Trends of Multimodal Learning Analytics: A Systematic Literature Review

UMAR BIN QUSHEM

Master's thesis

ITÄ-SUOMEN YLIOPISTO

School of Computing
Computer Science

June 2020

UNIVERSITY OF EASTERN FINLAND, Faculty of Science and Forestry, Joensuu
School of Computing,
Computer Science.

Umar Bin Qushem: Trends of Multimodal Learning Analytics- A Systematic Literature Review.
Master’s Thesis, 69 p., 2 appendices (3 p.)
Supervisor of the master’s Thesis: Solomon Sunday Oyelere, Postdoctoral researcher
June 2020

Abstract: Multimodal Learning Analytics (MMLA) has become a promising emerging concept, but higher education practitioners need to familiarize themselves with MMLA issues pertaining to the usage of multimodal approaches in the educational setting. The goal of this review analysis is therefore to identify, analyze and organize ‘MMLA’ literature, concentrating on how this new branch of Learning Analytics (LA) can broaden learning support without technology-mediated tools and the emergence of various educational platforms. This entry seeks to contribute to the knowledge discovery within Education Technology by exploring current methods and approaches to this immersive technique that can enable MMLA to thrive within and beyond the educational boundary, but also by highlighting key challenges currently experienced through the adoption and implementation processes. Thus, the Kitchenham and Charters methodology has been used to conduct this systematic review study on the Multimodal Learning Analytics and its support for educational activities. Analyzing and synthesizing the theoretical underpinnings of numerous studies, a total of 30 high-quality relevant literature meeting the inclusion criteria was considered. The findings of the review analysis presented several interesting methods and potential challenges identified by the different indicators and criteria which will be very useful for education providers or scientific community. The study also recommended a conceptual framework for what first-hand action would be required to develop an MMLA-based system that focuses on student learning and inclusive education.

Keywords: Multimodal Learning Analytics; Multimodal Learning; Educational Data; MMLA Methods, MMLA Challenges, Systematic Literature Review

CR Categories (ACM Computing Classification System, 1998 version):

K.3.1 **[Computer Uses in Education]**: Collaborative learning; Distance Learning; K.3.2 **[Computer and Information Science Education]**: Computer science education.

Acknowledgement

This thesis was done at the School of Computing, University of Eastern Finland during spring 2020.

First of all, I would like to express my deepest gratitude to my supervisor **Dr. Solomon Sunday Oyelere** for accepting to be the supervisor of my thesis work, for being very supportive during my hard time as well as being a source of motivation to finish the work in time. I am also deeply grateful to my thesis reviewers for their constructive feedback.

Meanwhile, I would also like to thank Associate Professor **Dr. Akram Zeki** who used to be my guide during my early career in higher education, has propelled my professional foundation into research and development activities. Likewise, I am also grateful to the 'EduTech' team at the School of Computing at UEF for giving me an opportunity to work with them and learn from them both personally and professionally.

I am sincerely grateful for the free-tuition and living scholarship of the University of Eastern Finland, which has supported and made it possible for me to receive an excellent education through the fine comprehensive School of Computing at the University of Eastern Finland without being financially worried.

Last but not least, thanks to the Almighty God who has allowed me to sustain, to give me everything I need to live. Finally, I am grateful to my family, especially **Kazi M. A. Kashem** and **Rashida Yesmin**, for the love, prayer and support they have provided me on this journey.

List of abbreviations

3D	Three-dimensional
ACM	Association for Computing Machinery
AR	Augmented Reality
Cmap	Concept Map
CASP	Critical Appraisal Skills Program
D	Document (Primary Study)
ECG	Electrocardiogram
EA	Educational Activities
EDM	Educational Data Mining
EEG	Electroencephalogram
ERIC	Education Resources Information Center
ESL	English as Second Language
GSR	Galvanic Skin Response
ICT	Information and Communications Technology
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of things
IWB	Interactive Whiteboard
LA	Learning Analytics
MDPI	Multidisciplinary Digital Publishing Institute
MMLA	Multimodal Learning Analytics
PhD	Doctor of Philosophy
RQ	Research Question
SG	Serious Game
SLR	Systematic Literature Review
STEM	Science, Technology, Engineering and Mathematics
TLCTS	A Tactical Language and Cultural Training System
UC	University of California
UEF	University of Easter Finland
USA	United States of America
VR	Virtual Reality

List of the Figures:

List of the Tables:

CONTENTS

1. INTRODUCTION

1.1 Background and contexts

The modern time has presented us a host of new benefits and obstacles for the Education Technology community which are particularly abundant in the different form of complex learning and learners' experiences outside the traditional settings i.e Classroom, Technology mediate learning. Although LA (Learning Analytics) and EDM (Educational Data mining) have contributed to analyze some of technology mediate learning data, but there was a hole and limited option to extend the understanding and collection of learning data in open environments (Worsley et al., 2016). However, an innovative technique called multimodal learning analytics was introduced with its multidimensional features to treat rarely considered multimodal data in various learning contexts. Our surrounding learning environments are now designed in such a way as to enable us to acquire different data for learning. As such, Learning Analytics is also very helpful in gathering data and helping learners of all ages. Although the usage of LA assists to classify online learning data for learners (Kausar et al., 2019), the multimodal learning analytics system, also referred to as the extension variant of LA, can be much more supportive in extracting comprehensive data such as when and how education partners of varied backgrounds experience challenges in the process of learning (Tamura et al., 2019). In order to unveil any important learning phenomena and to illustrate the experiences and behaviors of learners, the use of analytics is evident and commonly received given the widespread use of educational technology.

Multimodal Learning Analytics (Saqr, et al., 2020) has been proven to be a powerful tool for deriving complex interactions between people and computing devices across a wide range of sectors. Over the last 10 years, the advancement of multimodal systems and frameworks has started to evolve into mainstream pedagogy, either as a multimodal intelligent tutoring platform or as an interactive smart learning application. Nevertheless, there are hardly inclusive automated multimodal analysis performed in non-technology mediated environments (Worsley & Blikstein, 2015). In addition to sharing a similar ideology in the investigation of educational activities, the majority of work on LA and EDM has paid attention to computer-mediated, lettered, and structured tasks that have left an enormous opportunity and areas for exploration.

In this study, we argued that multimodal data extraction and the MMLA methodology may capitalize on the full potential by generating distinctive insights into the circumstances in which students participate in seamlessly communicating with peers. In addition, with the help of MMLA, education providers can create more opportunities for performance development, and schools can integrate new technologies into existing learning, while teachers can provide support for teaching in both the physical and digital worlds (Blikstein & Worsley, 2016; Haßler et al., 2016).

MMLA technologies tend to be complex in general, since they regularly include different stakeholders, different sources of information, or data handling activities. Besides the context of learning experience growth in today's educational world, the need for analysis of multimodal data derived from various physical environments — face to face engagements, group discussions, writing user stories, and from digital platforms — log files, multimedia based audio and video recordings, sensors based eye tracking, facial expression data and wearable devices, clickstreams, content interactions within the learning management (Pardoo & Kloss, 2011) has been gradually mushroomed. Focusing on learning student behavior and performance utilizing LA and EDM has been a popular practice over the last decade, but a limited data source may threaten a wider perspective of a deeper understanding of the needs and behavioral capacity of students to succeed on a long journey. In such cases, multimodal data-based learning is indeed the best solution for linking digital and physical interactions that can shed light on team-based cooperative learning as well as on measuring the collective sense of student individualistic performance (Martinez et al.,2011, Pijeira-Diaz, et al., 2016).

1.2 Theoretical framework

In perspective of the apparent increase in the value of MMLA, the literature lacks a comprehensive and systematic analysis of existing multimodal applications and methods used inclusively to support education and learning. That limitation was the primary driver of this research. I therefore undertake a systematic analysis of the literature with the aim of presenting a summary of the subject conducted in the field and highlighting potential problems which are not sufficiently addressed by current literature. In doing so, I have adopted the methodology suggested by Kitchenham and Charters (2007) as a guide and decided to pursue the similar but widespread systematic approach. The objective of this systematic literature review (SLR) is to investigate the multimodal learning frameworks connected with teaching, learning or the support of formal and informal education. The study discusses current or potential use cases reported in the literature and assesses the potential impact of MMLA on different education sectors. In fact, this often takes into consideration possible opportunities and difficulties that may result from the involvement of multimodal learning through several platforms.

1.3 Research aim and motivations

The motivation for this review arose from the fact that for theoretical frameworks to gain acceptance in the scientific community, empirical evidence is necessary. Consequently, an accredited overview of this new technology with its supporting features was needed for the audience. This review would contribute greatly to the providing of a sector assessment, thereby identifying areas for future studies that are likely to be of significant interest to both academics and practitioners. This research would extend the support of classification of students learning from technological systems to wide area of learning that students are involved on a regular basis. There are questions that arise concerning 'naturally learning contexts,' such as voice, writing, or gaze, which are sometimes ignored and left out. Thus, the MMLA research should also investigate these specific contexts in order to analyze and provide evidence for both cognitive and technology-mediated learning interactions. The integration of multimodal techniques, which are widely used in the multimodal communication and interaction community, should enable researchers to examine unscripted, complicated tasks in a more holistic way that we seek to identify and discuss here from existing MMLA studies.

1.4 Research objectives

- Firstly, the aim of the research is to outline an overview of the subject and research carried out on MMLA in the field of education. This research starts with the categorization of various types of work, sources of publication, years of publication distributed across databases. This also demonstrates the overview of the most common methodologies, theoretical approaches and research categories included in numerous current literatures.

- Secondly, to be able to assess the usage of MMLA to support education, the objective is to explore, in the literature, the current areas in which applications for MMLA, including specific technologies, methods and frameworks, have been implemented or created. In order to categorize the available research carried out in the past in terms of methodological approach, impact, method and available technology, the aim of this study to address this by examining the contribution and educational relevance.

- Thirdly, after discovering the different sectors and working environments of the Multimodal Learning Analytics, I would like to examine the current challenges faced by the adoption of the MMLA, which incorporates traditional learning practices. Particular focus will be given to the reporting of specific groups of people involved, different technologies applied, outcomes identified in different contexts.

1.5 Research questions

The following research questions are formulated for the purpose of the review in order to seek answers to above-mentioned objectives:

RQ1: What is the current state of the art of Multimodal Learning Analytics?
RQ2: What are the existing Multimodal Learning Analytics methods are being used in Education sectors?
RQ3: What are the challenges of Multimodal Learning Analytics in leveraging various learning practices?

1.6 Chapter summary and thesis structure

This thesis provides a systematic literature review covering the overview of MMLA and discusses the current methods and challenges in the field, which broaden support for the educational activities of students and learners. In addition, it also discovers the state-of-the-art of MMLA studies that describe the growing body of research in learning environments. Although these research efforts have contributed to the discovery of emerging techniques, approaches and challenges to multimodal learning analytics in educational institutions, the ultimate purpose of this analysis is to distinguish between MMLA, LA and EDM studies and to highlight a number of interesting and hidden features that are ignored in mainstream education technology research. This takes to the thesis structure as described in the following.

Chapter 02 gives the fundamental overview of Multimodal Learning Analytics and its taxonomy, followed by the essence of its use in educational activities.

Chapter 03 provides the foundation for the formulation of this systematic review based on the methodology adopted for this study.

Chapter 04 presents research findings on the state of technology, methods and challenges of Multimodal Learning Analytics in Education belonging to the primary studies.

Chapter 05 discusses the findings of three separate study questions and presents the author's viewpoint that would be evident to support MMLA studies in various learning environments.

Chapter 06 describes the overall outcome of the project, the limitations and the future opportunities that the MMLA may have in this area to promote greater diversity and inclusion in the evaluation of the learning and performance-based education sectors.

2. OVERVIEW OF MULTIMODAL LEARNING ANALYTICS

2.1 Definition

Learning is a dynamic and multi-dimensional activity (Wong, 2012). There are many structured and unstructured educational channels that generate data throughout the learning process that become possible with Multimodal Learning Analytics. This modern field of technology, like the MMLA, is a sub-field of learning analytics that handles data collection and integration from various sources, enabling a much wider angle observation of learning processes and various aspects of learning (Blikstein, P., & Worsley, M.,2016). In addition, the MMLA enables the assessment of engagement and particularities that are generally ignored by conventional analysis of learning methods, which often rely solely on computer-based data (Ochoa et al., 2017). Previously a known fact of LA, although LA stores data (learner cognitive and behavioral data and interaction in online learning activities) from a variety of sources collected and merged in a precise manner (Chatti et al., 2017), the implications are limited to the online community. Establishment of the Multimodal Learning Analytics, which promised the extension of LA in the capture of non-technology mediated settings of complex learning data relevant constructs as seen in a variety of learning environments, appears to be part of this. Thus, MMLA can simply refer:

> ***MMLA can be more broadly interpreted as a pipeline or platform that collects and exploits multimodal data to support educational activities through physical and digital environments utilizing the IoT technologies, wearable sensors, signal processing and facilitate Machine-based learning.***

Figure 01: Representation of the MMLA process between human modes and learning technologies

(Di Mitri et al., 2019)

Figure 01 shows a simple illustration of how the MMLA pipeline integrates various human modalities and different technologies to generate multimodal learning results. During the learning process, the MMLA technique used a variety of relevant learning theories, in particular multimodal learning, including text , speech, image, and sometimes haptic modes, while speaking, gesturing, moving, facial expressions and physiological signals are used as modalities. MMLA methods is very useful for examining and measuring student learning signals from a range of sources, such as programming activity assessments, learning user stories (Noel et al., 2018), storytelling in a more panoramic and comprehensive manner (Sadik, 2008).

2.2 The taxonomy of various types of data within Multimodal Learning Analytics

Considering the embedded sensor technology and assessment modalities within the MMLA architecture, three mutually inclusive areas may be considered to be the main concerns of MMLA, which are the evaluation of learner's understanding, the assessment of learner's interaction and physiology and the observation of learner's motives or perceptions. There are many techniques exist — online data mining, customer data mining, online surveys, etc.— however subsequent measures have been chosen for integration with MMLA as they are connected with the leading edge of technology and support the idea of "normal" evaluation (Zaiane, 2001). Thus, although any of these developments constitute a significant impact to research, the real value of including the varieties of specific field-based technology is to put at the forefront a wider range of non-conventional methods embodied by MMLA in a different format. Taxonomy of MMLA has been considered in a generic way due to the nature of the collection identified in both the primary studies and the open learning environment.

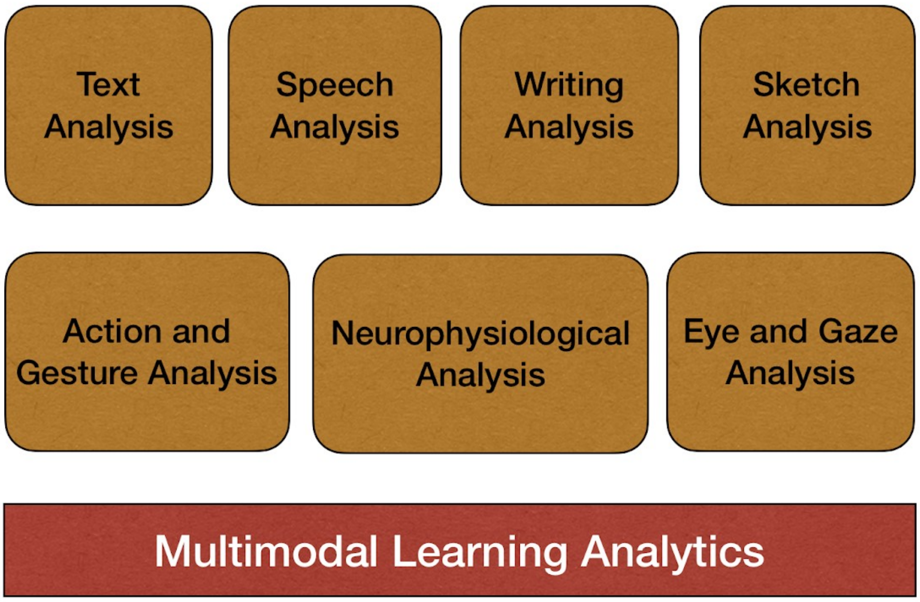


Figure 02: Taxonomy of various types of data for MMLA

1. **Text analysis-based learning data**

Despite the fact of being an uncommon trend in analyzing text based data from the students and learner's activity such as open-ended writing tasks, creating user stories, generated text from tests and exams, explanatory texts, this sector is a promising analytical field for MMLA, given that the transformation of this semiotic approach has shifted dramatically from the traditional to a digital form making it as easily accessible and interpretable. For example, Martin and Sherin (2013) used text-based multimodal analysis using the topic modeling and clustering technique to assess student success in generating ideas and scientifically describing explanations for different seasons.

