckin, R., Mavrikis, M., Porayska-Pomsta, K., McLaren, B., du Boulay, B., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2018; Volume 10948, pp. 513–516. ISBN 978-3-319-93845-5.
168. Okada, M.; Kuroki, Y.; Tada, M. Multimodal analytics to understand self-regulation process of cognitive and behavioral strategies in real-world learning. *IEICE Trans. Inf. Syst.* **2020**, *E103D*, 1039–1054. [[CrossRef](#)]
169. Okada, M.; Kuroki, Y.; Tada, M. *Multimodal Method to Understand Real-world Learning Driven by Internal Strategies*; Association for the Advancement of Computing in Education (AACE): Waynesville, NC, USA, 2016; pp. 1248–1257.
170. Chikersal, P.; Tomprou, M.; Kim, Y.J.; Woolley, A.W.; Dabbish, L. Deep Structures of Collaboration: Physiological Correlates of Collective Intelligence and Group Satisfaction. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, Portland, OR, USA, 25 February–1 March 2017; pp. 873–888.
171. Van Ginkel, S.; Gulikers, J.; Biemans, H.; Noroozi, O.; Roozen, M.; Bos, T.; van Tilborg, R.; van Halteren, M.; Mulder, M. Fostering oral presentation competence through a virtual reality-based task for delivering feedback. *Comput. Educ.* **2019**, *134*, 78–97. [[CrossRef](#)]
172. Tamura, K.; Lu, M.; Konomi, S.; Hatano, K.; Inaba, M.; Oi, M.; Okamoto, T.; Okubo, F.; Shimada, A.; Wang, J.; et al. Integrating Multimodal Learning Analytics and Inclusive Learning Support Systems for People of All Ages. In Proceedings of the Cross-Cultural Design. Culture and Society, Orlando, FL, USA, 26–31 July 2019; pp. 469–481.
173. Dias Pereira dos Santos, A.; Yacef, K.; Martínez-Maldonado, R. Let's Dance: How to Build a User Model for Dance Students Using Wearable Technology. In Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization, Bratislava, Slovakia, 9–12 July 2017; pp. 183–191.
174. Prieto-Alvarez, C.G.; Martínez-Maldonado, R.; Shum, S.B. Mapping learner-data journeys: Evolution of a visual co-design tool. In Proceedings of the 30th Australian Conference on Computer-Human Interaction, Melbourne, Australia, 4–7 December 2018; pp. 205–214.
175. Sharma, K.; Papamitsiou, Z.; Giannakos, M.N. Modelling Learners' Behaviour: A Novel Approach Using GARCH with Multimodal Data. In Proceedings of the Transforming Learning with Meaningful Technologies, Delft, The Netherlands, 16–19 September 2019; pp. 450–465.
176. Ochoa, X.; Chiluíza, K.; Granda, R.; Falcones, G.; Castells, J.; Guamán, B. Multimodal Transcript of Face-to-Face Group-Work Activity Around Interactive Tabletops. In Proceedings of the CrossMMLA@LAK, Sydney, Australia, 5–9 March 2018; pp. 1–6.
177. Ochoa, X.; Domínguez, F.; Guamán, B.; Maya, R.; Falcones, G.; Castells, J. 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-LAK '18, Sydney, Australia, 5–9 March 2018; pp. 360–364.
178. Roque, F.; Cechinel, C.; Weber, T.O.; Lemos, R.; Villarroel, R.; Miranda, D.; Munoz, R. Using Depth Cameras to Detect Patterns in Oral Presentations: A Case Study Comparing Two Generations of Computer Engineering Students. *Sensors* **2019**, *19*, 3493. [[CrossRef](#)]
179. Huang, K.; Bryant, T.; Schneider, B. Identifying Collaborative Learning States Using Unsupervised Machine Learning on Eye-Tracking, Physiological and Motion Sensor Data. In Proceedings of the 12th International Conference on Educational Data Mining (EDM), Montreal, QC, Canada, 2–5 July 2019.
