Article

A Learning Analytics Theoretical Framework for Virtual Reality Instructional Applications

Athanasios Christopoulos [1*], Nikolaos Pellas [2], Mikko-Jussi Laakso [1,]

[1] Center for Learning Analytics, University of Turku, 20014 Turku, Finland, atchri@utu.fi; milaak@utu.fi

[2] Department of Communication and Digital Media, University of Western Macedonia, 52100 Kastoria, Greece, aff00192@uowm.gr

* Correspondence: atchri@utu.fi

Abstract: The commercial popularity of Virtual Reality attracted educators' interest and brought new opportunities to the educational landscape. At the same time, Learning Analytics emerged with the promise to revolutionise the traditional practices by introducing ways to systematically assess and improve the effectiveness of instruction. However, the collection of 'big' educational data is mostly associated with web-based platforms as they offer direct access to learners' activities with minimal effort. On the antipode, the nature of VR limits the opportunities for such data collection. Hence, in the context of this work, we present a four-dimensional theoretical framework, that accounts the information that can be gathered from VR-supported instruction, and propose a set of structural elements which can be utilised for the development of a Learning Analytics prototype system. The outcomes of this work are expected to support practitioners to maximise the potential of their interventions and provide inspiration for new ones.

Keywords: Virtual Reality; Learning Analytics; STEM Education; Instructional Design.

1. Introduction

As newly statistical data shows STEM (Science, Technology, Engineering, and Mathematics) education is expanding rapidly in most developed countries and thus, the necessity to provide learners with well-designed instructional contexts becomes even more imperative [1]. This is aligned to the outcomes of studies [2] which stress on the importance of assisting learners to cultivate the acquired knowledge in-depth but the difficulties that instructional designers face when preparing laboratory exercises (including experiments and practice-based tasks) pertinent to the STEM fields cannot be disregarded either. For instance, field-based experiments require complex transportation to different locations whereas some of the laboratory-based tasks may be too dangerous (e.g. an electric shock caused due to the incorrect wiring of electrical wires in an electrical engineering course) or too expensive to perform in the real world (e.g. use of hard-to-acquire specialised equipment). In addition, the limited training, or the lack of awareness that students may have on matters related to lab safety and security further increase the risks for injuries or even fatalities [2,3].

Hence, in order to prevent such issues from occurring, the presence of the instructor is essential but, even then, the limited attention that individuals receive—e.g. due to the time-management constraints—has been reported as a factor which causes negative emotions and behaviour (e.g. frustration, dissatisfaction) [3,4]. Such shortcomings are linked to serious complications towards the theoretical knowledge development or the conceptual experience advancement—especially when abstract topics are under consideration—and may hinder students' confidence to apply such practices in the future [5].

A proposed solution to eliminate the impact of such drawbacks considers the adoption and integration of interactive technologies which can make the educational processes more efficient and effective [2,6]. This is also aligned to the belief that wants employees to be undertaking frequent training in simulated environments [7,8]. Nevertheless, the utilisation of digital solutions solves only the health and safety issues but does not guarantee the knowledge development or advancement [9].

As a result, the need to analyse the potential and the shortcomings of the computer-supported instructional strategy becomes also crucial.

However, understanding how to maximise the effectiveness of the instructional design strategy, based on the theory that each STEM subject imposes, is a complicated and demanding process. The solution to this matter is identified on the potential of the technological tools per se as they offer opportunities for the collection of large datasets which can provide information related to the educational context, the utilised instructional strategy and the behavior of the learners. Hence, by collecting and interpreting such information, content, and instructional designers as well as researchers and educators can increase the effectiveness of the learning strategy, facilitate the learning process, and prevent the development of misconceptions [10,11].

2. Background

2.1 Virtual Reality in STEM Education

Heim [12] argued over the potential of Virtual Reality (VR) by attributing its added value to three fundamental elements: interactivity, immersiveness and information intensity. Despite the passing of time since this claim was made, the empirical studies that have been performed ever since not only confirmed its validity but, also, revealed the additional benefits of such 'tools' can bring to the educational scenery [3,13]. Therein, in the context of this work, we adopt the definition that Gigante [14] coined, which defines VR as the computer-supported setting that enhances the real-world experience through the provision of multi-aesthetic stimuli (e.g. visual, audio, motion). In addition, we expand the notion of this definition by providing a brief overview of the VR-supported (educational) settings that are available to date (e.g. *room-scale VR* such as CAVE, *standalone-VR* such as Oculus Rift, HTC Vive, and *mobile-VR* such as Samsung Gear VR, Google Cardboard).

The aforementioned setups promote different levels of embodiment (immersion) and offer variable opportunities for knowledge acquisition and construction (information intensity) whereas, the inclusion of haptic sensors, brings additional opportunities for interactivity and thus, engagement. In a sense, this is what differentiates VR from other educational technology tools—i.e. the opportunity offered to learners to undertake both passive learning (e.g. observation of natural phenomena) and active learning activities (e.g. laboratory-related experiments) without spatiotemporal or time constraints. However, none of the above would have been possible without the rapid technological advancement of computing devices and the vast evolution of the VR [15].