2. Speech analysis-based learning data

Technology used as a speech recognition system, including smartphones, personal assistants, recorders) has also been developed to help researchers and educators assess student speech in a variety of ways. In speech analysis, it ranges from pre-language features such as speech time, pronounced terms or prosodic features such as tone (Prieto et al., 2016) to exact identification of voices in dialog based interactions between students and teachers (D'Mello et al., 2015). Considering the usage of speech-based technologies to study student activities such as improving reading competence (Beck & Sison, 2006), forecasting students' degree of knowledge in design skills in courses such as engineering (Worsley & Bilkstein, 2011), it opens up the prospect of becoming a measurable area of multimodal learning analysis.

3. Writing analysis-based learning data

Another form of text writing is handwriting that is also a potential field of producing multimodal data. We know this common technique as children begin to learn by writing in a small board or drawing in an art paper. Transforming traditional handmade texts into a recognizable object by means of a handwriting recognition system can provide a greater opportunity for Anthony and Koedinger (2007). They also pointed out that students at an early age prefer to take note of handwriting rather than using the keyboard or mouse during tutoring activities. Interactive whiteboard (IWB) is a good example in handwriting based multimodal approach, although it is often used by teachers, but features meant to be able to recognize teachers and students with timely writing in the fields of study and collective resources (Kuo et al., 2015).

4. Sketch analysis-oriented learning data

Sketching is a new medium of learning practice that many creative students, especially STEM students, use for framework building or mind mapping. One form of sketching based learning is concept mapping, which refers to visual representation of ideas in sequential, hierarchical, and comparative patterns (Hughes et al., 2003). The development of sketch-based technologies such as CogSketch (Forbus et al., 2009) and the interactive 3D sketching system proposed by Misry and Lipson (2007) shows the different levels of students in their computational and engineering concepts, which ultimately form a multimodal learning platform for students to understand cognitive minds and creativity.

5. Action and gesture analysis-based learning data

Action and Gesture based recognition is another form of multimodal learning analytics in the scientific community. Detection of human actions-based movement through cameras or sensors (Weinland et al., 2006; Yilmaz and shah, 2005) in IoT devices helps in tracking different facts about individual learning styles and behaviors. Such an example of student action can be captured by a frame-by-frame analysis of a teacher's video, where students' activity and degree of focus are measured using computer vision (Raca et al., 2014).

On the other hand, gesture recognition is now perceived as a very popular MMLA technique that leverages the infrared camera and accelerometers to balance geometric and visual variances of the camera. Based on this fact, Schlömer, Poppinga, Henze and Boll (2008) tested a technique to capture gestures of different available shapes known in the measurement index. In addition, the recent use of the Microsoft Kinect sensor has become increasingly popular in the analysis of human gestures and body movements. The test was one on student hand gestures pointing out 2 values to investigate their understanding of proportion and provide visual feedback (Howison et al., 2011) that identifies the use of sensor-based multimodalities to identify students and teachers' own behaviors for improvement.

6. Neurophysiological analysis-based learning data

In order to obtain physiological data, this modality can be separated into different parts of the body containing the heart, brain, and skin. Among the state of the neurophysiological medium, brain activity work is popular, using electroencephalogram (EEG) to determine the potential difference within the brain. In the work carried out by Preito and his team (2016), the EEG was used for the predictive analysis of teaching and social interaction activities, while the ECG – Electrocardiogram was used for the measurement of heart activity. Similarly, GSR-Galvanic skin response is referred to as the calculation of electrical conductivity of the skin. These approaches have been combined to detect the effect of naturalistic expression in the neurophysiological state.

7. Eye and gaze analysis-based learning data

Right after physiological substance, there are two more impactful areas that can be used to trace important cognitive and visual data associated with human learning, attitude, and interaction. These are eye and gaze. Analyzing eye gaze data allows to explain what sort of methods could be effective in optimizing student learning. For example, a group of student performances was investigated during the engineering game design, where eye-gaze generated data was a significant indicator of student success (Gomes et al., 2013).

The eye-tracking gadget could very well collect details as to where the participants paid attention to the content throughout the learning phase. In addition, eye-tracking is used to continually assess and enhance the design concept of technology-based learning in the field of educational scientific research (Jarodzka et al., 2017). Popular eye-tracking benchmarks include spatial criteria that show where students have concentrated, but there could be many key explanations as to why students have paid attention to motivation, challenges, organizations, and so on (Alemdag et al., 2018).

2.3 Existing literatures review

Studies on Multimodal Learning and Multimodal Learning Analytics are not new. However, the MMLA-related format or subject area is being perceived differently for different educational environments. Although there is no systematic review regarding the extensive classification of the MMLA methods and challenges. Nonetheless, some literature review research was conducted either as a study, direct reviews on Multimodal Learning Analytics or indirect reviews on multimodal data such as text, audio, physiological or sensor data in various ways.

As mentioned in the Introduction and Overview section above, Multimodal Learning Analytics has shown greater acceptance in analyzing learning environments outside of media technology platforms, but insufficient review has been conducted in MMLA 's Architecture. Shashi, Luis, Maria and Adolfo (2018) raised that concern and reviewed a total of 09 architecture related to Multimodal Learning Analytics applications and systems. They also presented several significant results for the development of potential MMLA architectures. No effective data processing and organizational strategy has been identified out of 9 architectures. They have therefore recommended that researchers should pay special attention when planning MMLA architecture to ensure more flexible system design and help value chain. At the other side, a new aspect of research called affective analytics, but its elements shared the traditional methodology with multimodal learning analytics such as sentiment-based data, emotion, and opinion. Nusrat, Li-Minn, Kah, D.M. Motiur, Tanveer (2020) performed a detailed study to integrate multimodal big data, which is often ignored. In numerous emotional models, different physiological data including sentiments, emotions, micro-expressions were analyzed as part of an affective analytics. They concluded that there is considerable demand for researching multimodal big data and value for such analysis, not only the areas of affective state in the human body analysis, but also the potential for learning different educational and environmental conditions that will help build more robust systems to analyze other multi-modalities.

Recently, another interesting review work on intercultural learning was carried out with the help of education technologies (Rustam, and Wayan, 2020). Researchers evaluated numerous past researches in that review work that considered various learning models including technology, methods, cultures, languages, learning activities. Their results highlighted that two different technologies, such as videoconferencing as a multimodal analytics tool and email as a learning analytics tool, were most significantly utilized in intercultural learning. Observing the potential of applications based on Human Hand Motion (HHM), they surveyed HHM's current technology and research to learn more about multimodal data, such as gestures, contact forces, complex motions, speed of body movement via multiple sensing technologies. Their research categorized numerous hand motions and distinguished contact-based sensing and non-contact-based sensing to promote rapid recognition while providing in depth sensing strategies of hand control glove, sensors, optical markers, ordinary cameras, etc.

Matthew (2013) reviewed natural human interactions in parallel and sequential manner involving computing technologies. In his studies he elaborated a personal point of view with multimodal integrations on human-computer interaction. Review showed the overall history of multimodal interactions, benefits of multimodal interactions, how to build a multimodal integrations system, challenges assisted with multimodal interactions. Matthew further stated that multimodal interaction solutions enable the identification of naturally occurring aspects of human language and actions through the usage of recognition-based technology, whereas multimodal interfaces are typically designed to provide normal and efficient interaction, although it turns out that multimodality has many different advantages.

2.4 Instruments for data collection in Multimodal Learning Analytics

In MMLA applications a broad variety of sensors is used for data collection, ranging from capturing motor and physiological behavior of students, spatial and environmental conditions in which learners are situated (Di Mitri et al., 2018). Providing an example case, the use of multimodal sensory data provides a glimpse into physiological measures that would be almost impossible for humans to perceive. In this way, the integration of low-cost sensors (Cornide-Reyes et al., 2019) will allow exposure to data on learners' interactions with every individual and their environments in physical space. As a result, many of the prospects for multimodal data capture and processing go beyond a simple desire for automation and are motivated instead by the need to deepen one's analysis of intrapersonal and interpersonal interactions, which cannot be feasible only with typical log data.

Furthermore, 'STREAMS' established by Ran Liu and John Stamper (2017) which are used to integrate and combine various sources of multimodal data, including logs data, transcripts, video, audio, eye-tracking, motion, physiology, text annotations etc. *'The Observer XT'* enables the calculation, collection and analysis of any pre-recorded video while generating accurate offline observations in the context of Multimodal interaction between students. This device can be used to acquire behavioral learning signals from the students or learners (Zimmerman et al., 2009). Subsequently, *'SÉANCE'* can predict affective educational conditions such as engagement, creativity, perceptual state of readership and several other evidence relevant to sentiment from a range of pre-existing human perceptions, social identification, and mental states (Liu et al., 2019). The 'DataPrism' framework (Fouse and Hollan, 2010) can support the assessment of a broad collection of time-based multimodal data, using student cognitive mode analysis to identify learner behavioral attitudes, whereas 'KINECT,' a gesture recognition device, uses 3D motion sensing camera depth (Hariharan et al., 2014) to measure object distance from the Kinect camera. For example, the student action or engagement can be interpreted from the E-Learning Class by incorporating gesture recognition using Kinect.

2.5 The use of Multimodal Learning Analytics in Education

The ultimate objective of Multimodal Learning Analytics is to leverage different aspects of existing traditional learning and educational activities. In the light of the study to investigate pedagogical support and transformation achieved so far in the field of education, MMLA has been adopted in different sectors. The use of various forms of multimodal learning analytics, implicitly or implicitly, was already made through the integration of multimodal data (touch, sensors, speech, pen input, camera) which enables different dimensions of learning to be followed and evaluated. Its support extends from computing-based education to enhancing student participation and a variety of more inclusive educational practices. These types of use and the incorporation of MMLA were found in the literature, especially in the primary studies (30 articles meeting the inclusion criteria) chosen for this research. Table 01 shows the indicators as a summary as follows.

Table 01: Support for teaching and learning discussed in the MMLA literatures

No.	Category of Supports	Primary Studies
1	Computing Based education	D1
2	Collaborative Writing of User stories	D3
3	Learning Game based behavioral data in classroom	D4
4	Performance measurement of students' scores on the self-assessment test	D6
5	Improving student's vocabulary and reading abilities	D7
6	Developing an interactive and customized learning environment	D8
7	Facilitating online education systems for students	D9
8	Enhancing feedback in learning context	D10
9	Differentiating student learning strategies	D15
10	Evaluating spatial issues for students and examiners through self-directed and 3D learning	D17
11	Encouraging classroom-based English learning (ESL) in elementary school	D18
12	Computational thinking understanding in physical fabrication, online media design and codes more generally	D20
13	Designing Mobile kits for wearable enhanced learning	D23
14	Collaborative opportunities for using latest applications	D25
15	Incorporation of language with other semiotic techniques to interpret of context in electronic controlled learning environments.	D27
16	Improving student engagement and learning through student involvement in scientific research, and collecting and analyzing real data	D29
17	Missing pedagogical concern from a multimodal, design-oriented, understanding of leaning	D30

2.6 Chapter summary

This chapter laid the foundation for Multimodal Learning Analytics. It described what the meaning of Multimodal Learning Analytics is and what the different types of techniques are considered to be relevant in this field. In addition, this study also pursued a deeper examination of the existing literature reviews and the tools used to collect various interesting data, information, and features for the purposes of the Multimodal Analysis. Finally, it presents exclusively the learning and pedagogical approaches carried out in different works, in particular the primary studies chosen in this study.

3. RESEARCH METHODOLOGY

3.1 Research protocol

In order to ensure a clear, replicable and scientific review of MMLA, I have adopted the process proposed by Briner and Denyer (2012) as well as some of the design elements of the PRISMA statement (Moher et al., 2009). It is necessary to establish a review protocol to confirm that the literature review is structured and to reduce bias among researchers. As such, this systematic review covers a variety of recent work that represents the purpose of this review and originate from the factors that contributed to this analysis. The review protocol is designed in such a way that the research questions describe the key areas of focus of the study. The analytical procedure initially involves the formulation of the research questions described in section 1.5. Note that a glimpse of the study is mentioned earlier. Subsequently, this protocol continues with the development of search terms, preceded by the 'Research Search Strategy', the 'Screening Databases', the 'Quality Assessment and Literature Extraction' and finally the 'Analysis and Synthesis' of this review.

3.2 Construction of search terms

The information below attributes is used for retrieving researches for the study

Educational attributes: Classrooms Learning, Collaborative Learning, Learning Objectives, student motivation, student profiling, semantic education, collaborative learning, educational activities.

Multimodal Learning Technologies: MMLA, Multimodal Data, Multimodal Learning, Mobile Eye-tracker, Hand gesture, Motion sensors, Machine Learning, Emotion detection, Video and speech video, Collaborative Learning. An illustration of the research question that includes the search terms above while searching: 'Use of sensing technology in [**Multimodal Learning]** to support [**collaborative learning]'**.

3.3 Literature search strategy

The strategy for the collection of relevant literatures is divided into two categories. In the first place, I constructed search terms on the basis of research questions to identify the searches for literature in a list of different scientific collections. In the second place, I have searched available literatures extensively and repeatedly across global databases of credible academic resources and publishers, including prominent scientific journals and international conference proceedings. The search keywords were initially defined on the basis of the research questions referred to in Section 1.5. Subsequently, the searching keywords were verified in accordance with the guidelines established by Kitchenham and the Charters (2007) for the review of the works. In order to increase the coverage of publications and ensure that no major primary studies were overlooked, I have chosen various keyword terms. In addition, I used 3 separate search queries to retrieve each of the selected databases and to aggregate findings based on research questions. Each search defined study pertaining to common educational attributes and multimodal learning technologies, all the keywords that were designed for specific query reasons.

Query Search:
Query 1: "Multimodal Learning" AND ("Student OR Classroom OR Semantic Web").
Query 2: "Learning Analytics" AND "Multimodal Data" AND ("Learning OR Education OR Collaborative Learning").
Query 3: "Learning Analytics" AND ("Multimodal Learning OR Multimodal Data") AND ("Learning OR Education OR Collaborative Learning OR Student OR Classroom OR Semantic Web").

Initially, a total of 9145 searches for literature were identified through the 3 rounds of query searches. However, an up-to - date search was conducted to restrict the literature from 2010 to 2020 in order to know the current trends and the expansion of MMLA studies. After combining all the searches and excluding duplicates, I have found 858 literatures for quality checks and subsequently obtained as primary study literature for the purpose of this review. As for the search process, some prominent databases and scientific libraries were used in the search process, such as ACM Digital Library, IEEE Explore, ScienceDirect (Elsevier), SpringerLink, Web of Science (Thomson Reuters) and ProQuest (Eric).

3.4 Literature selection in databases

The selection process for the primary study began with 858 papers scanned to Zotero software, where duplicates were automatically recognized and excluded. Prior to implementing a three-round screening process on the basis of the selection criterion (inclusion and exclusion), it should be noted that the selection of the databases were considered to be the most renowned databases, where a rigorous peer review is regarded to be one of the key policy areas. Criteria for determining the primary source measured which studies should be considered or omitted in each phase of the literature collection process. In addition, the selection criteria for each phase were improved to reduce the risk of missing related research. The general inclusion and exclusion parameters are described underneath and conditions for each phase could be found in Table 02.

General inclusion parameter: literature is included if it is: (e.g. peer-reviewed conference paper or proceedings, journal paper, scientific publication, case report, graduation dissertation, data articles, magazine, book chapter, video publications, patents and software) AND (e.g. Multimodal Learning Analytics and related support for education).

General exclusion parameter: literature is excluded if it is: (written in a language other than English, a summary, an abstract (extended), an undergraduate thesis or a complete book, duplicate and non-reviewed publications) AND (do not discuss multimodal learning and do not relate to the application of the MMLA).

Table 02: Inclusion and Exclusion conditions

No.	Inclusion conditions	Exclusion conditions
Phase 1	Meets the general parameter for inclusion	Meets the general parameter for exclusion
Phase 2	Meets the inclusion parameter in Phase 1 AND is from (Multimodal Learning or educational activities)	Meets the exclusion parameter in Phase 1 AND is certainly not from (Multimodal Learning or learning environments)
Phase 3	Meets the inclusion parameter in Phase 2 OR is discussing about MMLA tools	Meets the Phase 2 exclusion parameter AND (does not belong to any MMLA

	or methods has been used during inclusive learning and collaboration in educational activities.	related studies OR has learning elements but is not suitable OR is not particularly relevant to the multimodal learning process that integrates MMLA methods and educational activities)

Summary of the steps used to capture literatures:

There are few steps that I have followed in the collection of the literature body. In the first step in the selection of literature, I made initial searches from various well-known databases, which involve checking constructed keywords in the general search for the best possible extraction of literature. In the second step, I incorporated searches with inclusion and exclusion parameters where the previous steps are searched for keywords in the title, abstract, keywords. In the third step, I verified all the literature either to be refused or to be approved. The dismissed literature of the previous step is checked only for the terms "Multimodal data" and "Educational activities" in each of the titles of the literature, abstract and assessed if they are connected. This would prevent the rejection of papers with various combinations of search terms, such as "Multimodal data for Learning Analytics." As a result, the included literature obtained from the two earlier steps is validated manually by reviewing the abstract and the full text of the literature. In this step, the literature used in the evaluation provided results that were explicitly or implicitly relevant to the MMLA sector.

