Computers & Education: X Reality 4 (2024) 100048

Contents lists available at ScienceDirect

Computers & Education: X Reality

journal homepage: www.journals.elsevier.com/computers-and-education-x-reality

Development and testing of a model for explaining learning and learning-related factors in immersive virtual reality

Emmanuel Fokides [*], Panagiotis Antonopoulos

University of the Aegean, Department of Primary Education, Greece

ARTICLE INFO

Keywords:
Immersive virtual reality
Learning
Model

ABSTRACT

Many believe that immersive virtual reality (IVR) possesses transformative potential for a plethora of human activities that are mediated via technology, including education. In light of this, it is critically important to understand the determinants that influence learning in IVR environments and the interrelations among these determinants. For that matter, a model was developed that encapsulated ten potential factors influencing learning outcomes. Three hundred and thirty-four university students interacted with a purpose-built application that presented ancient Greek inventions through the use of head-mounted displays. Data analysis, adhering to a structural equation modeling approach, indicated that a multitude of factors exerted a positive influence on learning outcomes. These included the perceived quality of graphics, the perceived quality of feedback and content, and the perceived degree of interaction. Moreover, intrinsic motivation and the immersive experience that IVR provides also demonstrated a positive impact. Conversely, the perceived cognitive load and symptoms of simulator sickness manifested a negative impact. Interestingly, these factors did not appear to inhibit learners' motivation or their positive feelings. Age did not have any effect, while gender seemed to have an impact only on immersion. The implications of the results are also discussed.

1. Introduction

Virtual reality (VR) refers to a technology that simulates 3D environments, enabling users to explore and interact with their virtual surroundings in a way that approximates reality. The main objective is to create the feeling of "being" in a digital world even though users are physically present in the real one. VR can be experienced through a variety of displays (e.g., monitors, tablets, and smartphones; yet, the ongoing advancements in the field have allowed an alternative way for one to have VR experiences, namely with the use of head-mounted displays (HMDs). In recent years, these devices have become widely available and economically affordable. They not only enable stereoscopic vision (allowing the virtual world to be perceived in 3D), but they also support motion tracking, spatial audio, and, using controllers, interaction with the virtual world and haptic feedback. It is widely agreed that HMDs provide elevated levels of immersion (Makransky & Petersen, 2021). To distinguish the experiences offered by HMDs (as well as CAVE systems), the term "Immersive Virtual Reality" (IVR) was coined.

IVR's potential to transform a wide range of human activities (e.g., entertainment, gaming, real estate, business, commerce, work, social life, and healthcare) is widely recognized. Alongside other technologies (e.g., mixed and augmented reality), it can also revolutionize education. It can reshape the traditional teacher-student relationship and overcome spatial and temporal limitations (Lin, Wan, Gan, Chen, & Chao, 2022). IVR enhances students' autonomy and can increase their interest and engagement in the learning process (Kye, Han, Kim, Park, & Jo, 2021). Logically, there has been a shift in VR-related research toward exploring the affordances of IVR and its impact on learning. In the vast majority of cases, its impact has been reported as being positive. For instance, in the meta-analysis of 21 studies (Villena-Taranilla, Tirado-Olivares, Cozar-Gutierrez, & González-Calero, 2022), the authors found that the effect size was larger in IVR compared to semi- and non-immersive systems. Moreover, the effect size did not depend on the educational level and knowledge domain, while short interventions were considered more effective. However, there are meta-analyses reporting that, although the results were in favor of IVR, the effect size was small. For example, Coban, Bolat, and Goksu (2022) in their meta-analysis of 48 studies, concluded that not only the effect size was small, but it also differentiated depending on the level/field of education and the educational

* Corresponding author. 1 Dimokratias str, 85132, Rhodes, Greece.
E-mail addresses: fokides@aegean.gr (E. Fokides), pred19003@rhodes.aegean.gr (P. Antonopoulos).

https://doi.org/10.1016/j.cexr.2023.100048
Received 3 September 2023; Received in revised form 29 October 2023; Accepted 25 November 2023
Available online 9 December 2023

sources used. Similarly, Wu, Yu, and Gu (2020) in their meta-analysis of 35 studies, found that although IVR was more effective compared to non-immersive teaching/learning approaches, the effect size was small.