The integration of VR in the educational landscape is already playing a significant role as it has facilitated the application of contemporary instructional methods which enable learners to immerse themselves in the subject under investigation and thus, develop the cognitive strategies (e.g. problem-solving, critical thinking, creativity) that are essential in the 21st century [10,3]. Aligned to the notion of this claim, a common observation across the STEM education disciplines can be made with regard to the nature of the programs and the respective interventions which follow primarily the principles of the experiential learning model. It, therefore, comes with no surprise why such tremendous efforts have been made to integrate immersive technologies to every education level which involves matters related to the STEM disciplines. This is also in line with the conclusions that Pellas et al. [11] have drawn which attribute the successful integration of such technologies to the high degree of embodiment that users develop when interacting with the (digital) objects that have been customised in accordance to their personal needs and demands.

2.2 Instructional Design in Virtual Reality

Instructional design methods are comprised of strategies (e.g. instructor-guided, self-directed) and techniques (e.g. collaborative) which aim at helping educators to contextualise the learning process and learners to link the concepts under investigation with their prior knowledge and experiences [16,17]. In other words, instructional design helps learners to understand what kind of information is provided within a specific context, how this information can be translated into knowledge acquisition, and how the constructed knowledge can be applied more effectively into

practice [18,19]. The aforementioned processes are directly linked to the learning performance, which concerns the range of fluctuations in learners' knowledge development or behavior during the different stages of the intervention, and the learning outcomes (e.g. satisfaction, achievements, acquired knowledge/skills, competencies) that learners are expected to achieve at the end of the intervention [20].

The findings from the VR-supported educational studies are well-documented in the literature and so are the benefits that this technology brings to the learning process. Below, we provide a brief summary of the most important elements that influence the respective educational practices:

- *Student-centered learning*: Aligned to the principles of (Social) Constructivism and Constructionism, the visually rich environment and the experimental nature of VR enable students to develop strong mental representations of the information sources through hands-on and collaborative activities [21,22].
- *Self-directed learning*: By exploiting the affordances that the 3-Dimensional (3D) element offers, learners can investigate hypothetical and abstract concepts—which are difficult or even impossible to examine in the real-world—without spatial, time, and/or geographical boundaries [13,23,24].
- *Self-regulated learning*: By immersing learners in situations similar to the real-life context enables them to self-regulate the learning process in accordance with the challenges and difficulties they are facing [25,26].

2.3 Learning Analytics

Educational practitioners and scholars have attempted to define Learning Analytics (LA) from different perspectives and viewpoints. For instance, a portion of researchers [27-29] account them as an alternative method to gather student-generated data which can be utilised to provide personalised learning experiences. Others [30,31] set the focus on the patterns that can be developed from the students' learning behaviours so as to inform the future instructional design decisions. Long and Siemens [32] have proposed a definition which considers and rounds up the aforementioned perspectives by suggesting that LA is a method to collect longitudinal educational data and a process that utilises the collected data to optimise learning and the environment in which it occurs.

Regardless of the chosen definition, researchers from different disciplines and fields (e.g. Applied Statistics, Artificial Intelligence, Data Science) are working in collaboration so as to identify the diverse learning needs that students have and improve the present educational practices [32]. To achieve this goal, large sets of heterogeneous data—from different educational levels and sources—are collected, explored, and analysed using Machine Learning (ML) models. The outcomes of this process provide diverse, but equally useful, feedback to the educational stakeholders with regard to learners' performance, the shortcomings of the utilised instructional approach and the inadequacies of the course under investigation [33,34].

The added value of LA can be examined from different perspectives and points of view. Below we present the key-areas that LA influence, after considering the interests and the needs that the various stakeholders (e.g. learners, educators, instructional designers, policy makers) have:

- *Learners*: Alter the learning habits by identifying patterns and paths that can support the attainment of the learning objectives and ensure the achievement of the predefined goals.
- *Educators*: Improve the quality of teaching based on real-time and summative data that mirror learners' performance, involvement, and engagement over the course of the time.
- *Instructional designers*: Increase the quality of instruction based on the analysis of the elements that have been utilised the most, the feedback from the students on the provided interventions and the comments of the teachers.
- *Policy makers*: Develop clear and accurate awareness of the current and future tendencies so as to inform the subsequent decisions and policy making.

2.4 Rationale and Purpose

Accounting the above, the desirable outcome of this work is the provision of a conceptual framework that will offer educators and instructional designers suggestions related to the data that can be collected from different VR-supported teaching interventions and recommendations on the connections that may exist amongst them. Aligned to the end-goal of this project the main objectives of this work are split into three consecutive stages. In the context of this manuscript, we elaborate on and discuss the perquisites that characterise the requirements of the first stage as presented below:

1. Development of a theoretical framework which accounts the research gaps that have been identified from the examination of the relevant literature.
2. Proposition of an approach that can determine un/-trained students and trainee practitioners from different STEM fields while uncovering the most relevant variables related to this classification.
3. Identification of the most efficient ML models for the analysis of the error-related behaviors and the determination of the patterns that will improve the provided instruction.