3.5 Quality assessment and information retrieval

This quality assessment and information retrieval processes of this study were performed in parallel on the basis of guidelines for the systematic literature review. The information retrieval process initiated after the 'Primary Studies' was forwarded to the ATLAS.ti. After the search terms were constructed, search queries were executed in selected databases and journals, and 3 rounds of refined searches were also used (see Phases). Moreover, after reading the literature rigorously, if it was found to be relevant and interesting to address research questions, the literature was finally collected to be included as a primary study. In order to maintain the integrity of selected existing works, I have reviewed literatures on the basis of the quality assessment criteria set out in the study (Dybå & Dingsøyr, 2008) and reported by the Critical

Appraisal Skills Program (CASP). These criteria are measured with the mandate of three main issues related to quality – rigor, credibility and relevance. Pursuant to the adopted quality criteria, the collection of literature was used to demonstrate briefly.

Table 03: Quality Assessment Criteria

Attributes	Assessement Descriptions
Problems	Is there a problem the research is striving to solve?
Aims	Does the research have a clear aim to achieve?
Contexts	Does research belong to the activities of teaching and learning?
Research Design	Was the developed research design capable of addressing the aims of the study?
Methods	Do study methods include learning tools, data collection, potential techniques for analysis?
Research Type	Does the data analysis or the case study provide sufficient information to the audience?
Findings	Are there any significant findings presented in the results
Study Value	Does the study show any valuable contribution to the education sector?

3.6 Analysis and synthesis

From the initial set of 858 studies identified, the search process identified a total of 30 potential literature that were analyzed (Pereira et al., 2014; Seuring and Gold, 2012; Kache and Seuring, 2014), coded and critically evaluated in order to conduct a data analysis using different strategies (Spens, 2006; Guthrie et al., 2004). The research direction initially provided a descriptive statistical synopsis of the state of the MMA study field in RQ1. Subsequently, the recommendation of Cruzes and Dybå (2011) followed a thematic analysis of RQ2 and RQ3, while the literature was reviewed using ATLAS.ti.

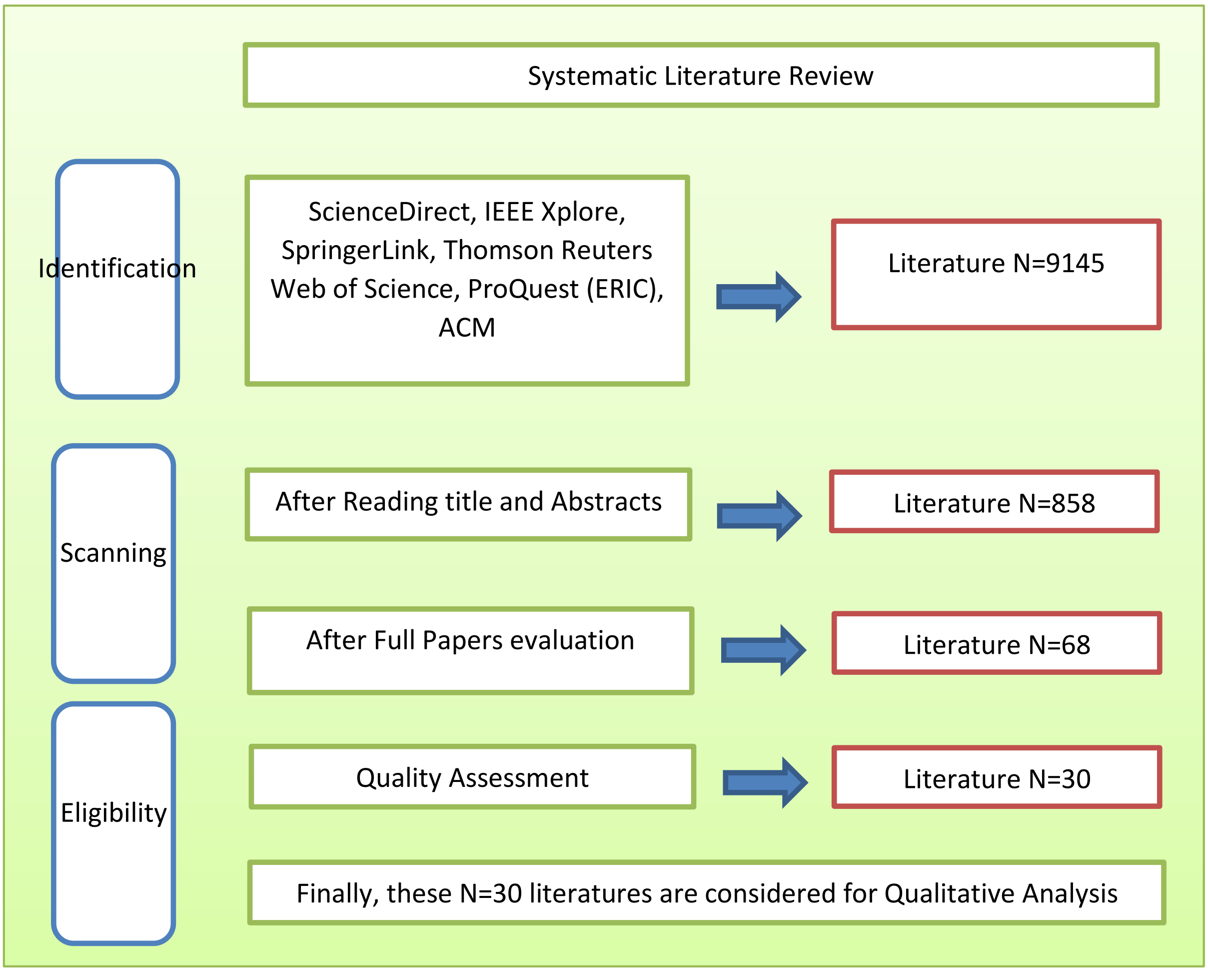


Figure 03: Schematic representation of the Systematic review process (Pereira et al., 2014, Casino et al., 2019)

3.7 Selected Primary Studies

Table 04. highlights following primary studies which are structured and collected on the basis of research questions presented in Section 1.5. This selective literature will be used to support the review work.

Table 04: Selected Primary Studies

ID	Authors	Title of the Primary Studies
D1	Schneider, B (2020)	Lowering Barriers for Accessing Sensor Data in Education: Lessons Learned from Teaching Multimodal Learning Analytics to Educators

D2	Cowling, M. A. (2020)	Mixed Reality Multimodal Learning Analytics
D3	Noel, R (2018)	Exploring Collaborative Writing of User Stories With Multimodal Learning Analytics: A Case Study on a Software Engineering Course
D4	Giannakos, M. N. (2019)	Multimodal data as a means to understand the learning experience
D5	Andrade, A. (2016)	Using Multimodal Learning Analytics to Model Student Behaviour: A Systematic Analysis of Behavioural Framing
D6	Sharma, K (2019)	Building pipelines for educational data using AI and multimodal analytics: A "grey-box" approach
D7	*Wang, S. P. (2018)*	Effects of Multimodal Learning Analytics with Concept Maps on College Students' Vocabulary and Reading Performance
D8	Junokas, M. J. (2018)	Enhancing multimodal learning through personalized gesture recognition
D9	Perveen, A (2018)	Facilitating Multiple Intelligences through Multimodal Learning Analytics
D10	Di Mitri, D. (2018)	From signals to knowledge: A conceptual model for multimodal learning analytics
D11	Cornide-Reyes, H. (2019)	Introducing Low-Cost Sensors into the Classroom Settings: Improving the Assessment in Agile Practices with Multimodal Learning Analytics
D12	Liu, R. (2019)	Learning linkages: Integrating data streams of multiple modalities and timescales
D13	Shankar, S. K. (2018)	A Review of Multimodal Learning Analytics Architectures
D14	Ochoa, X. (2016)	Augmenting Learning Analytics with Multimodal Sensory Data
D15	Worsley, M. (2015)	Leveraging Multimodal Learning Analytics to Differentiate Student Learning Strategies

D16	Blikstein, P. (2016)	Multimodal Learning Analytics and Education Data Mining: Using Computational Technologies to Measure Complex Learning Tasks
D17	Birt, J. (2019)	Piloting Multimodal Learning Analytics using Mobile Mixed Reality in Health Education
D18	Kuo, Fang-O. (2014)	Develop and Evaluate the Effects of Multimodal Presentation System on Elementary Student Learning Effectiveness: Within Classroom English Learning Activity
D19	Vujovic, M. (2019)	Motion Capture as an Instrument in Multimodal Collaborative Learning Analytics
D20	Richard, G. T. (2019)	Digital and Physical Fabrication as Multimodal Learning: Understanding Youth Computational Thinking When Making Integrated Systems Through Bidirectionally Responsive Design
D21	Fjørtoft, H. (2020)	Multimodal digital classroom assessments
D22	Samuelsen, J. (2019)	Integrating multiple data sources for learning analytics—review of literature
D23	Kusmin, M. (2019)	Engaging Students in Co-designing Wearable Enhanced Learning Kit for Schools
D24	Tamura, K. (2019)	Integrating Multimodal Learning Analytics and Inclusive Learning Support Systems for People of All Ages
D25	Stevenson, M. (2015)	Visualizing Solutions: Apps as Cognitive Stepping-Stones in the Learning Process
D26	Worsley, M. (2016)	Situating Multimodal Learning Analytics
D27	Tan, S. (2016)	Multimodal research: Addressing the complexity of multimodal environments and the challenges for CALL
D28	Guichon, N. (2017)	SHARING A MULTIMODAL CORPUS TO STUDY WEBCAM-MEDIATED LANGUAGE TEACHING
D29	Prieto, L. (2017)	Smart School Multimodal Dataset and Challenges

3.8 Generated Code Trees from content analysis

After identifying the primary studies, I performed a content analysis using ATLAS.ti, which was critical for performing qualitative studies and exploring the answers from the literature depending on the study's purpose. Figure 04 is a representation of the codes used in the content analysis during the search. All codes were grouped together and divided under a hierarchical tree where specific codes were narrowed down by: 'Research Questions Focus', 'MMLA base', 'Literature Related', 'Analysis Methods', 'Objectives', 'Learning Contexts', 'Technologies', 'Environments', 'Challenges', 'Methodologies'. Note that some of the codes are interconnected with certain groups of codes, if required.

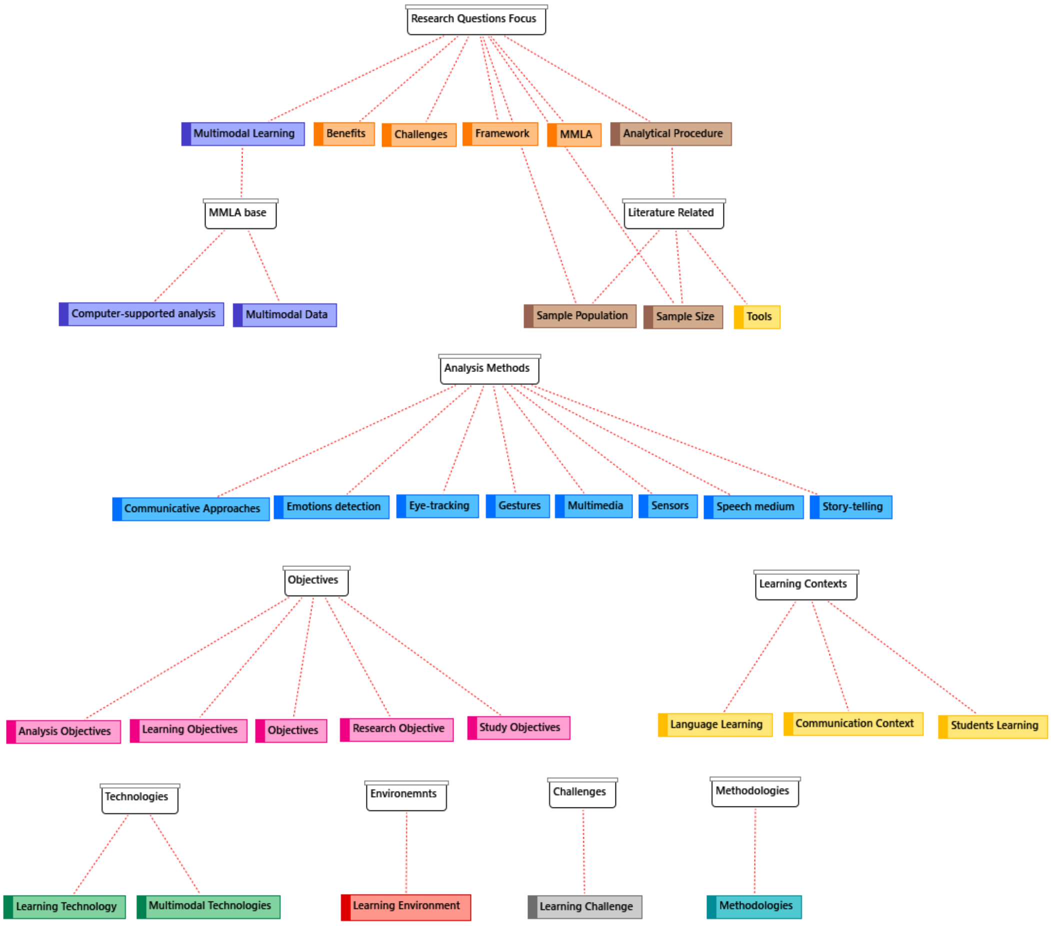


Figure 04: Representing the code trees from the content analysis

3.9 Chapter summary

This section described the overall research procedure and how the review was planned and carried out in accordance with the protocol adopted. Before and after the collection of literature, various tools and software were used to make the work easier. Within the scope of the methodology, the construction of search terms helped to find the appropriate literature, while the search strategy was used to narrow the collection of literature. In addition, I have also applied specific criteria and conditions for inclusion and exclusion in the selection of literature in databases. Finally, we have validated the overall procedures by the quality assessment, so that the review study could derive and analyze the findings of the research objectives.

4. FINDINGS

4.1 The state of the art of Multimodal Learning Analytics in primary studies

The purpose of the inclusion of the state of the art in this study to show the most current research has been conducted in this fields state of the art in research (Aesaert et al., 2013). As far as the systematic review is concerned, the output of the systematic review process and the results of existing literature, including the type of research, technology, years of such work distributions and other factors, can be discussed for the wider audience. RQ1 therefore analyzed the literature and the results of the review in a descriptive approach.

4.1.1 Primary study sources and number of publications

The Table 05 below summarizes how the MMLA related studies were identified in a variety of publications, ranging from journal articles, book chapters, conference papers, reference works, and workshop-based articles in well-known scientific databases. Although this field (Multimodal Learning Analytics) is certainly emerging, however, we have applied rigorous literature analysis and query rounds to find these numbers of entries in the following. Out of the list of 30 Primary Studies, 19 papers were collected as journal articles from peer-reviewed journals, two of which were collected from the 'Journal of Learning Analytics' (JLA) and the 'Journal of computer assisted Learning' (JCAL), making them the top two, among other publications. Additionally, 6 conference articles, 3 book chapters, 1 reference work and 1 workshop article were reported as the other form of publications in this study.

Table 05: Sources of Primary studies and number of papers published

Journals	**Papers**
Journal of learning analytics	2
Journal of computer assisted learning	2
IEEE Access	1
International Journal of Information Management	1
IEEE LATIN AMERICA TRANSACTIONS,	1
British Journal of Educational Technology	1
Journal of Educational Technology & Society	1
Journal of computer assisted learning	1
Turkish Online Journal of Distance Education-TOJDE	1
Sensors	1
Journal for STEM Education Research	1
European Association for Computer Assisted Language Learning	1
Language Learning & Technology	1
ACM Transactions on Computing Education	1
Computers & Education	1
Electronic Journal of E-Learning	1
Technology, Knowledge and Learning	1
Conferences	
IEEE 18th International Conference on Advanced Learning Technologies	1
IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)	1
Procedia-Social and Behavioral Sciences	1
International Conference on Human-Computer Interaction	1
Proceedings of International Conference of the Learning Sciences	1
Proceedings of the Fifth international conference on learning analytics and knowledge	1
Book Chapters	
Research and practice in technology enhanced learning	1
Perspectives on Wearable Enhanced Learning (WELL)	1
European Conference on Technology Enhanced Learning	1
Reference Work	
Encyclopedia of Educational Innovation	1
Workshop	
Joint proceedings of the sixth Multimodal Learning Analytics (MMLA) workshop	1

4.1.2 Country wise Multimodal Learning Analytics studies within Primary studies

Discovering the MMLA studies as state of the art, I have considered an approach to demonstrate how MMLA studies have spread over the last decade. Following the Table 06, it was clearly stated how the MMLA studies were conducted by the country. Table reported only 3 publications between 2010 and 2015, taking into account the early stage of the field. Interestingly, the USA dominated this field of research, followed by Australia and Norway. Statistically, the research carried out by the researcher in the USA amounted to a total of 8 making 22.22% of the total contributions as per country, while Australia and Norway showed great interest in the entries, with a figure of 13.89 % and 11.11 % each. Chile and Estonia shared 8.33% of the contribution interest in the field, while Taiwan and Brazil shared about 6%. On the other hand, there are many countries, such as France, Spain, Pakistan, the Netherlands, Ecuador, Japan, Sweden, and Switzerland, which contributed 2.78% individually, though perhaps the lowest.