180. Reilly, J.M.; Schneider, B. *Predicting the Quality of Collaborative Problem Solving through Linguistic Analysis of Discourse*; International Educational Data Mining Society: Montreal, QC, Canada, 2019; pp. 149–157.
181. Schneider, B. Unpacking Collaborative Learning Processes During Hands-on Activities Using Mobile Eye-Trackers. In Proceedings of the 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, Lyon, France, 17–21 June 2019; pp. 70–79.
182. Schneider, B.; Dich, Y.; Radu, I. Unpacking the relationship between existing and new measures of physiological synchrony and collaborative learning: A mixed methods study. *Int. J. Comput. Support. Collab. Learn.* **2020**, *15*, 89–113. [[CrossRef](#)]

183. Andrade, A. Understanding student learning trajectories using multimodal learning analytics within an embodied-interaction learning environment. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17), Vancouver, BC, Canada, 13–17 March 2017; pp. 70–79.
184. Limbu, B.; Schneider, J.; Klemke, R.; Specht, M. Augmentation of practice with expert performance data: Presenting a calligraphy use case. In Proceedings of the 3rd International Conference on Smart Learning Ecosystem and Regional Development—The Interplay of Data, Technology, Place and People, Aalborg, Denmark, 22–24 May 2018; pp. 23–25.
185. Larmuseau, C.; Vanneste, P.; Cornelis, J.; Desmet, P.; Depaepe, F. Combining physiological data and subjective measurements to investigate cognitive load during complex learning. *Frontline Learn. Res.* **2019**, *7*, 57–74. [[CrossRef](#)]
186. Hassib, M.; Khamis, M.; Friedl, S.; Schneegass, S.; Alt, F. Brainatwork: Logging cognitive engagement and tasks in the workplace using electroencephalography. In Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia, Stuttgart, Germany, 26–29 November 2017; pp. 305–310.
187. Olsen, J.; Sharma, K.; Aleven, V.; Rummel, N. Combining Gaze, Dialogue, and Action from a Collaborative Intelligent Tutoring System to Inform Student Learning Processes. In Proceedings of the 13th International Conference of the Learning Sciences (ICLS) 2018, London, UK, 23–27 June 2018; pp. 689–696.
188. Zhu, G.; Xing, W.; Costa, S.; Scardamalia, M.; Pei, B. Exploring emotional and cognitive dynamics of Knowledge Building in grades 1 and 2. *User Model. User Adapt. Interact.* **2019**, *29*, 789–820. [[CrossRef](#)]
189. Maurer, B.; Krischkowsky, A.; Tscheligi, M. Exploring Gaze and Hand Gestures for Non-Verbal In-Game Communication. In Proceedings of the Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play, Amsterdam, The Netherlands, 15–18 October 2017; pp. 315–322.
190. Srivastava, N. Using contactless sensors to estimate learning difficulty in digital learning environments. *UbiCompISWC 19* **2019**. [[CrossRef](#)]
191. Sharma, K.; Leftheriotis, I.; Giannakos, M. Utilizing Interactive Surfaces to Enhance Learning, Collaboration and Engagement: Insights from Learners' Gaze and Speech. *Sensors* **2020**, *20*, 1964. [[CrossRef](#)] [[PubMed](#)]
192. Howard, S.K.; Thompson, K.; Yang, J.; Ma, J.; Pardo, A.; Kanasa, H. Capturing and Visualizing: Classroom Analytics for Physical and Digital Collaborative Learning Processes. Proceeding of the 12th International Conference on Computer Supported Collaborative Learning, Philadelphia, PA, USA, 18–22 June 2017; pp. 801–802.
193. Muñoz-Soto, R.; Villarroel, R.; Barcelos, T.; de Souza, A.A.; Merino, E.; Guiñez, R.; Silva, L.A. Development of a Software that Supports Multimodal Learning Analytics: A Case Study on Oral Presentations. *J. Univers Comput. Sci.* **2018**, *24*, 149–170.

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).