In the second stage, we plan on using the proposed framework to design a functional prototype of a VR learning tool which will be applied in various iterations within the STEM education fields for evaluation purposes. Finally, in the third stage, it is expected that a complete training and assessment session will be provided by utilising solely the recommendations of the LA measurements.

To achieve the abovementioned objectives, we will use different quantitative approaches on the basis of which the student models will be shaped from the information that can be retrieved from the VR application and the companion Learning Management System (LMS).

The proposed methodology is also comprised of three parts which are:

1. Use of statistical analysis models to classify trained and untrained students after collecting data from several VR-supported training sessions. The initial dataset will include information related to the course design, the learners' profile, and the interactions that the students had during the training task. For the construction of the final model it is expected that several statistical models will be considered so as to increase the prediction accuracy and reliability of the results.
2. Use of different Feature Importance Analysis (FIA) methods to identify the most effective classifiers per task, the relevant variables, and their impact on determining students' success or failure for the task under consideration.
3. Use of an Exploratory Data Analysis (EDA) tool to identify the relationships between the recorded errors. To this end, the clustered information will be exported visually so as to develop different hypotheses related to the underlying reasons that drive these relations. For the visual representation we are planning to apply the LA guidelines that Baker and Yacef [35] have proposed.

The use of ML techniques will benefit all the research and development stages as it will enable us to answer the questions that educators usually raise and the concerns that instructional designers usually have. A few indicative examples of such matters are provided below:

*   How to assess the skill cultivation between novice and expert students in VR-supported STEM training scenarios?
*   How to select the most appropriate design elements and instructional concepts, so as to increase the efficacy and efficiency of the VR-based application, in accordance to the difficulty level of the topic and the abilities of the learners?
*   How to aggregate the best practices, so as to perform error diagnosis, when an LMS is used in conjunction?
*   How to provide timely support to low-performing or additional opportunities for development to high-performing students?

3. The Theoretical Framework

3.1 Framework Analysis

According to Hevner *et al.* [36] the Design Science research methodology is amongst the most appropriate methods for the development of an Information Technology or Information System artifact which, in this case, is a framework. The main principle of this approach suggests the deconstruction of important problems on the basis of which sound (technical) solutions can be developed. Therein, during the extensive literature review that has been conducted in the context of this work, we identified a set of issues that have not been addressed and therefore, provided the foundation to design the main requirements of this framework as presented below:

1. LA models are applied primarily to data that originate from LMS without considering alternative or supplementary tools.
2. The main sources for data collection take into account the information that derive either from the technological or the pedagogical perspective of the tool/intervention but disregard partially or even completely the psychological one.
3. Relevant studies examine the correlations that may exist between a finite set of dependent variables (e.g. demographics, credits, grades) against non-classified parameters which are relevant to specific contexts and fields. This endangers the essence and further evolution of LA as it prevents the collection and sharing of large and homogenous datasets.
4. By cross-examining the latest (systematic) literature reviews it became apparent that there is still lack of a universally accepted comprehensive framework and/or system capable of providing the involved stakeholders with suggestions on what kind of data should collect or recommendations on how to interpret such data so as to evaluate specific elements and improve their practices.

Hence, the proposed framework (Figure 1) blends the aforementioned points by integrating the use of LA models for the processing and cross-examination of the information related to: the technical affordances of the utilised tools, the instructional design choices that practitioners make, the psychological elements that influence the learning process.
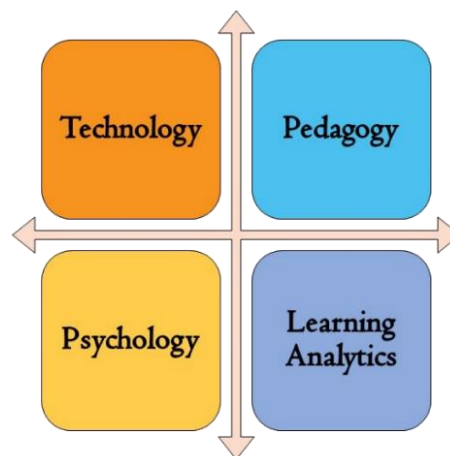
Figure 1. The four-dimensional framework for VR-based Learning Analytics.

3.2 Design Decisions

The information that can be collected from each category are illustrated in Figure 2 (abstract level) and elaborated further in the following sections.