Table 06: Country-wise yearly primary study in MMLA between 2010 and 2020

Country	2010-2015	2016	2017	2018	2019	2020	(%)
USA	1	3	–	1	2	1	22,22
Australia	1	1	–	–	2	1	13,89
Norway	–	–	–	–	3	1	11,11
Chile	–	–	–	2	1	–	8,33
Estonia	–	–	1	1	1	–	8,33
Taiwan	1	–	–	1	–	–	5,56
Brazil	–	–	–	2	–	–	5,56
France	–	–	1	–	–	–	2,78
Spain	–	–	–	–	1	–	2,78
Paskistan	–	–	–	1	–	–	2,78
Netherland	–	–	–	1	–	–	2,78
Japan	–	–	–	–	1	–	2,78
Sweden	–	–	–	–	1	–	2,78
Switzerland	–	–	1	–	–	–	2,78
Ecuador	–	1	–	–	–	–	2,78

4.1.3 Databases per type of publication among Primary studies

Searching in databases is somehow crucial for systematic review. Figure 05 represents the search history of databases that primary studies belongs to. From the year of 2010 to 2020, there are many renowned databases included in this study such as ScienceDirect, ACM, IEEE, Springer, Wiley, Eric, JSTOR, MDPI, Cambridge University Press, UC Berkeley, HAL Archive, and lastly CEUR-WS. In 2014, the ScienceDirect indexed MMLA-related entry analysis was just 1, while 2 entries were indexed by ACM and Eric in 2015, reporting just 3 studies in the first quarter of the decade. However, in 2018 and 2019, almost all of the studies were indexed by the top-tier databases where Springer was the highest (4 entries) in 2019 and IEEE was the second highest (3 entries) in 2018 in terms of publication. Among the databases, Springer, IEEE, and ScienceDirect had demonstrated considerable interest in the MMLA research, which reflected in these collected primary studies.

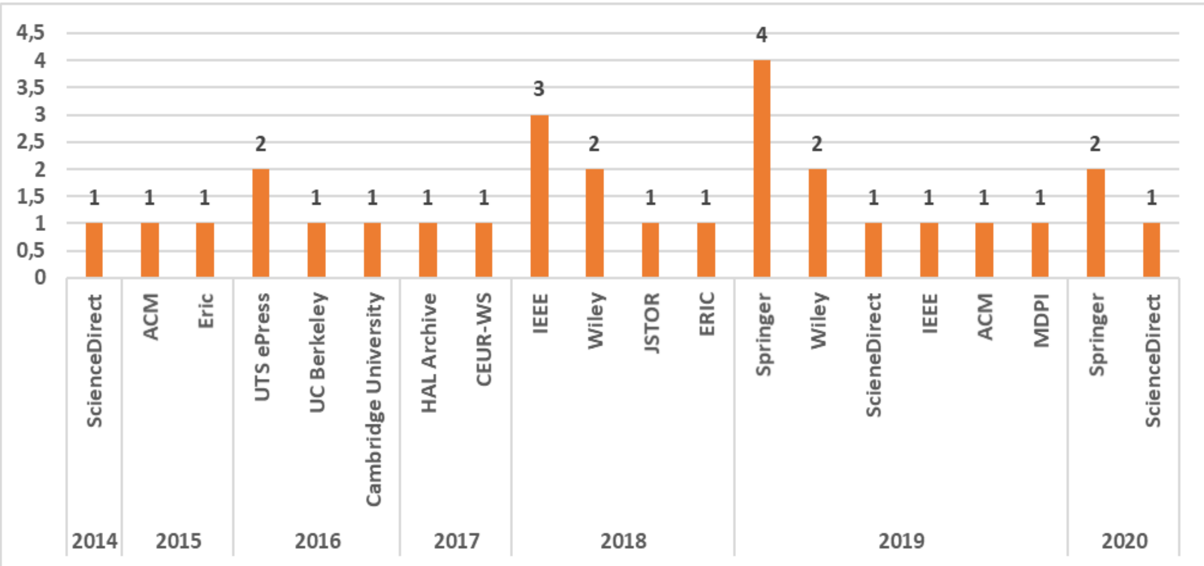


Figure 05: Analysis of selected databases for each year, per type of primary study

4.1.4 Word-cloud and the potential keyword extractor: Primary studies

This word cloud in Figure 06 visualizes keywords or factors related to Multimodal Learning Analytics. Primary studies chosen for the review study included keywords such as learning, education, multimodal, students, analytics, collaboration, digital, performance, understanding, and so on, which provided an overview of the discussion in those literatures.

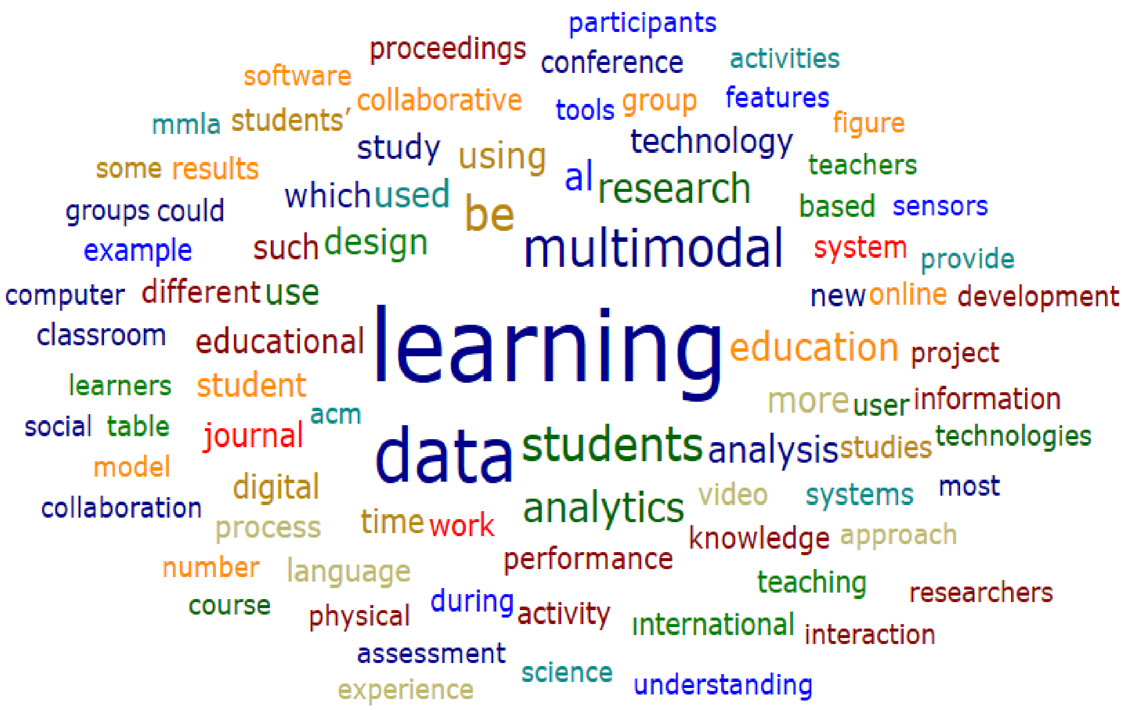


Figure 06: A word cloud generated from the Primary studies

For the sake of content analysis, the keywords mentioned below were designed to extract the appropriate findings for analysis. Terms were a combination of the initial search keywords as well as the keywords in the original literature. Based on the Figure 7, The learning environment and learning technology were seen to be a critical component, as these primary studies reflect a significant range of education sectors that MMLA has so far been active in. Among other keywords, some keywords provided a greater response, such as Multimodal Learning, sensors and behavior. In identifying the main findings on the basis of research questions, ATLAST.it explored terms such as MMLA, Tools, Pedagogical approach, eye-tracking, gestures, Multimedia in primary studies that were very useful.

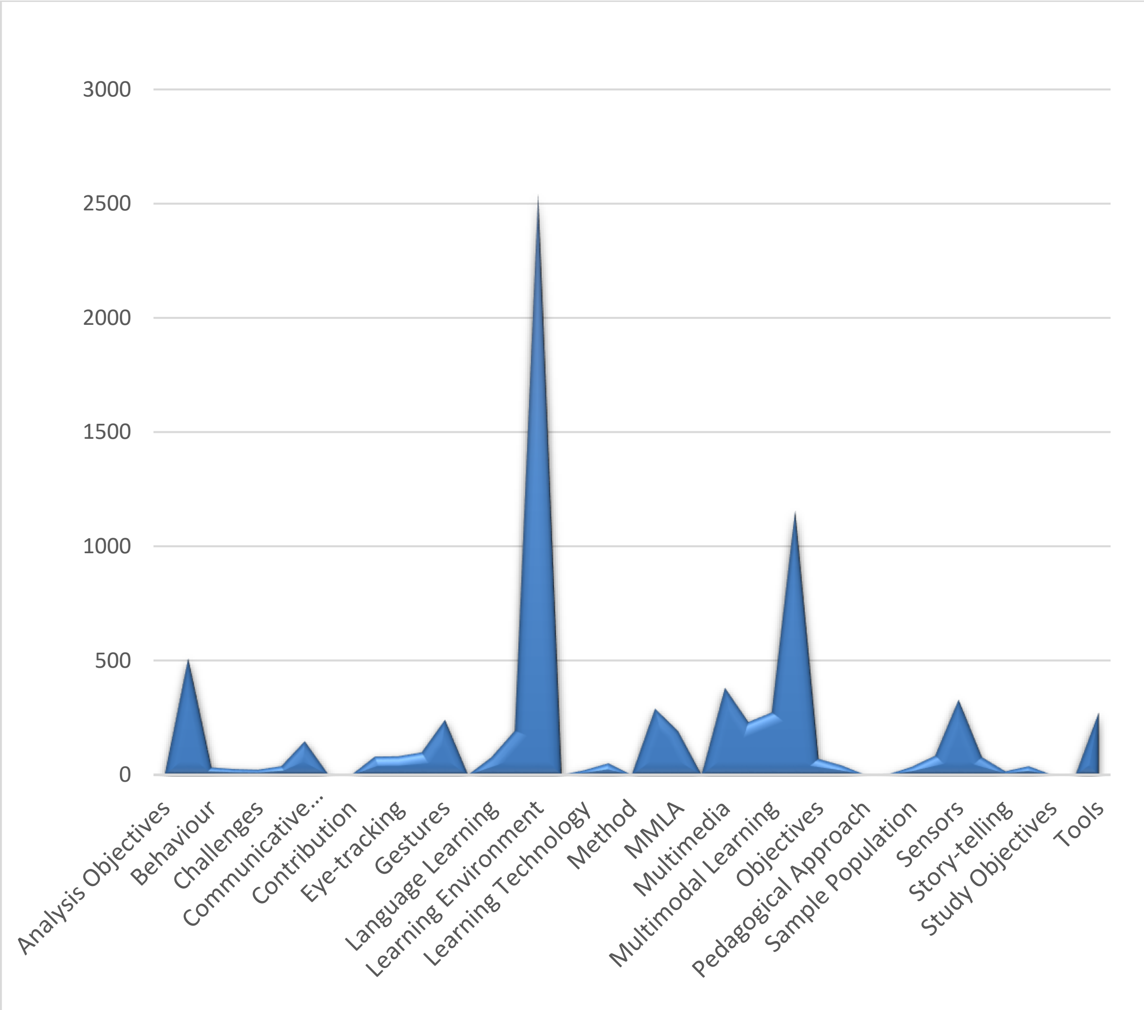


Figure 07: Potential content analysis keyword extractor

4.1.5 Theory, research category and research methodologies

following the analysis of the MMLA related studies, this study identified a number of theoretical approaches, different types of research and methodologies within primary studies. These figures illustrated those internal developments and demonstrated the foundation of the MMLA studies undertaken during the last decade. Figure 08 showed that there were 7 theoretical approaches used in the overall body of research in the field of observation (1), Model (3), Mixed-Framework and Model (3), Framework (6), Case Study and Pilot Study (1), Case Study (3) and Analysis (1). Moreover, category of researches included in Figure 09, were Action (1), Case study (3), Conceptual (1), Descriptive (1), Empirical (5), Experimental (12), Review (5), and Theoretical (2). Finally in Figure 10, the research methodologies used in MMLA studies used in MMLA studies were Theory establishment (1), Survey (1), Scenario-based research design (1), Questionnaire and Interviews (1), Quantitative (1), Qualitative (4), Mixed mode research (7), Design and Learning Technology (5), Data Analysis (2) and Action research (1).