Setting aside knowledge gains, a positive impact was noted on skills acquisition (e.g., Queiroz, Nascimento, Tori, & da Silva Leme, 2018). The participants in several studies reported elevated levels of enjoyment and positive feelings (e.g., Butt, Kardong-Edgren, & Ellertson, 2018; Caro, Carter, Dagli, Schissler, & Millunchick, 2018), engagement and interest (e.g., Abichandani, Mcintyre, Fligor, & Lobo, 2019; Bertrand, Bhargava, Madathil, Gramopadhye, & Babu, 2017), motivation (e.g., Rupp et al., 2019), self-confidence (Tzanavari, Charalambous-Darden, Herakleous, & Poullis, 2015, pp. 423–427), creativity and active learning (Caro et al., 2018; Erturk & Reynolds, 2020). All the above probably contributed to the positive learning outcomes. Indeed, Huifen et al. (2021, pp. 320–325) in their meta-analysis of 45 studies, concluded that IVR motivated students, elicited their interest, and engaged them with the learning material, probably because it can offer first-hand experiences and opportunities for situated learning.

Still, as indicated in a literature review (Atsikpasi & Fokides, 2022) further research is needed to comprehend the educational potential of IVR. Firstly, as we stated above, its impact on learning, although positive, appears to be modest, suggesting room for improvement. Secondly, there is a multitude of factors involved in shaping learning outcomes, as we will elaborate on in the following sections. The main challenge lies in the fact that researchers have considered different factors, a limited number of them, or have not examined learning *per se* but rather focused on users' intentions or perceptions. Thus, it remains unclear which factors play a decisive role and, more importantly, how they interact. Given the complexity of the issue, what we felt was missing was a model that incorporates a substantial number of factors influencing learning outcomes, illustrating their interactions and their effect on learning.

To address these issues, in mid-2021, we initiated a research project to develop a series of highly immersive VR applications and assess their impact on learning. The initial outcome of this project was the development (and subsequent testing) of the MLES scale (Fokides, 2023). This scale was the result of a review of studies related to the educational use of IVR, with a focus on the scales and questionnaires used by researchers for data collection. In this study, we present the results of the project's second phase, in which we developed a model for examining the interactions and effects of factors we theorized to influence one's learning in IVR. We present the steps we followed for developing the model, gathering data, and analyzing them, together with the subsequent discussion of the results in the coming sections.

2. Related work

While research on the impact of IVR on learning and skills is quite extensive, there are far fewer studies that have attempted to develop models to explain how learning occurs. Therefore, we expanded our review of related work to include studies conducted within the context of the Metaverse, which, besides IVR, encompasses various technologies.

In a study aimed at predicting users' intention to use the Metaverse (without explicitly specifying which technology was used) within the context of medical education (Almarzouqi, Aburayya, & Salloum, 2022), the researchers considered factors from the Technology Acceptance Model (TAM, Davis, Bagozzi, & Warshaw, 1989). They also incorporated personal innovativeness, perceived compatibility, user satisfaction, perceived triability, and perceived observability. Upon analyzing data collected from 1858 university students, the study revealed that perceived usefulness, perceived ease of use, and user satisfaction were the key determinants.

Again, within the realm of medical training, and employing the TAM as the foundation for their model, researchers examined students' perceptions concerning the utilization of the Metaverse in their training (Alawadhi et al., 2022). The data collected from 435 participants indicated that personal innovation and perceived enjoyment significantly impacted perceived ease of use and perceived usefulness. Furthermore, these latter two factors had a significant influence on students' intention to use Metaverse applications. Fussell and Truong (2022) tested an extended TAM, to determine the factors influencing flight students' intentions to use VR for their training. Having a sample of 489 such students, they concluded that the original TAM factors together with factors relevant to VR technology had the strongest relationships and impact on attitude and behavioral intention to use VR.