In the first category (*Technology*) we consider matters related to the design and development of the VR-supported interventions, such as:

- the *software toolkits* utilised for the development of the VR application (e.g. Unity, Vuforia, Maya, .Net, Photoshop)

- the *specifications* of the *hardware* equipment utilised for the conduct of the interventions (e.g. smartphone, tablet, laptop, desktop PC, Head-Mounted Display)
- the *type of the VR* approach (e.g. HMD-based, CAVE, 360°) and the companion equipment (e.g. QR code sheets, VR-enabled laboratory handbooks or discipline-related specialized equipment)
- the *supplementary resources* that may be required for the conduct of the intervention (e.g. multimedia resources, web-based educational platforms, 3D models)

In the second category (*Pedagogy*), we contemplate the potential connection across the instructional decisions that practitioners make when designing the educational activities, such as:

- the *learning theories* on the basis of which the design of the intervention will rely on (e.g. Constructionism, Cognitivism, (Social) Constructivism or Embodied cognition),
- the *instructional strategies* (learning models) that gravitate the didactic essence of the respective theories (e.g., activity-based, experiential, collaborative, situated, problem-based, game-based, agent-based learning) and instructional techniques utilised for the conduct of the intervention (e.g. lecture, demonstration, seminar, tutorial, case study) [16], and
- the *evaluation focus points* related to the effectiveness and efficiency of the application, the intervention, and the instructional approach (e.g. learners' performance, learning outcomes, learning gains).
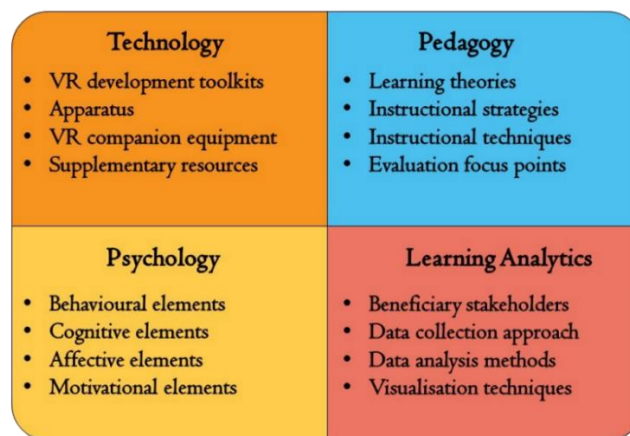
Figure 2. Classification parameters for each dimension.

In the third category (*Psychology*), we consider the psychological dimensions that are connected to the pedagogical dimensions, engagement, and influence the learning process, such as:

- the *behavioral* elements (e.g. the impact/effect of reinforcement),
- the *cognitive* elements (e.g. attention and memory span, problem-solving ability),
- the *affective* elements (e.g. interest, attachment, satisfaction, emotional investment, degree of arousal), and
- the *motivational* elements (e.g. self-belief, self-regulation, self-efficacy, self-goals, self-concept, self-esteem, situational interest)

In the fourth category (*Learning Analytics*), we consider the steps that are related to the data gathering and analysis process, such as:

- the *information* that can be collected from the different stakeholders (e.g. administrators, educators, students, assessment tools),
- the *data collection approach* which includes information related to the research method (e.g. experimental, quasi-experimental, non-experimental) and the research methods utilised (e.g. qualitative, quantitative, mixed),
- the *data analysis approach* which includes the use and combination of different methods (e.g. Item Response Theory, Cognitive Diagnosis, Evolutionary Algorithms) and Educational Data Mining models (e.g. Decision Tree, Naive Bayes Classifier, *k*-Nearest Neighbor), and
- the *data visualisation models* for the dissemination of the processed data (e.g. graphs/charts, scatterplots, sociograms, tag clouds, signal lights).

3.3. Tools overview

Aligned to the intended goal, in this section, we provide an overview of the functional requirements and specifications of the proposed system (Figure 3).
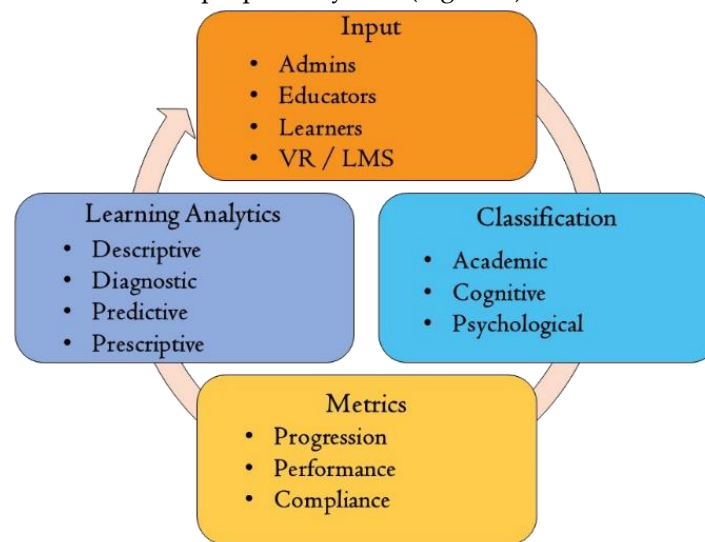
Figure 3. Overview of the system data processing approach.