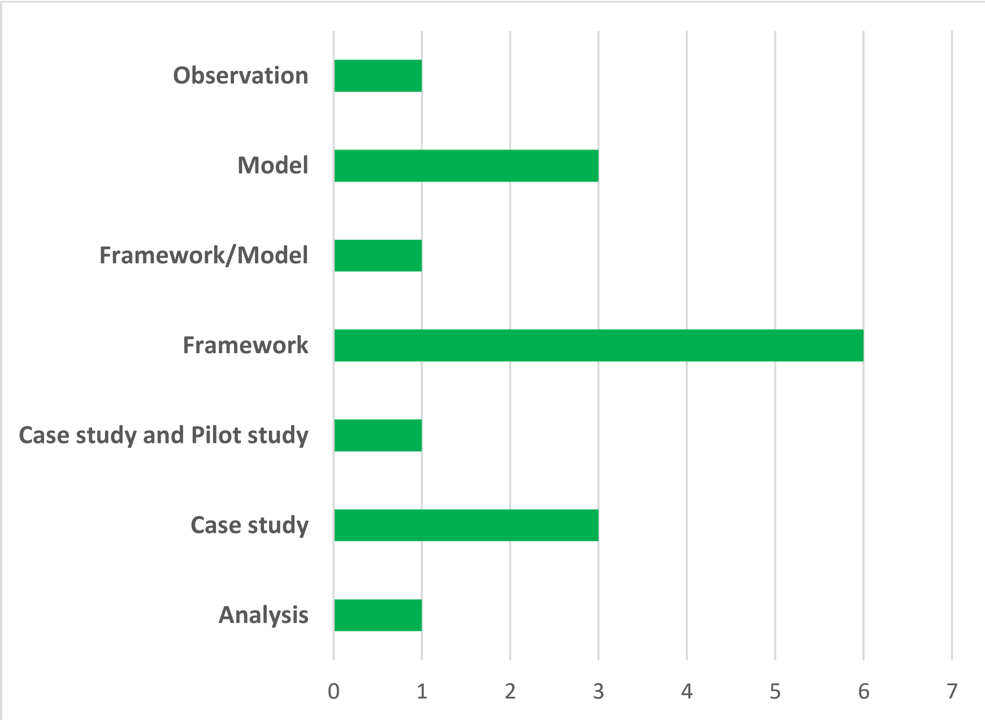


Figure 08: Theoretical Approaches conducted within MMLA related literature

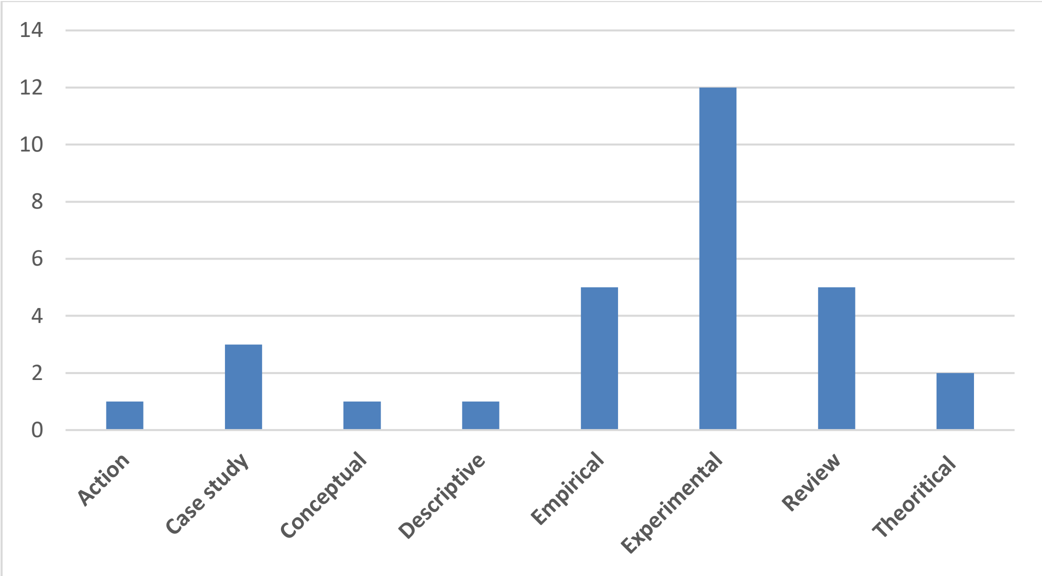


Figure 09: Distribution of research category among MMLA literatures

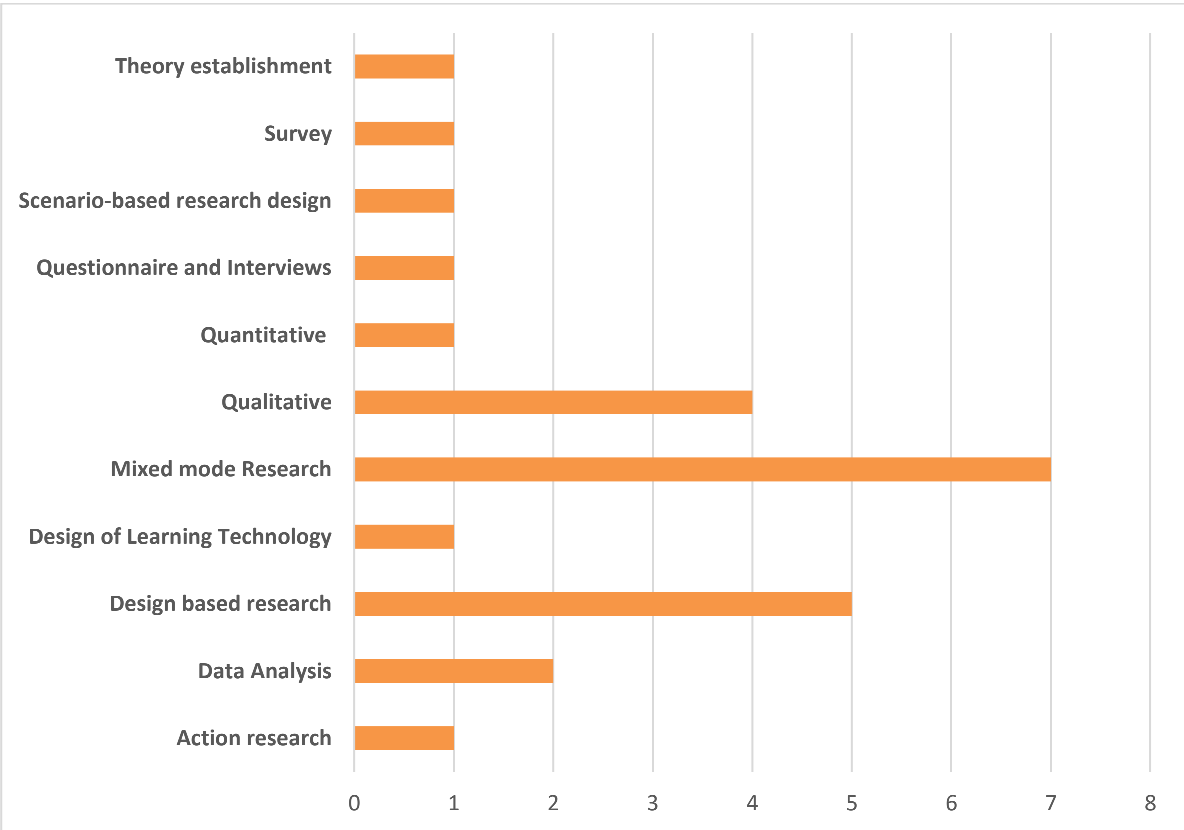


Figure 10: Research Methodologies adopted in Primary studies (MMLA)

4.2 Potential Multimodal Learning Analytics methods to support Multimodal Learning

This study explored various types of multimodal learning analysis methods when investigating the existing learning environment. Exploring the different types of MMLA representation in the latest available platforms and educational settings, it was highly important to categorize and define these methods and approaches in a hierarchical manner. Potential MMLA methods and approaches with various tools and techniques elaborated how Multimodal Learning Analytics has been used directly and indirectly in the past. In the following, the classification of potential methods can be seen in reference to the primary studies mentioned.

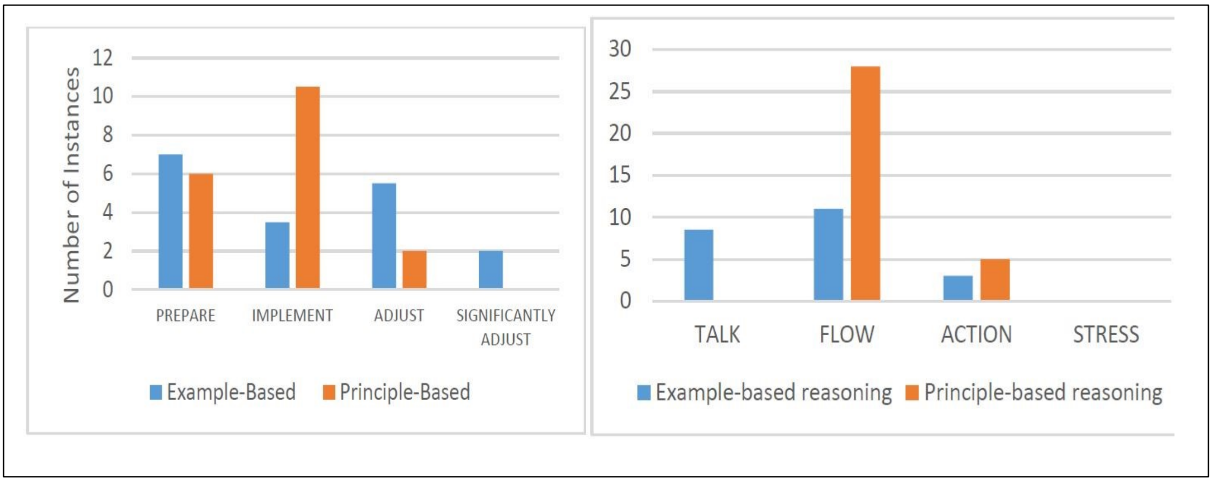
Based on the Table 07 below, these potential methods were clustered under different indicators on the basis of behavioral, educational, physiological and digitally supported methods. Any of the comprehensive literature that directly implemented multimodal learning techniques or MMLA methods were being collected and analyzed within the framework of this review. The remaining methods and approaches of the MMLA are included as a summary for further reading in the appendix.

Table 07: Classification of potential methods according to indicators

Indicators	Potential Methods	Primary Studies	Target group	Sample Size
Behavioral	Modelling Student Behaviour	D5	School graders	30
	Example based reasoning, and principle-based reasoning	D15	Primary students	20
Educational	Learning language using vocabulary and through reading comprehension	D7	Students	70
	Digital texts writing, taking pictures, creating audio and video recordings, modal practices	D30	NA	505
Physiological	Physiologies of public speakers and their connection to student engagement	D1	Students	32
	Click-streams, Keystrokes, EEG, Eye-tracking, Gaze, Video, Wristband	D4	Participants	17
	Sensor data (Eye-tracking, EEG, Facial expressions, wristband data)	D6	Students	32
	User-defined gestures, accessing data- skeleton positions, and kinematics features	D8	NA	21
	Physical indicators (gaze direction, the distance between learners and the speed of movement)	D19	Pair of groups	3
	Eye Tracking measurement, EEG measurement	D24	Participants	8
Digital	Audio recording	D3	Students	60
	Multiple modalities data and contexts (individual vs. collaborative classroom learning).	D12	Students	59
	Observational dashboard illustrating MMLA, data recording and summative analytics	D17	NA	6
	Interactive whiteboard (IWB) integartion within Multimodal presentation system	D18	Pupils	134
	Sensors kits	D23	Students	645

4.2.1 Behavioral methods

The differentiation of student learning strategies using uni-modal and multimodal methods is considered a milestone in the communities of computer science and learning analytics. Researcher Marcelo Worsley with Paulo Blikstein (2015) extended the scope of MMLA integration in student behavioral and reasoning state to discover multimodal practices. The study presented two groups of determinants (Figure 11) before analyzing the state of the design study: "example based reasoning– utilizing instances from the real world as a point of entry for the task of fixing; and principle-based reasoning – using engineering fundamentals as the grounds for one's design" (Marcelo et al., 2015). Although 9-12 graduate students and 8 graduate students were considered to take part in the study, students were randomly selected to



different groups.

Figure 11: Common behavior by conditions reported on hand annotated data (left) and Multimodal sensors data analysis (right) (Worsley and Blikstein, 2015)

Data collection from the target population was carried out using three specific MMLA tools – a Kinect sensor for recording speech, video and motion, and a web-based camera for monitoring student activity for materials and lastly electrodermal activation sensor for measuring stress. Based on their applied methods, they employed hand-annotated and multimodal sensor-based data analysis to state the difference of behaviors of learning strategies in between two groups of students. Their study concluded that the hand-annotated 'IMPLEMENT' presented a comprehensive view of how principle reasoning can relate to success and learning. However, the evidence of the multimodal sensing technology provided a far more concrete description of the two investigational groups. It has also been observed that the use of similar analytical techniques in both experiments may be used in the study of differences between two experimental circumstances. Furthermore, Alejandro, Ginette, Joshua (2016) tried to model student behavior using Multimodal Learning Analytics. Their study was to frame learning behavior (body, motions, and gaze) in rich video data of student interviews as a cluster and determine how activity of learners to enhance our understanding of the relationship between observed behaviors and learning. Participants were 30 students from the U.S. elementary school who were interviewed by the researcher, provided they were involved in a project named honeybees collection and nectar.

Figure 12: An experimental proof of the Hesitant behavioral frame (Body backwards or straight, gaze on paper or away, fidgeting, hedging language, soft voice) (Alejandro et al., 2016)

Their analysis using statistical models and matching algorithms has demonstrated that MMLA technique can help to identify clusters of easily observable behaviors in students' interview activities. In addition, the clustering of behavioral frames containing learners' physiological learning status reveals a very promising area of research because they tend to contribute to concrete student learning interventions. Characteristics of interactions in the hesitant frame on the figure (Figure 12) indicate that the student typically does not make a lot of eye contact, chooses to restart sentences, and offers tentative answers. They have concluded the calm behavioral frame appears to co-occur with more demonstrations of competence in the form of mechanistic reasoning while the transitions between frames during the interview indicates grouping students into a small set of profiles correlated with performance.

4.2.2 Educational methods

The engagement of students in learning a variety of courses has already been a common phenomenon. As a result, enormous ways of focusing on active student learning have been recognized. While the Estonian Tallinn University agreed to set up a ‘Smart School project’ (D30), the Latin American research group unveiled a 'low-cost classroom sensor' to analyze the development of core skills among students in the software engineering course (D11). Learning through a game-based platform, whether it is for language learning or learning in active collaboration with colleagues, is no longer new. Such examples of these implementations include TLCTS – a tactical language and cultural training system (Lewis and Johnson, 2010) and Serious Game (SG) like Lego, designed to enhance collaborative reasoning and teamwork capabilities that include pre-and post-graduate capabilities.

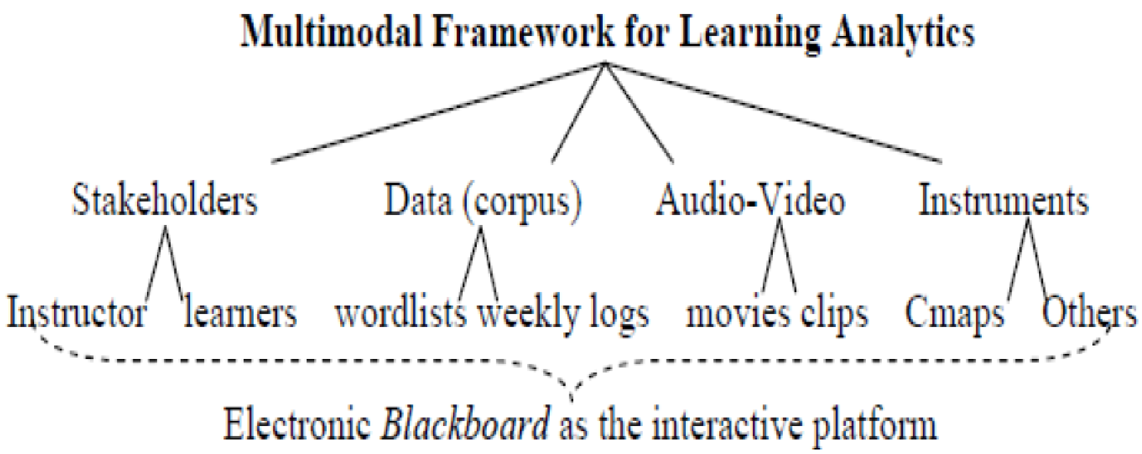


Figure 13: A MMLA based framework promoting language education (Wang and Chen, 2018)

Shih-Ping Wang and Yih-Lan Chen from National Taiwan University of S&T (2018) proposed a framework (Figure 13) that integrate multimodal learning with the concept map (Cmap) approach to assist in language learning (remembering vocabulary and successful reading) for students. The experiment includes course design, measurement and lastly analysis. A total of 70 students divided into two different classes participated in the reading and vocabulary study, which was further evaluated on the basis of the weekly logs used as corpus to observe their learning behavior. According to the framework, both groups adopted a different approach to reading and vocabulary, while one group considered Cmap, the second group preferred a more traditional approach to vocabulary memorization. Nevertheless, the ultimate achievement of the study was the integration of the framework in collecting different groups of student's performances. The MMLA based framework demonstrated a significant effect on learning abilities that students can partake in language learning. Additionally, the instruments used for the study is also a recommendation to advance students' reading skills.

4.2.3 Physiological methods

Recently, action-oriented analysis has been reported to be quite convincing in illustrating student experience during lessons (Mangaroska and Giannakos, 2018; Blikstein and Worsley, 2016). It was known that when an action is made, regardless of whether or not it is done, a person creates valuable knowledge which is often not included in the design of education technology (e.g., cognitive ability, eye movement, facial gesture). These physiologically focused Multimodal Learning Analytics may provide useful insights (e.g., consumer focus and psychological state) into advanced learning experience analysis (Pantazos and Vatrapu, 2016). Terms like Multimodal system and integrated framework that have been coined in extended modern learning environments. Multimodal data extractor features tools, such as eye-tracking and EEG measurement were used in a range of MMLA applications to support the data collection of hybrids learning systems combining online and offline. Tamura (2019) originally discovered a procedure to examine the hybrid academic landscape for university students. Although his discovered framework enabled learners and teachers to engage using a web-based interface, it was not appropriate for learners of all ages and there was limited choice to track or collect information about learners' abilities to work in materials or to navigate lessons properly.

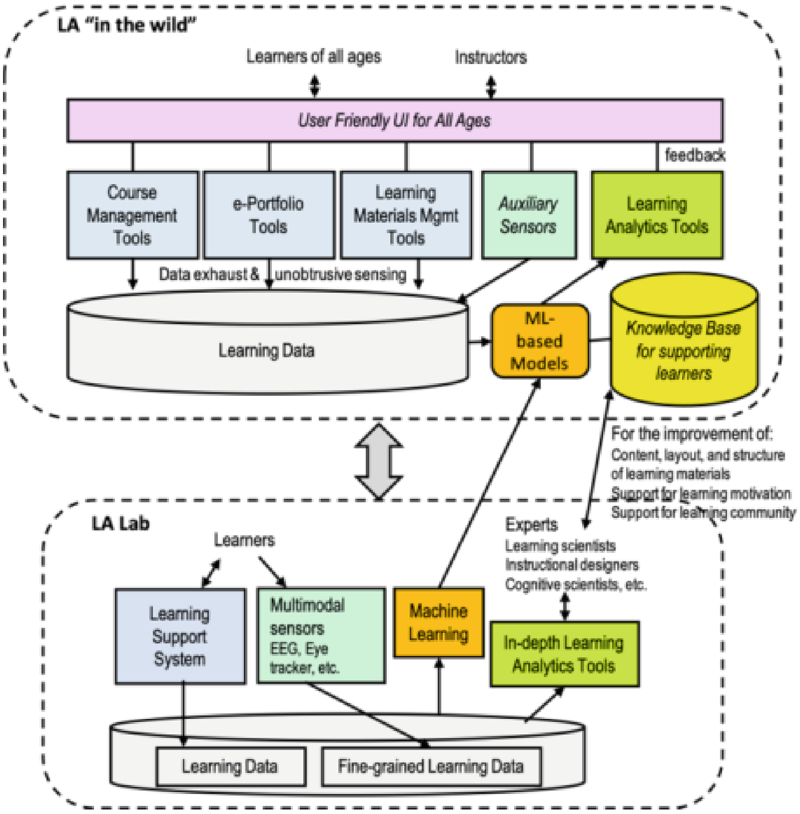


Figure 14: An integrated MMLA architecture supporting different age level learners (Tamura et al., 2019)

Thus, he updated and proposed MMLA architecture (Figure 14) and prototype development that incorporated multimodal sensors including EEG and eye tracker, Dual-Tablet. The system was designed to reduce operating layers and make a handy visual representation for many people, especially those who are older in age or have difficulty learning online platforms (Tamura et al., 2019).