In the context of history teaching, Villena-Taranilla, Cózar-Gutiérrez, González-Calero, and Diago (2023) explored the causal relationships between perceived attention, perceived ease of use, prior knowledge, perceived utility, attitude towards the use of IVR, and perceived enjoyment, having as a sample 111 primary school students. They found that prior knowledge did not explain students' intention to use IVR, while perceived attention had an impact (mediated through perceived utility).

In a systematic literature review that encompassed 41 studies related to the educational applications of the Metaverse published from 2011 to 2022, the authors concluded that the Technology Acceptance Model (TAM) was the most frequently employed (21 cases), followed by the Unified Theory of Acceptance and Use of Technology (UTAUT) (Alfaisal, Hashim, & Azizan, 2022). The authors also pointed out that VR was the most frequently used technology (32 cases).

Studies that have developed models to examine learning outcomes do exist. For example, Fokides and Atsikpasi (2018) tested a model to explain learning in informal learning settings. They employed a desktop VR application, and their sample consisted of 612 individuals spanning a wide age range. Their findings indicated that motivation, perceived ease of use, perceived usefulness, and enjoyment served as predictors of learning outcomes. Similar results were observed in another study with a target group of 437 primary school students (Fokides, 2017).

Other studies developed more intricate models. Of interest is the Cognitive Affective Model of Immersive Learning (Makransky & Petersen, 2021). Despite being a theoretical model that lacks empirical testing, the authors present persuasive arguments that representational fidelity, control factors, and immersion contribute to the development of presence and agency, which they considered as the psychological affordances of learning in IVR. Furthermore, they posited that self-efficacy, interest, motivation, cognitive load, embodiment, and self-regulation can facilitate knowledge acquisition in IVR environments.

In another study, the authors did not include actual learning as a variable; instead, they examined students' perceived learning (Makransky & Lilleholt, 2018). They recruited 104 university students who interacted with both the desktop and immersive versions of an application (using GearVR, which allowed only rotational tracking). The authors concluded that immersion led to higher levels of presence, increased enjoyment, and heightened motivation, resulting in higher perceived learning outcomes. Immersion also influenced perceived cognitive benefits, once again leading to higher perceived learning outcomes. In another study, Huang, Roscoe, Craig, and Johnson-Glenberg (2022) compared both Oculus Go and Oculus Rift headsets, although their sample size was rather small for model testing (77 undergraduate students). The results indicated that prior knowledge was the most significant determinant of learning outcomes. Immersion positively impacted the user experience and motivation, while its effect on cognitive engagement remained unclear.

Within the context of desktop VR, Makransky and Petersen (2019) recruited 199 undergraduate students attending a genetics course. They concluded that two pathways influenced learning outcomes: one stemming from VR features such as control and representational fidelity, leading to presence, motivation, and self-efficacy, and another pathway leading from VR features to usability, cognitive benefits, and self-efficacy. Again, in the context of desktop VR, Lee, Wong, and Fung (2010) worked with a sample of 210 senior high school students who utilized a desktop VR dissection simulator. According to the authors, motivation and presence, in conjunction with reflective thinking, cognitive benefits, control, and active learning, affected learning.

It appears that a substantial body of research has focused on the development and testing of models aimed at examining acceptance or intention to use VR/IVR as an educational platform. Given that our primary goal was to assess the impact of factors on learning, we determined that these models were not well-aligned with our objectives, although they did provide valuable insights. Moreover, the technology predominantly employed in the aforementioned studies was desktop or non-fully immersive VR. Additionally, in some studies, the sample size was small or borderline; therefore, we have some reservations regarding the validity of their results. What we also found troubling was that in some cases, the participants did not actually use VR/IVR systems. Instead, they viewed video presentations of such systems (e.g., Fussell & Truong, 2022). In our perspective, this approach does not enable one to fully experience or even understand the pros and cons of these technologies.