To facilitate the data collection, analysis and interpretation process, we classify the information that can be retrieved from the different stakeholders (Table 1) and further associate them with a set of dimensions (i.e. academic, cognitive, and psychological) and metrics (i.e. progression, performance, and compliance) (Figure 3). Moreover, we provide a set of examples related to the proposed dimensions and their respective interrelations.

Table 1. Indicative examples of the data that can be gathered from the stakeholders.

Stakeholder	Data	Examples
Administrators	Demographics	gender, age, family status
	Educational	history, pathway, study field, level
	Academic	enrollments, dropouts, credits
Educators	Curriculum	materials, learning objectives/outcomes
	Activity	tasks, assignments, assessments
Learners	Personal info	interests, motivation, expectations
	Deliverables	coursework, crafts
VR / LMS	Sensory	motion, visual, auditory, haptic
	Traces	attendance, engagement, resources' use

For instance, to measure matters related to the academic dimension, the primary data collection will include information related to students' management skills (e.g. use of resources), their prior knowledge with the scientific subject (STEM) and experience with the digital learning tools (e.g. VR, LMS) as well as their attitude towards the learning process (e.g. attendance, participation, interaction with the peers) and their learning competence (i.e. time to develop and integrate the acquired knowledge and skills). The primary data sources will include information originating from the students' interaction with the VR application and the LMS as well as self-reported cues related to their short- and long-term plans or goals (e.g. academic, personal, professional, monetary).

As regards the measurement of matters related to the cognitive dimension, self-reported data related to the ways that students regulate their efforts (e.g. strategies, tactics, habits) will be collected using validated instruments and further correlated with their learning outcomes (including the identification of misconceptions and knowledge gaps) using Artificial Intelligence techniques and ML models.

Finally, for the measurement of matters related to the psychological dimension, the focus will be set on learners' behavioral patterns which will be recorded from the onboard sensors of the devices

that will be utilised for the conduct of the interventions (e.g. smartphones, tablets) and other wearables (e.g. HMDs). Such data will include information related to learners' interactions (e.g. app use log, visual attention span, emotion recognition, textual communication records) and mobility patterns (e.g. frequency and duration of time spent at various locations while interacting with the learning tool).

Table 2. The analytics stages as described by Howson *et al.* [37]

Analytics	Description	Outcome
Descriptive	What happened?	Insights into historical patterns of behavior/performance.
Diagnostic	Why did it happen?	In-depth evaluation of the examined data.
Predictive	What could happen in the future?	Identify trends in data and predict future behavior.
Prescriptive	How should we respond in future events?	Generate recommendations and make decisions based on the computational findings of algorithmic models

For the interpretation of the collected data, we will use the model that Howson *et al.* [37] propose (Table 2), along with the literature recommendations related to the implementation of ML models in LA practices (Table 3).

Table 3. Examples of the ML models to address main issues concerning relevant educational topics.

Goals	ML models	References
Feedback to educators' & instructional designers' scenarios.	Decision Trees, Random Forest	[38,39]
Investigation of learners' behavior during & after the VR-supported intervention.	Naïve Bayes	[40]
Course adaptation & learning recommendations based on learners' behavior.	Decision Trees, Random Forest	[41,42]
Assessment of the VR-supported learning material & content.	Decision Trees, Random Forest, Naïve Bayes	[43,44]
Prediction of student's learning performance.	Decision Trees, Logistic Regression, Support Vector Machines	[45,46]

4. Contribution and implications

4.1. Conceptual implications

The current study also contributes to the existing body of literature by providing a range of parameters which stream from the proposed theoretical framework and could improve the teaching and learning practices. These parameters are:

1. Orchestration of instruction by teachers and reflection on the utilised strategies from the originals available to them.
2. Evaluation methods to assess not only the students' performance but, also, that of teachers with regard to the mode of operation and practices followed in both formal and informal contexts.
3. Provision of personalised suggestions and appropriate structures to support the implementation of similar scenarios in the future.
4. Development of deep understanding on the core elements that influence the educational process and adaptation of the educational resources based on the needs and interests of the students.
5. Assessment of the course curriculum with particular focus on the parameters that affect the success and the effectiveness of the interventions in STEM training tasks.

6.   Support from the administration for the reshaping of the educational units and allocation of financial resources for the development of VR applications in formal teaching conditions.

4.2. Theoretical implications

We provide a number of theoretical implications with regard to the development of a universal LA system tailored to the VR configuration setups. The following points are expected to guide our future decisions but, also, provide instructions to those researchers, educators, and instructional designers who are willing to contribute towards this effort:

1.   The decisions related to the data collection should be driven by the principles of the applied instructional design method. Hence, the involved stakeholders are encouraged to provide detailed information about the utilised instructional approaches, the educational subjects that were under investigation, and the analysis methods that have been followed for the examination of the correlations. In doing so, the repetition of the intervention to similar contexts will be facilitated and that will support the future research efforts to validate (collectively) the gathered information so as to develop well-grounded generalisations.
2.   The potential of interactions should be examined holistically and not just unilaterally (i.e. both between the users and the VR system and among the users themselves). Under this consideration, we recommend the cross-examination and correlation clustering of different pedagogical and psychological elements using ML models so as to aid the development of prototype profiles and allow the systematic mapping of the factors that influence students' outcomes and performance.
3.   The classification of the gathered information should be done in accordance to the areas of interest of the different beneficiaries (i.e. administrators, instructional designers, teachers, students) and the outcomes should be disseminated in accordance to the data analytics maturity scale that Howson *et al.* [37] proposed (i.e. descriptive, diagnostic, predictive, and prescriptive analytics). In doing so the involved stakeholders will be able to determine the suitability and the effectiveness of the intervention and thus, perform any adjustments that may be required prior to designing or implementing new interventions.