4.2.4 Digitally supported methods

Digital media and tools are very common outside school settings. Collecting data from physical classroom was a challenging process (Giannakos, 2016). Although online mediate learning and multimedia-based tools and IoT technology made it possible in recording students' activities (Okubo et al. 2016). The idea of Multimodal digital and smart classroom required much more digital connections of ICT tools in the classroom. Multimedia technology such as audio recorder, video recorder, and IoT-based technologies such as stickers, eye trackers, wearable wristbands (D4), a gesture recognition system (D8), analyzes learning or teaching processes in traditional classroom, and promotes awareness, regulation, and observation processes among the learning process stakeholders (D23).

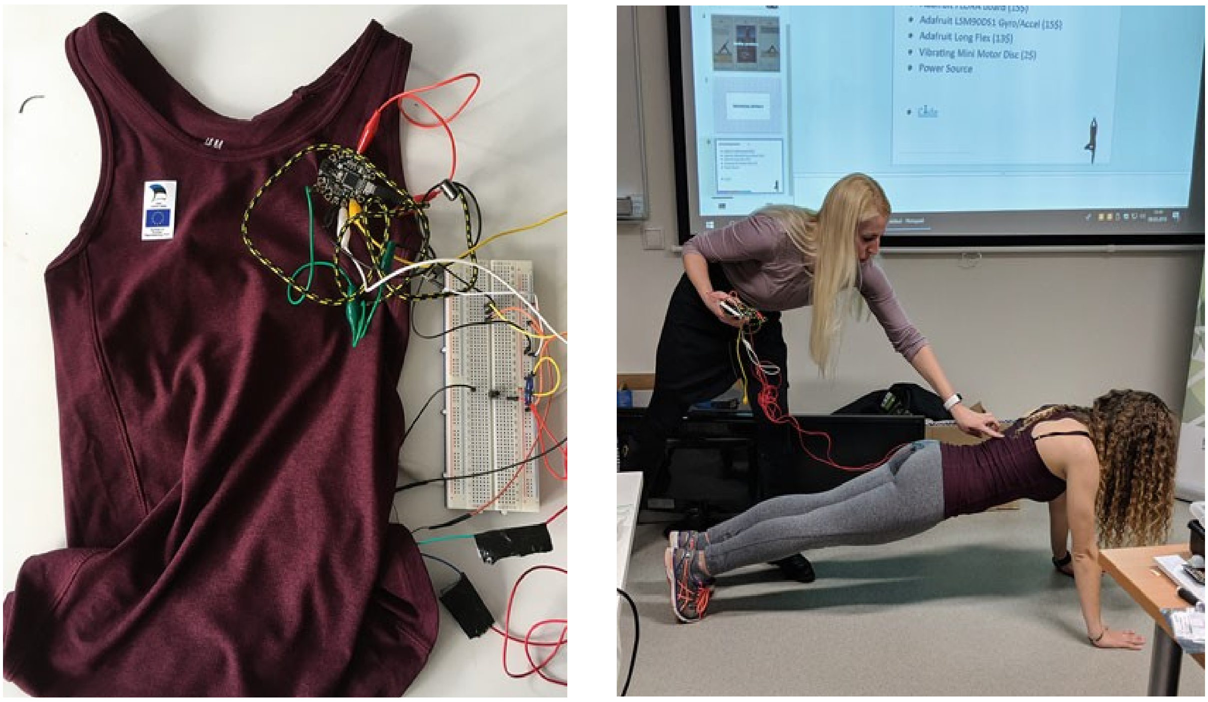


Figure 15: Wearable learning kit's prototype development and testing (Kusmin et al., 2019)

INNOVATORIUM referred as a smart schoolhouse aimed to innovate the STEM learning process, conducted a survey through skype prior creating scenarios and later constructing personas within the process of building a wearable kit (Figure 15). Survey participants involved 179 participants including 159 students and 15 teachers. In the meantime, students were engaged in seeking an answer knowing a 'course by performing'. In addition, the Henning Fjørtoft (D21) multimodal class assessment carried out another approach using various example-based experiments such as creating videos to learn mathematics using a different modality medium (Figure 16), describing fiction as a type of interview that reflects a different culture from the novels on Television programs. These ubiquity of interactive technologies in several school environments helped teachers to combine learning and evaluation through the implementation of modern classroom-based assessment. Multimodal System supporting different digital services extended the evaluation of practices of teachers to provide a larger array of representations.

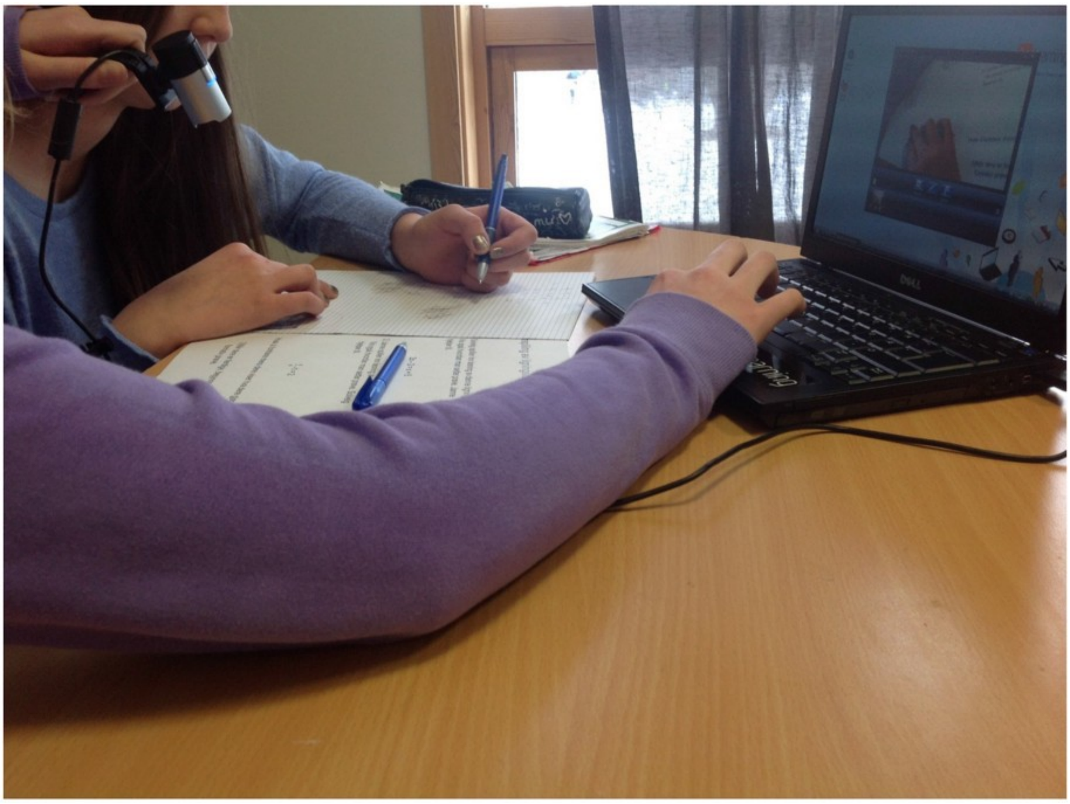


Figure 16: Student making a video for studying mathematics through multimodal approaches (Fjørtoft, 2020)

4.3 Potential challenges of Multimodal Learning Analytics in leveraging various learning practices

While these research efforts contributed to the development of Multimodal Learning Analytics in the Education Community and for more inclusive learning activities. Several aspects with the MMLA was overlooked. In order to appreciate the demands of constantly changing multimodal methods and strategies that continue to change and impact conventional teaching and learning activities, the difficulties and obstacles observed in the MMLA studies must be recognized. Thus, our review study aimed to find these challenges in RQ3 which is describes in the following.

Table 08: Classifications of identified obstacles and challenges in the literature based on the criteria

Primary Studies	Potential Challenges	Criteria
D1	Data privacy; use of student data.	Data Related
D2	Data Storage; Privacy and security; machine learning capability.	
D4	Data collection and consent from the participants.	
D13	Complex data in MMLA; lack of support for the data organization.	
D26	Data privacy, cost, data synchronization, and data capturing process	
D28	Ethical issues in sensitive, confidential and private data.	
D3	Quality of the recordings in audio extraction.	Environmental
D8	Personalizing gestural interactions.	
D11	Assessment of core skills development; ambient noise; mobility of sensors.	
D21	Fieldwork; coverage of area.	
D29	Smart school physical settings; user identification.	
D9	Tracing multiple intelligences from a course learning experience.	Learning and Pedagogical
D10	Lack of understanding of multimodality learning data; ambiguity and misalignment of terminology.	
D12	Student common misconceptions; limitations of the tutor interface.	
D15	Connection of learning science with learning practices.	
D16	Complexities in computational approaches; actionable deas; learning theories.	
D20	Interchanging various coding blocks for learning or understanding of computational concepts.	
D7	The need for training for the use of a multimodal approach.	
D27	Framework for language learning analysis; digital tools for the analysis of complex intersections.	Technical
D17	Interface challenge for the Mixed-Reality applications.	
D18	Managing learning materials in less time in traditional classrooms.	
D22	Lack of technological solutions for the combination of multiple data sources.	
D24	Device operability.	

4.3.1 Data related challenges

MMLA was already conducted in a variety of domains where different types and modalities of datasets were used, ranging from behavior of students, collaborative interactions, teachers' training, remote teaching system operations, and the use of multimodal sensing technologies in classroom activities. Based on the studies identified, Data privacy (D1), dealing with ethical issues while collecting data (D4) from students found very concerning in this sector. It has been also reported that although data storage (D2) over privacy concern made it harder for doing research work, lack of support (D13) in managing confidential and private data (D28), capturing naturalistic data, feedback and its synchronization (D26) was something that need to be addressed.

4.3.2 Environmental challenges

Suitable environment was needed not only to support comprehensive education for students and teachers, but also when MMLA-based methods were established to track interaction from personalized gesture (D8) to understand student behavior. The challenges identified in traditional classrooms were also difficult to assess the development of core skills and quality recording (D3) for ambient noise or sensor mobility in outdoor activity (D11). In addition, environmental support focused on infrastructure such as setting up a physical school, implementing a user identification (D29), or fieldwork to reach less centralized local areas (D21) with low accountability disclosed several key sectors to explore before experimenting with MMLA studies.

4.3.3 Learning and pedagogical challenges

Apart from the specific challenges of MMLA, there was a several main constraints related to learning and pedagogical practices (D15). For examples, when an instructor taught a mathematical equation via an online tutorial system, it was not difficult for the students to wrongly interpret the equation (missing parts) due to the interface problem (D12) rendering it a learning challenge for the student and a tracing challenge for the MMLA video based system (D9). Nonetheless, a lack of understanding of multimodal learning details, ambiguous and confusing terminologies (D10), complexities of computational approaches to actionable ideas and learning theories (D16, D20) in the implementation of multimodal technology reflected a small fraction of challenges in the multimodal research group that demonstrated the need for user training to learn how to incorporate multimodal infrastructure into reality (D7).

4.3.4 Technical challenges

Despite the above-mentioned challenges, the shortage of digital tools and appropriate techniques made the analysis of complex multimodal data and multidimensional interactions (D27) considerably crucial. For example, a customized framework for the language learning analysis (D27) required extraction capacity of data from the interaction log to measure the usefulness of the instructor medium. In addition to the physical challenge of setting up the aforementioned physical classroom, the management of different learning materials (offline and online) for the traditionally established classroom (D18) needed additional technical sufficiency to meet the demands. Similarly, with regard to the learning challenge of accessing study materials through multimodal systems, the design of interfaces, such as mixed-reality applications (AR: Augmented Reality, VR: Virtual Reality) (D17) concerned the operation of the devices of those applications (D24) and the combination of multiple data sources in the solutions implemented (D22).

4.4 Chapter summary

This segment discussed the results of the review processes and presented answers to the research questions established. The review analysis was conducted using non-statistical methods and extracted the results from the primary studies chosen. The results of the findings revealed the application, tools and methodologies used for Multimodal Learning Analytics (MMLA) and investigated the main challenges it has in the education sectors. It should be remembered that almost all primary studies had multiple target groups, numerous data collection techniques, technologies or more than a research objective and theoretical approach.

5. DISCUSSION

5.1 RQ1. What is the state of art of Multimodal Learning Analytics?

From the research findings according to RQ1 above, several studies related to Multimodal Learning Analytics were extracted. Particular attention was paid to studies that were more recent in such literature, limited to 2010-2020. In a perspective of SLR, we used several peer-reviewed literatures form the reputed database in collecting supported primary studies for the review. In the section 4.1, some of findings covered number of literatures found within MMLA studies with their sources. Note that, in order to maintain a quality study, it considered a combination of literature, including the 'Journal', the 'Book Chapter', the 'Workshop', the 'Conference Paper', the 'Reference Work' to cover a comprehensive search. In addition, our search highlighted that 22.2 % of the MMLA studies conducted in the USA made it the top-notch in the list country wise. The most preferred database for publishing research on MMLA was the Springer database, while the 'Journal of Learning Analytics' published works were among the top of the journals. Finding the state of the field in primary studies, framework-based work as a theoretical approach, experimental as a research category, mixed-mode research as adopted methodologies were recognized as the highest in number.

5.2 RQ2. What are the existing Multimodal Learning Analytics methods are being used in Education sectors?

During the review of 30 collected literature, the content analysis using ATLAS.ti gathered existing MMLA methods and techniques. Initially, auto coding feature of ATLAST.ti helped to explore various relevant, semi-relevant and irrelevant methods of Multimodal Learning Analytics among the conducted literatures. All the methods found in studies was classified within a 4-indicator scale based on the purpose of the methods used. Potential methods indicators classified as behavioral methods, Educational methods, Physiological methods, and Digitally supported methods. Behavioral methods helped to find multimodal data from different learning behaviors, such as the use of example-based reasoning and principle-based reasoning to define student success correlations. Educational methods helped to classify the application of the MMLA used in learning practices, classrooms, or different forms of communication in educational activities, such as writing digital texts. Physiological methods helped to find different kinds of physiological state used to track multimodal learning data, such as eye tracking, gesture data collection, defining the speed of movement during lecture listening. Finally, digitally supported methods were used to determine the methods whereby technologies and tools used either to assist learning process or to collect learning data of the learners for use in multimodal analysis. The study has demonstrated that the various physiological and cognitive styles of students can be examined using specific multimodal learning methods, such as eye-tracking, gesture recognition, sentimental condition. Identifying student engagement when studying languages or participating in group conversations in a multimodal context offers valuable perspectives for teachers and service providers (Saqr et al., 2020) that help facilitate the inclusiveness of multimodal activities in the learning process. In addition, this study also included some of the interactive devices or IoT-based technologies such as wearable learning kits (Kusmin et al., 2009) that will open a new space to support student learning from their physiological data.

5.3 RQ3. What are the challenges of Multimodal Learning Analytics in leveraging various learning practices?

This analysis identified many of the current obstacles involved in conducting the Multimodal Learning Analysis or the implementation of the MMLA system. A number of structured and unstructured issues and issues were studied in primary literature. However, the analysis revealed these challenges in a more constructive way by introducing a 'challenge classifier' based on the criteria. The results of the study also showed that some of the studies directly highlighted the challenges, some of which were explicitly mentioned in the studies. The challenge classifier listed the different challenges of MMLA studies based on criteria such as data related challenges, environmental challenges, learning and pedagogical challenges and lastly, technical challenges. Moreover, the challenges identified in the literature analysis were positively associated with the successes and failures of multimodal learning technologies that evaluate specific student performances (Balogun et al., 2019) and interactions in the multimodal learning environment. In addition, this systematic study not only represented the ongoing challenges of either data privacy or ethical issues, but also identified environmental (Qushem et al., 2017), pedagogical and technological challenges that demanded proper examination and workable solutions within the educational boundary.

5.4 Procedural conceptual framework

This proposed procedural conceptual framework (Figure 17) is designed to help service providers in the education community gain a deeper understanding of how MMLA can transform student educational activities while maintaining the proper use of MMLA-based methods and systems.