3. Model formation

The first issue we had to resolve when trying to develop our model was which factors to include. Although, as we stated in the preceding section, previously tested models provided us with useful ideas about which factors are important, we decided to follow an approach similar to the one when developing the MLES scale. The starting point of our endeavor was (i) the literature review of the educational uses of HMDs (Atsikpasi & Fokides, 2022) because it was closely related to how we view IVR and (ii) the literature review of Alfaisal et al. (2022), because, although it reviewed all types of Metaverse technologies, it mostly focused on VR. We accessed and examined the 128 studies included in both reviews in terms of the research questions/hypotheses they set, as well as any other factors they examined besides learning that their authors considered important in shaping learning outcomes. Indeed, we found that a multitude of other factors were of interest. Readers can find the complete list of the most commonly used in Appendix I.

The next step was to come up with larger constructs by grouping factors if possible. We concluded that past research focused on: (i) the technical aspects of the applications and devices, such as ease of use, perceived usability and control, the perceived quality of the graphics and perceived realism of the applications, and perceived degree of interaction, (ii) the perceived quality of the feedback and content, (iii) immersion and presence, (iv) the effects of simulator sickness, (v) enjoyment and positive/negative feelings in general, (vi) engagement and motivation, (vii) subjective usefulness of the applications and subjective impact on learning, (vii) perceived cognitive load, (ix) collaboration and social interactions, and (x) self-efficacy. Several other factors were examined just once or could not fit into the above categories. On the basis of the above, we decided to include in our model the factors listed below.

Perceived feedback and content quality: Interactivity, usability, and the sense of realism depend on the proper design of multimodal feedback cues (auditory, visual, and haptic) (Faeth & Harding, 2014). Indeed, such cues can have a positive impact on user experience and task accuracy (Faeth & Harding, 2014; Kobayashi et al., 2016). Equally important is the presentation of information and learning content (e.g., help screens, text, images, and audio), which helps users feel guided and avoid confusion regarding what to do, where to find information, and what to learn. Furthermore, studies have shown that the quality of teaching content presented through AR/VR applications significantly affected their effectiveness (Portman, Natapov, & Fisher-Gewirtzman, 2015; Potkonjak et al., 2016).

Perceived quality of the virtual environment's graphics: Clearly, users expect realistic representations of virtual environments. Representational fidelity pertains not only to the realism of 3D objects but also to how realistically they behave (Lee et al., 2010). It can reduce or even eliminate users' disbelief, which is the feeling of being in a non-real environment (Mystakidis, 2022), and enhance their sense of immersion. Research has supported a positive correlation between the realism of the virtual environment and various factors, including immersion (Mystakidis, 2022; Parong & Mayer, 2018), motivation (Lee et al., 2010), and learning outcomes (e.g., Harrington, 2012; Kim & Ahn, 2021).

Perceived ease of use/control of the virtual environment: Perceived ease of use refers to users' belief that a given system or technology will not require too much effort to complete tasks with it. In the context of TAM, it impacts users' attitudes (Davis et al., 1989). However, in the context of VR, it was found that it was not significantly associated (Abd Majid & Mohd Shamsudin, 2019). Lee et al. (2010) suggested that for VR to enhance one's learning experience and have an impact on their learning, it must be considered user-friendly. Similarly, Asad, Naz, Churi, Guerrero, and Salameh (2022) concluded that the user-friendliness of VR ensures its implementation and, in turn, enhances experiential learning.

Perceived degree of interaction with the virtual environment: Users should be allowed to interact with the environment ubiquitously. Moreover, the devices used for interaction have to emulate how users interact with objects in reality, moving beyond the mouse and keyboard that do not accurately reflect body movements (Duan et al., 2021). In the context of IVR applications, controllers are commonly used, allowing users to interact with virtual objects (e.g., touch, manipulate, and operate). Interactions are further enhanced through the use of motion-tracking technology and devices that allow users to freely move in all directions in addition to rotational motion (Atsikpasi & Fokides, 2022) and also interact with virtual objects (Maereg, Nagar, Reid, & Secco, 2017). Given that, users become active learners (Mystakidis, 2022). Studies indicated that the adequacy and realism of interactions can positively impact the effectiveness of VR/AR applications (Portman et al., 2015; Potkonjak et al., 2016).