4.3. Practical implications

The inadequacy of the literature to provide recommendations with regard to the data types that can be collected from immersive technologies as well as the absence of a distributed system—capable of collecting, analysing and determining the appropriateness and effectiveness of the VR-supported interventions in STEM education—motivated this initiative on the basis of which we provide a set of practical implications which could help developers to better understand the functional requirements of such VR-supported LA systems:

1.   VR technology produces a huge amount of data but not all of them are meaningful to the context of educational studies. For exemplification purposes we summarise the data sources that are pertinent to the aim of the proposed LA system followed by some indicative examples:
     - visual (e.g. eye motion tracking)
     - auditory (e.g. pitch/intensity of the environmental noise levels)
     - haptic (e.g. movement, rotation, force)
     - network (e.g. packet loss, time delay)
2.   The essence of the educational VR applications relies on the provision of immediate feedback which offers answer-revision opportunities and leads to errorless learning. In the same vein, a comprehensive implementation of a visual LA dashboard is expected to influence the learning dynamics (e.g. motivation, competitiveness, goal orientation) and impact positively learners' outcomes, achievements, and performance.

5. Discussion and Conclusion

The potential of VR in STEM education has already attracted practitioners' interest by demonstrating its power to support the conduct of safe, interactive, and engaging learning experiences. At the same time, the LA domain is gaining more and more ground as it has immense potential to improve the teaching and learning practices [24,28,29]. However, while a substantial body of literature provides guidance and recommendations with regard to the potential of LA in various educational settings, the attempts to integrate LA in the context of immersive technologies are limited and scarce.

Therein, in an effort to tackle this literature limitation, we outline the foundations of a four-dimensional theoretical framework by accounting the multifaceted layers of the learning process (Technology, Pedagogy, Psychology) as well as the respective parameters and constructs that should be taken into account when gathering the so-called 'big data' (LA). To this notion, we pay particular attention on the impact that the different instructional strategies and methods have on students [16] and the opportunities they bring to create personalised learning patterns. In addition, we set the focus on the importance of assessing learners' actions and interactions in real-time so as to provide timely feedback and feedforward. Finally, considering the complex difficulties that the global education system is currently facing due to the COVID-19 outbreak, it becomes imperative to propose and plan on the development of a prototype that will support educators and instructional designers to create engaging and cross-disciplinary online/remote learning activities both for formal and informal contexts.

Conclusively, by highlighting these conceptual design and developmental elements, we envision that researchers, educators and educational technology entrepreneurs will further consider these relations—when evaluating the potential of the utilised instructional approach—so as to take full advantage of the data that can be collected from such tools and platforms and thus, reach the maximum potential of the VR technology especially when applied to STEM education subjects.

References

1. STEM Education Data. Available online: https://nsf.gov/nsb/sei/edTool/explore.html (Accessed 6 Oct 2020).
2. Pellas, N.; Kazanidis, I.; Konstantinou, N. *et al.* Exploring the educational potential of three-dimensional multi-user virtual worlds for STEM education: A mixed-method systematic literature review. *Educ Inf Technol* **2017**, *22*, 2235–2279. https://doi.org/10.1007/s10639-016-9537-2
3. Pellas, N.; Kazanidis, I.; Palaigeorgiou, G. A systematic literature review of mixed reality environments in K-12 education. *Educ Inf Technol* **2020**, *25*, 2481–2520. https://doi.org/10.1007/s10639-019-10076-4
4. Potkonjak, V.; Gardner, M.; Callaghan, V. *et al.* Virtual laboratories for education in science, technology, and engineering: A review. *Comp & Educ* **2016**, *95*, 309-327. https://doi.org/10.1016/j.compedu.2016.02.002
5. Makransky, G.; Borre-Gude, S.; & Mayer, R. Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *J Comput Assist Learn* **2019**, *35*(6), 691-707. https://doi.org/10.1111/jcal.12375
6. Christopoulos, A.; Conrad, M.; Shukla, M. Increasing student engagement through virtual interactions: How?. *Virtual Real* **2018**, *22*(4), 353-369. https://doi.org/10.1007/s10055-017-0330-3

7.   Ayala García, A.; Galván Bobadilla, I.; Arroyo Figueroa, G. *et al.* Virtual reality training system for maintenance and operation of high-voltage overhead power lines. *Virtual Real* **2016**, *20*(1), 27-40. https://doi.org/10.1007/s10055-015-0280-6