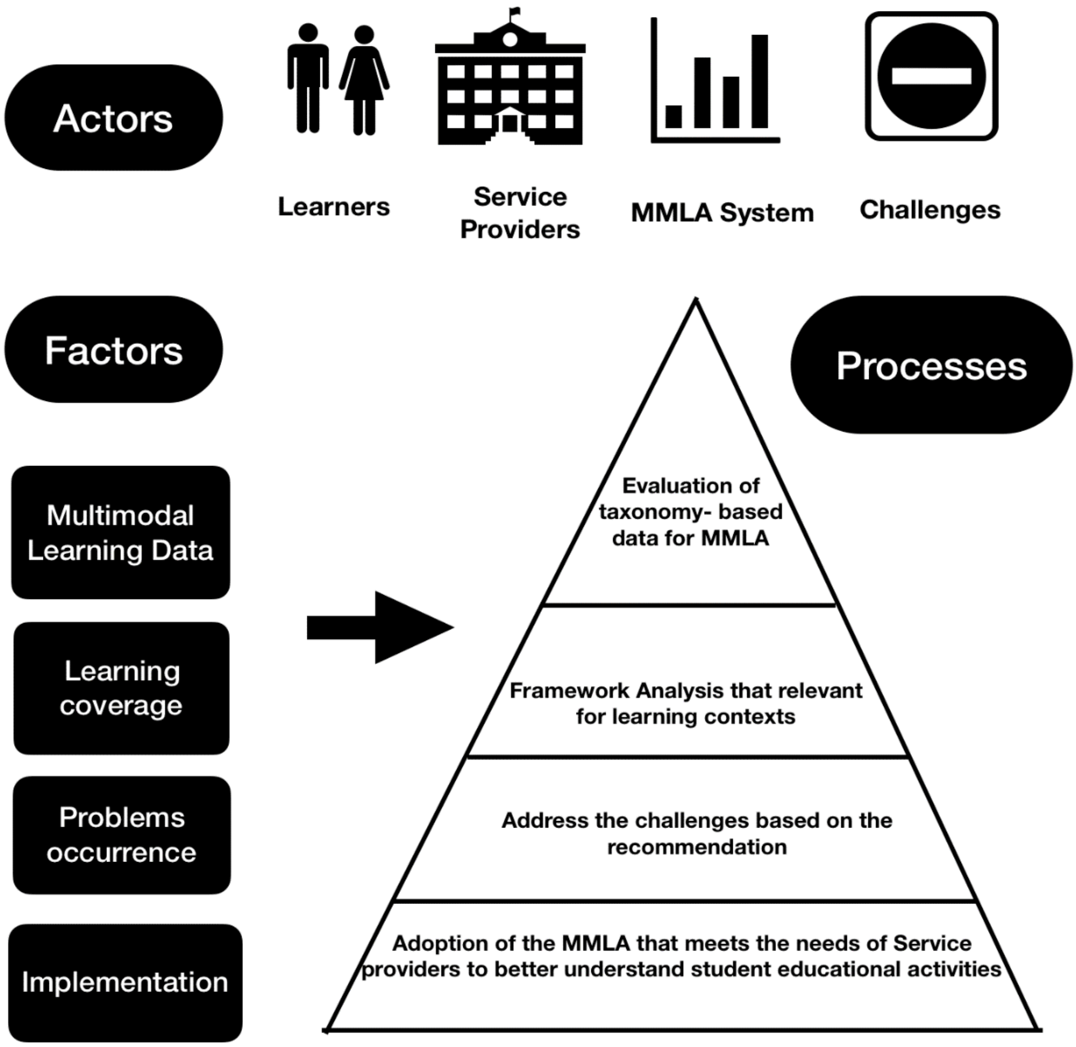


Figure 17: A Procedural conceptual framework for MMLA implementation

This framework for the implementation of the MMLA was established on the basis of the findings of the review identified in this study. The proposed framework includes different actors from teaching and learning platforms, traditional classrooms, student engagement in online and offline learning communities. In different environments and contexts, these actors are mutually inclusive in addressing a wide range of multimodal interactions, leveraging usage of multimodal data (Qushem et al., 2017) in creating multimodal data. As a central theme of the framework to leverage various forms of learning practices, learning tools and methods in and for multimodal learning, it sets out procedural guidelines dealing with multimodal data, integration of learning, challenges prior to the implementation of the MMLA-based inclusive learning environment. In the light of the above example, the procedural framework can be shown as follows.

Actors:

Framework consists of a total of 4 components or actors (direct and indirect) that are 'Learners', 'Service providers', 'MMLA system' and 'Challenges' (Figure 17). These components have an impact within the procedural framework at different stages. Learners includes students in academic courses, web-based tutorials, experiments, interviews, language learning groups, while service providers are different authorities in the existing education community, such as physical schools, tutors on different online MOOC-based platforms such as Coursera, EdX, Udemy and others. Subsequently, the 'MMLA system' and the 'Challenges' are often indirect actors found in the study.

Factors:

Actors of the framework are directly and indirectly affected by factors such as multimodal learning data, learning coverage, problems of occurrence and architecture implementation (Figure 17). Multimodal Learning Data — text and writing data, speech data, drawing data, gesture data, neurophysiological data, eye, and gaze data — included in various learning settings. Learning coverage defines the contextual aspects that multimodal learning and analysis have been established in this study. Problems that have occurred in pedagogical settings, not only between instructors and learners, but also between service providers during observation of their environmental set-up in multimodal learning. As a result, implementation has always been challenged by other factors and raises concerns.

Processes:

In Figure 17, the triangular shape represents the overall processes contributing to the successful implementation of Multimodal Learning Analytics in support of educational activities. Learners generate large volumes of learning data across a number of learning contexts, including face-to -face conversations in the study of various languages, using example and principle-based reasoning by knowledge-based technologies, utilizing audio and video records to monitor specific deductive and interactive learning approaches. Moreover, individual, and group-based learning to solve complex learning tasks and to analyze student movement and behavior through EEG, Eye-tracking, Gaze, or wristband also produces new multidimensional learning data. While these learning approaches have taken place, they create opportunities for Multimodal Learning Analytics. Traditionally built MMLA-based systems do not seem to be able to capitalize on these big data of learning traces for various systematic and environmental challenges. Challenges such as data storage, privacy, complex computing, lack of support, organizational support, lack of training, increased assessment of outdoor skills and many others should be addressed to the use of the MMLA infrastructure in the education sector. If either service providers are able to address the above-mentioned challenges by building an MMLA system, or if they adopt an MMLA system that identifies those challenges, the Education Community will benefit from this new but emerging technology by implementing the Multimodal Learning Analytics functionality in order to understand the educational activities of students.

5.5 Chapter summary

This analysis of the findings initially aimed at discovering the state of art related to MMLA Studies. Targeted literature analysis later provided a comprehensive overview of multimodal learning practices in educational activities through the adoption of MMLA techniques and methods. Based on these activities of Multimodal Learning Analytics, various barriers and challenges evolved, which was reported in this review study. In order to know the implications of research as a new source of knowledge in this educational community, all research questions have been reviewed and resolved. On the basis of the findings, a procedural conceptual framework has also been established to provide a new guide for the educational community.

6. CONCLUSION

6.1 Contributions and Limitations

This study revealed the different methods used by Multimodal Learning Analytics not only to demonstrate how the use of multimodal data can be beneficial in educational activities, but also to uncover namouras challenges that exist in the MMLA-related learning community. This review analysis was focused mainly on describing the overall picture of this new field: MMLA, which is distinguished from other experimental or model-based research. It described a set of modalities that have been the subject of multimodal analysis over the last decade, as well as modalities that have recently emerged as new data streams through which researchers can study human interaction and behavior.

Looking at the subject areas, MMLA is quite new to studies such as Learning Analytics and Educational Data Mining. However, the area of multimodal learning analytics will keep expanding tremendously, partially due to the ability to process excessive amounts of data and a broad range of research components. As a consequence of improved efficiency and exposure to data, multimodal learning analytics can offer a better way of learning activity trends, contexts, and experiences. The research also showed that learning analytics underlines the need for a deeper understanding of how to interpret data with a focus on maximizing outcomes and use data to enhance the learning process at all stages. In addition, it illustrates a framework that will be a roadmap for general service providers to address existing problems in the application of the MMLA.

Although the study will provide good support for future research work in Multimodal Learning Analytics, some limitations should also be noted. This study examined the current state of the application of MMLA, methods and challenges in the field of education. The literature review in this research was performed in a systematic and qualitative way to define various uses for MMLA in the context of learning. The review includes a total of 30 reference documents to justify the overall results, which might have been even higher if the research covered MMLA studies not only for students but also for specific educational stakeholders. In addition, data collection from different databases was more difficult when some of the literature found that access was limited.

6.2 Recommendations and future works

After conducting the review analysis, it can be suggested that MMLA, with the help of the latest technologies such as IoT devices, can capture and analyze student learning processes and educational activities while providing learners with an assessment platform for better learning experience. Considering the potential of the MMLA to provide support to a group of educational stakeholders (teachers , students, tutors, administrators), this systematic review sought to establish the importance of the Multimodal Learning Data Analysis and the MMLA system in teaching and learning.

Literatures have exhibited that, also in conventional schools, instructors participate in major multimodal activities to emphasize and de-emphasize various concepts in a lecture. In the same manner, students utilize a range of approaches to illustrate their expertise and, most specifically, to acquire an understanding of the subject field. Although end-user methods and strategies in multimodal learning analytics shared some similarities with unimodal analytics, there should be a higher need for multimodal analysis to support the increased difficulty of data acquisition in MMLA. Our educational environment should also provide importance to teaching in the context of instruction, the development of learning environments and the assessment of such activities should always be regarded and embraced in the best way possible.

After the review, analysis and closed observation of the dissemination and application of Multimodal Learning Analytical Approach in the scientific community, there are plenty of opportunities and challenges that are simultaneously accountable for future success in this field. Future work should therefore focus more on the incorporation of multimodal techniques in less technology mediating sectors, where it requires researchers to look at unscripted, complex tasks in more holistic ways. Such problems, far from being extraordinary, are the subject of active lines of research that continuously offer creative ways to increase the sustainability and viability of MMLA activities. Finally, educational technology researchers and professionals must consider the use of various modalities in their own studies and methods to mitigate open problems such as privacy or data storage of multimodal sensing data. The multimodal LA community should explore different multimodal technologies and frameworks that address existing challenges while dramatically improving student learning activities and performance that can lead to inclusive learning.

REFERENCES

1. Aesaert, K., Vanderlinde, R., Tondeur, J., & van Braak, J. (2013). The content of educational technology curricula: A cross-curricular state of the art. Educational Technology Research and Development, 61(1), 131–151. https://doi.org/10.1007/s11423-012-9279-9
2. Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. Computers & Education, 125, 413–428. https://doi.org/10.1016/j.compedu.2018.06.023
3. Andrade, A., Delandshere, G., & Danish, J. A. (2016). Using Multimodal Learning Analytics to Model Student Behaviour: A Systematic Analysis of Behavioural Framing. Journal of Learning Analytics, 3(2), 282–306.
4. Anthony, L., Yang, J., & Koedinger, K. R. (2007). Adapting handwriting recognition for applications in algebra learning. Proceedings of the International Workshop on Educational Multimedia and Multimedia Education, 47–56. https://doi.org/10.1145/1290144.1290153
5. Balogun, O. S., Oyelere, S. S., & Atsa'am, D. D. (2019, November). Data analytics on performance of computing students. In Proceedings of the 19th Koli Calling International Conference on Computing Education Research (pp. 1-2).
6. Beck, & Sison, J. (2006). Using Knowledge Tracing in a Noisy Environment to Measure Student Reading Proficiencies. International Journal of Artificial Intelligence in Education, 16(2), 129–143.
7. Birt, J., Clare, D., & Cowling, M. (2019). Piloting Multimodal Learning Analytics using Mobile Mixed Reality in Health Education. 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH), 1–6. https://doi.org/10.1109/SeGAH.2019.8882435
8. Blikstein, P. (2013). Multimodal learning analytics. Proceedings of the Third International Conference on Learning Analytics and Knowledge, 102–106. https://doi.org/10.1145/2460296.2460316
9. Blikstein, P., & Worsley, M. (2016). Multimodal Learning Analytics and Education Data Mining: Using Computational Technologies to Measure Complex Learning Tasks. Journal of Learning Analytics, 3(2), 220–238. https://doi.org/10.18608/jla.2016.32.11
10. Briner, R. B., & Denyer, D. (2012). Systematic Review and Evidence Synthesis as a Practice and Scholarship Tool. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199763986.013.0007
11. Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. Telematics and Informatics, 36, 55–81. https://doi.org/10.1016/j.tele.2018.11.006
12. Chatti, M. A., Muslim, A., & Schroeder, U. (2017). Toward an Open Learning Analytics Ecosystem. In B. Kei Daniel (Ed.), Big Data and Learning Analytics in Higher Education: Current Theory and Practice (pp. 195–219). Springer International Publishing. https://doi.org/10.1007/978-3-319-06520-5_12
13. Cornide-Reyes, H., Noël, R., Riquelme, F., Gajardo, M., Cechinel, C., Mac Lean, R., Becerra, C., Villarroel, R., & Munoz, R. (2019). Introducing Low-Cost Sensors into the Classroom Settings: Improving the Assessment in Agile Practices with Multimodal Learning Analytics. Sensors, 19(15), 3291. https://doi.org/10.3390/s19153291

14. Cowling, M., & Birt, J. (2020). Mixed Reality Multimodal Learning Analytics. In M. A. Peters & R. Heraud (Eds.), Encyclopedia of Educational Innovation (pp. 1–7). Springer Singapore. https://doi.org/10.1007/978-981-13-2262-4_149-1
15. Crossley, S. A., Kyle, K., & McNamara, D. S. (2017). Sentiment Analysis and Social Cognition Engine (SEANCE): An automatic tool for sentiment, social cognition, and social-order analysis. *Behavior research methods*, *49*(3), 803-821.
16. Cruzes, D. S., & Dyba, T. (2011). Recommended Steps for Thematic Synthesis in Software Engineering. 2011 International Symposium on Empirical Software Engineering and Measurement, 275–284. https://doi.org/10.1109/ESEM.2011.36
17. D'Mello, S. K., Olney, A. M., Blanchard, N., Samei, B., Sun, X., Ward, B., & Kelly, S. (2015). Multimodal Capture of Teacher-Student Interactions for Automated Dialogic Analysis in Live Classrooms. Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, 557–566. https://doi.org/10.1145/2818346.2830602
18. Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. Journal of Computer Assisted Learning, 34(4), 338–349. https://doi.org/10.1111/jcal.12288
19. Di Mitri, D., Schneider, J., Specht, M. M., & Drachsler, H. J. (2019, August). Multimodal Pipeline: A generic approach for handling multimodal data for supporting learning. In *First workshop on AI-based Multimodal Analytics for Understanding Human Learning in Real-world Educational Contexts*.
20. Dybå, T., & Dingsøyr, T. (2008). Empirical studies of agile software development: A systematic review. Information and Software Technology, 50(9), 833–859. https://doi.org/10.1016/j.infsof.2008.01.006
21. Fjørtoft, H. (2020). Multimodal digital classroom assessments. Computers & Education, 152, 103892. https://doi.org/10.1016/j.compedu.2020.103892
22. Fouse, A. S., & Hollan, J. D. (2010, August). DataPrism: a tool for visualizing multimodal data. In Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research (pp. 1-4).
23. Forbus, K., Gentner, D., Jee, B., Sageman, B., & Uttal, D. (2009). Drawing on experience: Use of sketching to evaluate knowledge of spatial scientific concepts. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 31, No. 31).
24. Giannakos, M. N. (2016). Orchestrating 21st Century Learning Ecosystems Using Analytics. In CrossLAK (pp. 47-51).
25. Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. International Journal of Information Management, 48, 108–119. https://doi.org/10.1016/j.ijinfomgt.2019.02.003
26. Gomes, J., Yassine, M., Worsley, M., & Blikstein, P. (2013, July). Analysing engineering expertise of high school students using eye tracking and multimodal learning analytics. In Educational Data Mining 2013.
27. Guichon, N. (n.d.). SHARING A MULTIMODAL CORPUS TO STUDY WEBCAM-MEDIATED LANGUAGE TEACHING. Language Learning, 21.
28. Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. Journal of Computer Assisted Learning, 32(2), 139–156. https://doi.org/10.1111/jcal.12123
29. Hariharan, B., Padmini, S., & Gopalakrishnan, U. (2014, May). Gesture recognition using Kinect in a virtual classroom environment. In *2014 Fourth International*