Cognitive load: The cognitive load theory is based on the notion that working memory has a finite capacity (Sweller, 2020). Thus, the educational material facilitates learning when it puts as little burden as possible on working memory; otherwise, learning becomes difficult. The educational material should aim at reducing the exogenous cognitive load (i.e., information not related to the nature and basic structure of the information, but to the way it is presented) and increasing the germane cognitive load, as this frees up working memory resources (Leppink, Paas, Van Gog, van Der Vleuten, & Van Merrienboer, 2014; Novak, Daday, & McDaniel, 2018). In the research of Cecotti, Day-Scott, Huisinga, and Gordo-Pelaez (2020), the cognitive load of an immersive VR application was assessed as low. Armougum, Orriols, Gaston-Bellegarde, Joie-La Marle, and Piolino (2019) concluded that there is no difference between the cognitive load a virtual environment causes and the corresponding real one. Yet, others suggested that IVR applications can lead to increased cognitive load (Makransky, Terkildsen, & Mayer, 2017). Ibili and Billinghurst (2019) found that the perceived usefulness and perceived ease of use had a strong positive relationship with germane cognitive load.

Simulator sickness: A commonly reported negative side effect of using HMDs is simulator sickness (motion sickness and cybersickness are also terms used). When users move in a virtual environment with the help of controllers, their brains are faced with a contradiction. On the one hand, through vision, they perceive that their bodies are moving, but on the other hand, their brains have not given any movement commands (Gallagher & Ferrè, 2018). Moreover, the signals received from their vestibular systems signify that they are stationary. The symptoms of simulator sickness include pallor, nausea, disorientation, and sometimes vomiting. Studies have indicated that the symptoms decrease with visual fidelity (de Winkel, Talsma, & Happee, 2022). Importantly, simulator sickness negatively impacts one's motivation to perform (Johnson, 2005), learning, presence, and engagement (Atsikpasi & Fokides, 2022; Maraj, Badillo-Urquiola, Martinez, Stevens, & Maxwell, 2017). However, some have suggested that the performance of psychomotor and cognitive tasks was not affected (Johnson, 2005). Conversely, others have reported that even mild simulator sickness negatively impacted mental workload and learning outcomes (Hsin et al., 2022).

Immersion/presence: Immersion refers to the technical affordances of a system that delivers an environment that gives users a sense of reality (Rasimah, Nurazean, Salwani, Norziha, & Roslina, 2015). Presence is the perceptual illusion that there is no medium between the user and the virtual environment and relates to users' perceptions of reality (Baños et al., 2004). In our study, we decided to treat immersion and presence as a single factor. We believe that, although not identical, these concepts are very closely related; immersion reflects the system's technical qualities that allow one to feel present in a virtual environment (Witmer, Jerome, & Singer, 2005). Therefore, we expected both to be affected by the same factors and have similar (if not identical) effects on other factors. Immersion and presence can improve learning outcomes (Kim & Ahn, 2021). Due to immersion and presence, learners become mentally and emotionally engaged with the virtual environment (Lindgren, Tscholl, Wang, & Johnson, 2016), and these feelings help them acquire self-directed learning experiences (Jeon & Jung, 2021, pp. 361–368). Furthermore, immersion and presence, along with the on-demand repetition of an application, can help users better understand the learning content (Maas & Hughes, 2020).

Positive feelings/enjoyment: Studies have demonstrated that learners find VR applications enjoyable (e.g., Barry et al., 2015; Fokides & Zampouli, 2017; Moriuchi, Landers, Colton, & Hair, 2021; Pagano et al., 2020). Some theories suggest that the immersive and interactive experiences students have in VR lead to more enjoyable learning, which, in turn, allowed them to learn more effectively (e.g., Lehikko, 2021; Parong & Mayer, 2021). In our view, it is not just enjoyment that can lead to better learning but positive feelings in general. For example, Kim and Ahn (2021) found that, among other factors, satisfaction had a positive impact on learning in a VR application. Hedonic motivation (i.e., the feeling of happiness when using a technology tool, Brown & Venkatesh, 2005) may also play a role, as users experience feelings of self-fulfillment and fun (Ramírez-Correa, Rondán-Cataluña, Arenas-Gaitán, & Martín-Velicia, 2019).