8.   Villagrá-Arnedo, C. J.; Gallego-Durán, F. J.; Llorens-Largo, F. *et al.* Improving the expressiveness of black-box models for predicting student performance. *Comput. Hum. Behav.* **2017**, *72*, 621-631. https://doi.org/10.1016/j.chb.2016.09.001

9.   Christopoulos, A.; Conrad, M.; Shukla, M. Between virtual and real: Exploring hybrid interaction and communication in virtual worlds. *Int J Soc. Media Interact. Learn. Environ*. **2016**, *4*(1), 23-42. https://doi.org/10.1504/IJSMILE.2016.075038

10.   Christopoulos, A.; Kajasilta, H.; Salakoski, T.; Laakso, M.J. Limits and virtues of educational technology in elementary school mathematics. *J. Educ. Technol. Syst.* **2020**, *49*(1), 59-81. https://doi.org/10.1177%2F0047239520908838

11.   Pellas, N.; Fotaris, P.; Kazanidis, I. *et al.* Augmenting the learning experience in primary and secondary school education: a systematic review of recent trends in augmented reality game-based learning. *Virtual Real* **2019**, 23, 329–346. https://doi.org/10.1007/s10055-018-0347-2

12.   Heim, M. *Virtual realism*. Oxford University Press: Oxford, UK, 1998.

13.   Wolfartsberger, J. Analyzing the potential of Virtual Reality for engineering design review. *Autom. Constr.* **2019**, *104*, 27-37. https://doi.org/10.1016/j.autcon.2019.03.018

14.   Gigante, M. Virtual reality: Definitions, history and applications. In *Virtual Reality Systems*; Earnshaw, R.A., Gigante, M.A., Jones, H. Eds.; Academic Press: USA, **1993.** https://doi.org/10.1016/C2009-0-22372-5

15.   Virtual reality (VR) – statistics & facts. Available online: https://www.statista.com/topics/2532/virtual-reality-vr (Accessed on 6 October 2020).

16.   Akdeniz, C. Instructional process and concepts. In *Theory and Practice: Improving the teaching process*. Springer: Singapore, **2016**. https://doi.org/10.1007/978-981-10-2519-8

17.   Yeh, S. C.; Hwang, W. Y.; Wang, J. L. *et al.* Study of co-located and distant collaboration with symbolic support via a haptics-enhanced virtual reality task. *Interact. Learn. Environ*. **2013**, *21*(2), 184-198. https://doi.org/10.1080/10494820.2012.705854

18.   Christopoulos, A.; Conrad, M.; Shukla, M. Objects, Worlds, and Students: Virtual Interaction in Education. Educ Res Int, **2014**, *2014*, 1-20. https://doi.org/10.1155/2014/318317

19.   Christopoulos, A.; Conrad, M.; Shukla, M. Interaction with educational games in hybrid virtual worlds. *J. Educ. Technol. Syst.* **2018**, *46*(4), 385-413. https://doi.org/10.1177%2F0047239518757986

20.   Moore, K. *Classroom teaching skills*. McGraw-Hill: MA, Boston, 2007.

21.   Concannon, B. J.; Esmail, S.; Roduta Roberts, M. Head-Mounted Display Virtual Reality in Post-Secondary Education and Skill Training: A Systematic Review. *Proceedings of the Frontiers in Education International Conference*, **2019**, 4:80. https://doi.org/10.3389/feduc.2019.00080

22.   Segura, R. J.; del Pino, F. J.; Ogáyar, C. J. *et al.* VR-OCKS: A virtual reality game for learning the basic concepts of programming. *Comput. Appl. Eng. Educ.* **2020**, *28*(1), 31-41. https://doi.org/10.1002/cae.22172

23.   Johari, A.; Chen, C.J.; Toh, S.C. A feasible constructivist instructional development model for virtual reality (VR)-based learning environments: Its efficacy in the novice car driver instruction of Malaysia. *Educ. Technol. Res. Dev.* **2005**, *53*, 111–123. https://doi.org/10.1007/BF02504861

24.   Pellas, N.; Kazanidis, I. Online and hybrid university-level courses with the utilization of Second Life: Investigating the factors that predict student choice in Second Life supported online and hybrid university-level courses. *Comput. Hum. Behav.* **2014**, *40*, 31-43. https://doi.org/10.1016/j.chb.2014.07.047

25.   Chen, Y.L.; Hsu, C.C. Self-regulated mobile game-based English learning in a virtual reality environment. *Comput. Educ.* **2020**, *154*, 103910. https://doi.org/10.1016/j.compedu.2020.103910

26.   Pellas, N. The influence of computer self-efficacy, metacognitive self-regulation and self-esteem on student engagement in online learning programs: Evidence from the virtual world of Second Life. *Comput. Hum. Behav.* **2014**, *35*, 157-170. https://doi.org/10.1016/j.chb.2014.02.048

27.   Junco, R.; Clem, C. Predicting course outcomes with digital textbook usage data. *Internet High Educ* **2017**, *27*, 54–63. https://doi.org/10.1016/j.iheduc.2015.06.001

28.   Siemens, G. Learning analytics: The emergence of a discipline. *Am Behav Sci* **2013**, *57*, 1380–1400. https://doi.org/10.1177%2F0002764213498851

29.   Xing, W.; Guo, R.; Petakovic, E. *et al.* Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory. *Comput. Hum. Behav.* **2015**, 47, 168–181.