Conference on Digital Information and Communication Technology and its Applications (DICTAP) (pp. 118-124). IEEE.
30. Howison, M., Trninic, D., Reinholz, D., & Abrahamson, D. (2011). The mathematical imagery trainer: From embodied interaction to conceptual learning. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1989–1998. https://doi.org/10.1145/1978942.1979230
31. Hughes, C. A., Maccini, P., & Gagnon, J. C. (2003). Interventions that positively impact the performance of students with learning disabilities in secondary general education classrooms.
32. Hwang, G.-J., Hung, P.-H., Chen, N.-S., & Liu, G.-Z. (2014). Mindtool-assisted in-field learning (MAIL): An advanced ubiquitous learning project in Taiwan. Journal of Educational Technology & Society, 17(2), 4–16.
33. Jarodzka, H., Gruber, H., & Holmqvist, K. (2017). Eye tracking in Educational Science: Theoretical frameworks and research agendas. http://dx.doi.org/10.16910/jemr.10.1.3
34. Junokas, M. J., Lindgren, R., Kang, J., & Morphew, J. W. (2018). Enhancing multimodal learning through personalized gesture recognition. Journal of Computer Assisted Learning, 34(4), 350–357. https://doi.org/10.1111/jcal.12262
35. Kache, F., & Seuring, S. (2014). Linking collaboration and integration to risk and performance in supply chains via a review of literature reviews. Supply Chain Management: An International Journal, 19(5/6), 664–682. https://doi.org/10.1108/SCM-12-2013-0478
36. Kausar, S., Oyelere, S.S., Salal, Y.K., Hussain, S., Cifci, M.A., Hilcenko, S., Iqbal, M.S., Wenhao, Z., Huahu, X. Mining Smart Learning Analytics Data Using Ensemble Classifiers. International Journal of Emerging Technologies in Learning (iJET), [S.l.], v. 15, n. 12, p. pp. 81-102, jun. 2020. ISSN 1863-0383. Available at: <https://online-journals.org/index.php/i-jet/article/view/13455>. Date accessed: 28 Jun. 2020. doi:http://dx.doi.org/10.3991/ijet.v15i12.13455.
37. Kitchenham, B., & Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering.
38. Kuo, F.-O., Yu, P.-T., & Hsiao, W.-H. (2015). Develop and Evaluate the Effects of Multimodal Presentation System on Elementary Student Learning Effectiveness: Within Classroom English Learning Activity. Procedia - Social and Behavioral Sciences, 176, 227–235. https://doi.org/10.1016/j.sbspro.2015.01.465
39. Kusmin, M., Kusmin, K.-L., Laanpere, M., & Tomberg, V. (2019). Engaging Students in Co-designing Wearable Enhanced Learning Kit for Schools. In I. Buchem, R. Klamma, & F. Wild (Eds.), Perspectives on Wearable Enhanced Learning (WELL) (pp. 97–120). Springer International Publishing. https://doi.org/10.1007/978-3-319-64301-4_5
40. Liu, R., Stamper, J., Davenport, J., Crossley, S., McNamara, D., Nzinga, K., & Sherin, B. (2019). Learning linkages: Integrating data streams of multiple modalities and timescales. Journal of Computer Assisted Learning, 35(1), 99–109. https://doi.org/10.1111/jcal.12315
41. Liu, R., & Stamper, J. C. (2017). Multimodal Data Collection and Analysis of Collaborative Learning through an Intelligent Tutoring System. In *MMLA-CrossLAK@ LAK* (pp. 47-52).
42. Mangaroska, K., & Giannakos, M. (2019). Learning Analytics for Learning Design: A Systematic Literature Review of Analytics-Driven Design to Enhance Learning. IEEE

Transactions on Learning Technologies, 12(4), 516–534. https://doi.org/10.1109/TLT.2018.2868673

43. Martin, T., & Sherin, B. (2013). Learning Analytics and Computational Techniques for Detecting and Evaluating Patterns in Learning: An Introduction to the Special Issue. Journal of the Learning Sciences, 22(4), 511–520. https://doi.org/10.1080/10508406.2013.840466

44. Martínez, R., Collins, A., Kay, J., & Yacef, K. (2011). Who did what? Who said that? Collaid: an environment for capturing traces of collaborative learning at the tabletop. Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, 172–181. https://doi.org/10.1145/2076354.2076387

45. Masry, M., & Lipson, H. (2007). A sketch-based interface for iterative design and analysis of 3D objects. ACM SIGGRAPH 2007 Courses on - SIGGRAPH ’07, 31. https://doi.org/10.1145/1281500.1281542

46. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLOS Medicine, 6(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

47. Munoz, R., Schumacher Barcelos, T., Villarroel, R., Guinez, R., & Merino, E. (2018). Body Posture Visualizer to Support Multimodal Learning Analytics. IEEE Latin America Transactions, 16(11), 2706–2715. https://doi.org/10.1109/TLA.2018.8795111

48. Noel, R., Riquelme, F., Lean, R. M., Merino, E., Cechinel, C., Barcelos, T. S., Villarroel, R., & Munoz, R. (2018). Exploring Collaborative Writing of User Stories With Multimodal Learning Analytics: A Case Study on a Software Engineering Course. IEEE Access, 6, 67783–67798. https://doi.org/10.1109/ACCESS.2018.2876801

49. Nouri, J. (2019). Students Multimodal Literacy and Design of Learning During Self-Studies in Higher Education. Technology, Knowledge and Learning, 24(4), 683–698. https://doi.org/10.1007/s10758-018-9360-5

50. Ochoa, X., & Worsley, M. (2016). Augmenting Learning Analytics with Multimodal Sensory Data. Journal of Learning Analytics, 3(2), 213–219. https://doi.org/10.18608/jla.2016.32.10

51. Ochoa, X., Lang, A. C., & Siemens, G. (2017). Multimodal learning analytics. The Handbook of Learning Analytics, 1, 129-141.

52. Okubo, F., Hirokawa, S., Oi, M., Shimada, A., Kentaro, K., Yamada, M., & Ogata, H. (2016). Learning activity features of high performance students. CEUR Workshop Proceedings, 1601, 28–33. https://kyushu-u.pure.elsevier.com/en/publications/learning-activity-features-of-high-performance-students

53. Pantazos, K., & Vatrapu, R. (2016). Enhancing the Professional Vision of Teachers: A Physiological Study of Teaching Analytics Dashboards of Students’ Repertory Grid Exercises in Business Education. 2016 49th Hawaii International Conference on System Sciences (HICSS), 41–50. https://doi.org/10.1109/HICSS.2016.14

54. Pardo, A., & Kloos, C. D. (2011). Stepping out of the box: Towards analytics outside the learning management system. Proceedings of the 1st International Conference on Learning Analytics and Knowledge, 163–167. https://doi.org/10.1145/2090116.2090142

55. Pereira, T., Eldering, J., Rasmussen, M., & Veneziani, A. (2014). Towards a theory for diffusive coupling functions allowing persistent synchronization. Nonlinearity, 27(3), 501–525. https://doi.org/10.1088/0951-7715/27/3/501

56. Perveen, A. (2018). Facilitating Multiple Intelligences Through Multimodal Learning Analytics. Turkish Online Journal of Distance Education, 19(1), 18–30. https://doi.org/10.17718/tojde.382655
57. Pijeira-Díaz, H. J., Drachsler, H., Järvelä, S., & Kirschner, P. A. (2016). Investigating collaborative learning success with physiological coupling indices based on electrodermal activity. Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, 64–73. https://doi.org/10.1145/2883851.2883897
58. Prieto, L. P., Rodriguez Triana, M. J., Kusmin, M., & Laanpere, M. (2017). Smart School Multimodal Dataset and Challenges (CONF). Joint Proceedings of the Sixth Multimodal Learning Analytics (MMLA) Workshop and the Second Cross-LAK Workshop Co-Located with 7th International Learning Analytics and Knowledge Conference; CEUR. https://infoscience.epfl.ch/record/229375
59. Prieto, L. P., Sharma, K., Dillenbourg, P., & Jesús, M. (2016). Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors. Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, 148–157. https://doi.org/10.1145/2883851.2883927
60. Quek, F., Shi, Y., Kirbas, C., & Wu, S. (2002). VisSTA: A tool for analyzing multimodal discourse data. In Seventh International Conference on Spoken Language Processing
61. Qushem, U. B., Zeki, A. M., & Abubakar, A. (2017, August). Successful Business Intelligence System for SME: An Analytical Study in Malaysia. In IOP Conference Series: Materials Science and Engineering (Vol. 226, No. 1, p. 012090).
62. Qushem, U. B., Zeki, A. M., Abubakar, A., & Akleylek, S. (2017, October). The trend of business intelligence adoption and maturity. In 2017 International Conference on Computer Science and Engineering (UBMK) (pp. 532-537). IEEE.
63. Raca, M., Tormey, R., & Dillenbourg, P. (2014). Sleepers' lag—Study on motion and attention. Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, 36–43. https://doi.org/10.1145/2567574.2567581
64. Richard, G. T., & Giri, S. (2019). Digital and Physical Fabrication as Multimodal Learning: Understanding Youth Computational Thinking When Making Integrated Systems Through Bidirectionally Responsive Design. ACM Transactions on Computing Education, 19(3), 1–35. https://doi.org/10.1145/3243138
65. Sadik, A. (2008). Digital storytelling: A meaningful technology-integrated approach for engaged student learning. Educational Technology Research and Development, 56(4), 487–506. https://doi.org/10.1007/s11423-008-9091-8
66. Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—Review of literature. Research and Practice in Technology Enhanced Learning, 14(1), 11. https://doi.org/10.1186/s41039-019-0105-4
67. Saqr, M., Viberg, O., Nouri, J., & Oyelere, S. S. (2020). "Multimodal temporal network analysis to improve learner support and teaching", In Companion Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK20).
68. Schlömer, T., Poppinga, B., Henze, N., & Boll, S. (2008). Gesture recognition with a Wii controller. Proceedings of the 2nd International Conference on Tangible and Embedded Interaction, 11–14. https://doi.org/10.1145/1347390.1347395
69. Schneider, B., Reilly, J., & Radu, I. (2020). Lowering Barriers for Accessing Sensor Data in Education: Lessons Learned from Teaching Multimodal Learning Analytics to Educators. Journal for STEM Education Research. https://doi.org/10.1007/s41979-020-00027-x

70. Seuring, S., & Gold, S. (2012). Conducting content-analysis based literature reviews in supply chain management. Supply Chain Management: An International Journal, 17(5), 544–555. https://doi.org/10.1108/13598541211258609

71. Shankar, S. K., Prieto, L. P., Rodriguez-Triana, M. J., & Ruiz-Calleja, A. (2018). A Review of Multimodal Learning Analytics Architectures. 2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT), 212–214. https://doi.org/10.1109/ICALT.2018.00057

72. Sharma, K., Papamitsiou, Z., & Giannakos, M. (2019). Building pipelines for educational data using AI and multimodal analytics: A "grey-box" approach. British Journal of Educational Technology, 50(6), 3004–3031. https://doi.org/10.1111/bjet.12854

73. Shoumy, N. J., Ang, L. M., Seng, K. P., Rahaman, D. M., & Zia, T. (2020). Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals. *Journal of Network and Computer Applications*, *149*, 102447.

74. Shadiev, R., & Sintawati, W. (2020). A review of research on intercultural learning supported by technology. *Educational Research Review*, 100338.

75. Spens, K. M., & Kovács, G. (2006). A content analysis of research approaches in logistics research. International Journal of Physical Distribution & Logistics Management, 36(5), 374–390. https://doi.org/10.1108/09600030610676259

76. Stevenson, M., Hedberg, J., Highfield, K., & Diao, M. (2015). Visualizing Solutions: Apps as Cognitive Stepping-Stones in the Learning Process. Electronic Journal of E-Learning, 13(5), 366–379.

77. Tamura, K., Lu, M., Konomi, S., Hatano, K., Inaba, M., Oi, M., Okamoto, T., Okubo, F., Shimada, A., Wang, J., Yamada, M., & Yamada, Y. (2019). Integrating Multimodal Learning Analytics and Inclusive Learning Support Systems for People of All Ages. In P.-L. P. Rau (Ed.), Cross-Cultural Design. Culture and Society (Vol. 11577, pp. 469–481). Springer International Publishing. https://doi.org/10.1007/978-3-030-22580-3_35

78. Tan, S., O'Halloran, K. L., & Wignell, P. (2016). Multimodal research: Addressing the complexity of multimodal environments and the challenges for CALL. ReCALL, 28(3), 253–273. https://doi.org/10.1017/S0958344016000124

79. Turk, M. (2014). Multimodal interaction: A review. *Pattern Recognition Letters*, *36*, 189-195.

80. Vujovic, M., Tassani, S., & Hernández-Leo, D. (2019). Motion Capture as an Instrument in Multimodal Collaborative Learning Analytics. In M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider (Eds.), Transforming Learning with Meaningful Technologies (Vol. 11722, pp. 604–608). Springer International Publishing. https://doi.org/10.1007/978-3-030-29736-7_49

81. W. Lewis, & Johnson. (2010). Serious Use of a Serious Game for Language Learning. International Journal of Artificial Intelligence in Education, 20(2), 175–195. https://doi.org/10.3233/JAI-2010-0006

82. Wang, S.-P., & Chen, Y.-L. (2018). Effects of Multimodal Learning Analytics with Concept Maps on College Students' Vocabulary and Reading Performance. Educational Technology & Society, 21(4), 12–25.

83. Weinland, D., Ronfard, R., & Boyer, E. (2006). Free viewpoint action recognition using motion history volumes. Computer Vision and Image Understanding, 104(2), 249–257. https://doi.org/10.1016/j.cviu.2006.07.013

84. Wong, L.-H. (2012). A learner-centric view of mobile seamless learning. British Journal of Educational Technology, 43(1), E19–E23. https://doi.org/10.1111/j.1467-8535.2011.01245.x
85. Worsley, M., & Blikstein, P. (2011, July). What's an Expert? Using Learning Analytics to Identify Emergent Markers of Expertise through Automated Speech, Sentiment and Sketch Analysis. In EDM (pp. 235-240).
86. Worsley, M., & Blikstein, P. (2015). Leveraging multimodal learning analytics to differentiate student learning strategies. Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15, 360–367. https://doi.org/10.1145/2723576.2723624
87. Worsley, M., Abrahamson, D., Blikstein, P., Grover, S., Schneider, B., & Tissenbaum, M. (2016). Situating multimodal learning analytics. In Proceedings of International Conference of the Learning Sciences, ICLS (Vol. 2, pp. 1346-1349).
88. Xue, Y., Ju, Z., Xiang, K., Chen, J., & Liu, H. (2018). Multimodal human hand motion sensing and analysis—a review. *IEEE Transactions on Cognitive and Developmental Systems*, *11*(2), 162-175.
89. Yilmaz, A., & Shah, M. (2005). Recognizing human actions in videos acquired by uncalibrated moving cameras. Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, 1, 150-157 Vol. 1. https://doi.org/10.1109/ICCV.2005.201
90. Zaiane, O. (2001). Web Usage Mining for a Better Web-Based Learning Environment. ERA. https://doi.org/10.7939/R3736M20P
91. Zimmerman, P. H., Bolhuis, J. E., Willemsen, A., Meyer, E. S., & Noldus, L. P. (2009). The Observer XT: A tool for the integration and synchronization of multimodal signals. Behavior research methods, 41(3), 731-735.

APPENDICES

Appendix 1: Summary of methods found within Primary Studies

Indicators	Potential methods	Primary Studies
Behavioral	Integration of activity logging, eye-tracking; and video processing	D26
	Complex learning task, study on motion and attention, Designing an automated assessment, modelling student behavior	D14
	Example based reasoning, and principle-based reasoning	D15
	Modeling Student Behavior	D5
	Virtual, Logical Mathematical, Spatial visual, Kinetic, Musical, Interpersonal, Intrapersonal	D9
Educational	Learning language using vocabulary and through reading comprehension	D7
	Multimodalities for learning and learning theories	D10
	Setting up a classroom that can generate multimodal dataset	D29
	face to face conversation, language learning experiences	D27
	Digital texts writing, taking pictures, creating audio and video recordings, modal practices	D30
Physiological	Click-streams, Keystrokes, EEG, Eye-tracking, Gaze, Video, Wristband	D4
	Eye Tracking measurement, EEG measurement	D24
	From text to gestures, gesture, voice, text, video, and audio	D25
	User-defined gestures, and accessing data of skeleton positions, kinematics features	D8
	Low-cost sensors	D11
	Physical indicators such as gaze direction, the distance between learners and the speed of movement/reactions	D19

	Physiologies data on public speaking and student engagement	D1
	Sensor data (Eye-tracking, EEG, Facial expressions, wristband data)	D6
Digital	Audio and video recordings for instructors and apprentices.	D28
	Audio interviews and video data analysis using a deductive constant comparison approach	D20
	Audio recording	D3
	Interactive whiteboard (IWB) integration within Multimodal presentation system	D18
	Multiple modalities data and contexts (individual vs. collaborative classroom learning).	D12
	Observational dashboard illustrating MMLA, data recording and summative analytics	D17
	Sensors kits	D23
	Spatial Learning Analytics, Multimodal Mixed reality learning	D2

Appendix 2: Working Directory in ATLAS.ti Software