30. Drachsler, H.; Kalz, M. The MOOC and learning analytics innovative cycle (MOLAC): A reflective summary of ongoing research and its challenges. *J Comput Assist Learn* **2016**, *32*, 281–290. https://doi.org/10.1111/jcal.12135

31. Slade, S.; Prinsloo, P. Learning analytics: Ethical issues and dilemmas. *Am Behav Sci* **2013**, *57*(10) 1510–1529. https://doi.org/10.1177%2F0002764213479366

32. Long, P.; Siemens, G. Penetrating the fog: Analytics in learning and education. EDUCAUSE Review **2011**, *46*, 31–40. Available online: https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education (Accessed on 6 October 2020).

33. Peña-Ayala, A. Learning analytics: A glance of evolution, status, and trends according to a proposed taxonomy. *WIREs: Data Mining and Knowledge Discovery* **2018**, *8*(3), https://doi.org/10.1002/widm.1243

34. Sciarrone, F. Machine Learning and Learning Analytics: Integrating Data with Learning. *Proceedings of the 17th International Conference on Information Technology Based Higher Education and Training*, **2018**, Olhao, pp. 1-5, doi: 10.1109/ITHET.2018.8424780.

35. Baker, R. S.; Yacef, K. The State of Educational Data Mining in 2009: A Review and Future Visions. *J Educ Data Min* **2009**, *1*(1), 3-17. https://doi.org/10.5281/zenodo.3554657

36. Hevner, A.; March, S.; Park, J. *et al*. Design Science in Information Systems Research. *MIS Quarterly* **2004**, *28*(1), 75-105. doi:10.2307/25148625

37. Howson, C.; Sallam, R. L.; Richardson, J. L. *et al.* Magic quadrant for analytics and business intelligence platforms. Available online: http://www.sift-ag.com/wp-content/uploads/Qlik-Gartner-2019.pdf (Accessed 6 Oct 2020).

38. Chung, J. Y.; Lee, S. Dropout early warning systems for high school students using machine learning. *Child Youth Serv Rev* **2019**, *96*, 346–353. https://doi.org/10.1016/j.childyouth.2018.11.030

39. Gray, C.C.; Perkins, D. Utilizing early engagement and machine learning to predict student outcomes. *Comput Educ* **2019**, *131*, 22–32. https://doi.org/10.1016/j.compedu.2018.12.006

40. Ruipérez-Valiente, J.A.; Muñoz-Merino, P.J.; Alexandron G. *et al.* Using Machine Learning to Detect 'Multiple-Account' Cheating and Analyze the Influence of Student and Problem Features. IEEE Trans. Learn. Technol. **2019,** *12*(1), 112-122. doi: 10.1109/TLT.2017.2784420.

41. Asif, R.; Merceron, A.; Ali, S. A. *et al.* Analyzing undergraduate students' performance using educational data mining. *Comput. Educ.* **2017**, *113*, 177-194. https://doi.org/10.1016/j.compedu.2017.05.007

42. Pliakos, K.; Joo, S.H.; Park, J. Y. *et al.* Integrating machine learning into item response theory for addressing the cold start problem in adaptive learning systems. *Comput. Educ.* **2019**, *137*, 91-103. https://doi.org/10.1016/j.compedu.2019.04.009

43. Alexandron, G.; Yoo, L.Y.; Ruipérez-Valiente, J.A. *et al.* Are MOOC Learning Analytics Results Trustworthy? With Fake Learners, They Might Not Be!. *Int J Artif Intell Educ* **2019**, *29*, 484–506 https://doi.org/10.1007/s40593-019-00183-1

44. Zhang, H.; Huang, T.; Lv, Z. *et al.* MOOCRC: A Highly Accurate Resource Recommendation Model for Use in MOOC Environments. *Mobile Netw Appl* **2019**, *24*, 34–46. https://doi.org/10.1007/s11036-018-1131-y

45. Yang, F.; Li, F. W. Study on student performance estimation, student progress analysis, and student potential prediction based on data mining. *Comput. Educ.* **2018**, *123*, 97-108. https://doi.org/10.1016/j.compedu.2018.04.006

46. Yin M.S.; Haddawy P.; Suebnukarn S. *et al.* Automated outcome scoring in a virtual reality simulator for endodontic surgery. *Comput Methods Programs Biomed.* **2018,** *153*, 53-59. https://doi.org/10.1016/j.cmpb.2017.10.001