Motivation: In general, motivation is considered a key factor with a considerable impact on learning. Metaverse applications have been found to be useful tools for increasing learners' motivation (Go, Jeong, Kim, & Sin, 2021; Jeon & Jung, 2021, pp. 361–368). This is likely because they provide flexible, engaging, and dynamic learning environments, allowing users to have unique experiences that foster their motivation (Erturk & Reynolds, 2020). Moreover, as these applications allow for innovative learning approaches, they may impact motivation through self-directed learning (Jeon & Jung, 2021, pp. 361–368). The enjoyable nature of these applications may also be a motivating factor (Barry et al., 2015). Others have suggested that Metaverse applications enhance motivation because they offer more immersive and interactive experiences that promote active learning (Diaz, Saldaña, & Avila, 2020).

Perceived knowledge gains: According to TAM, perceived usefulness refers to users' belief that a given technology will help them improve their performance (Davis et al., 1989). The perceived usefulness of VR has been found to significantly influence students' attitudes and intentions to use this technology (Abd Majid & Mohd Shamsudin, 2019). Similarly, Shen, Xu, Sotiriadis, and Wang (2022) concluded that perceived usefulness (along with other factors) was a determinant of students' intention to use AR/VR applications and learn through them. In the context of our research, we view perceived usefulness as users' subjective perception that the virtual environment helped them increase their academic performance and that they learned something while viewing/interacting with it. Consequently, we hypothesized that perceived knowledge gains might have an impact on the actual learning outcomes.

We also considered the inclusion of self-efficacy and collaboration/social interactions, but we decided not to include them at this stage. During the development of the MLES scale, we found that only a handful of the participants had an understanding of what IVR is, and even fewer had immersive VR experiences or experience in using HMDs. Therefore, we believed that examining self-efficacy as a contributing factor would be pointless, given that we expected it to show no variance. Of course, as IVR gains momentum and attracts the interest of users, self-efficacy should be considered. Regarding collaboration and social interactions, we did not include them because the application we developed for the study was not a multi-user one, and therefore, no collaboration and social experiences were possible. Once again, as this factor is an important one, we strongly recommend its inclusion in multi-user scenarios.

4. Method

4.1. Research questions

Given that the simultaneous examination of a considerable number of factors might change their interactions and effects as these were reported in past research, we examined all possible relationships. As a result, we formed a total of 49 research questions (Fig. 1).

- RQ1a-g. Do the perceived feedback and content quality (Feedback/Content) affect: the perceived cognitive load of the learning material (Cognitive Load) (RQ1a), users' simulator sickness (Simulator Sickness) (RQ1b), their feelings of immersion and presence to the virtual environment (Immersion/Presence) (RQ1c), their positive feelings (Positive Feelings) (RQ1d), their motivation to learn (Motivation) (RQ1e), their perceived knowledge gains (Gains) (RQ1f), and the learning outcomes (RQ1g)?
- RQ2a-g. Does the perceived quality of the virtual environment's graphics (Graphics) affect: the Cognitive Load (RQ2a), Simulator Sickness (RQ2b), Immersion/Presence (RQ2c), Positive Feelings (RQ2d), Motivation (RQ2e), Gains (RQ2f), and the learning outcomes (RQ2g)?
- RQ3a-g. Does the perceived ease of use/control of the virtual environment (Control) affect: the Cognitive Load (RQ3a), Simulator Sickness (RQ3b), Immersion/Presence (RQ3c), Positive Feelings (RQ3d), Motivation (RQ3e), Gains (RQ3f), and the learning outcomes (RQ3g)?
- RQ4a-g. Does the perceived degree of interaction with the virtual environment (Interaction) affect: the Cognitive Load (RQ4a), Simulator Sickness (RQ4b), Immersion/Presence (RQ4c), Positive Feelings (RQ4d), Motivation (RQ4e), Gains (RQ4f), and the actual learning outcomes (RQ4g)?
- RQ5a-f. Does the perceived cognitive load of the learning material (Cognitive Load) affect: Simulator Sickness (RQ5a), Immersion/Presence (RQ5b), Positive Feelings (RQ5c), Motivation (RQ5d), Gains (RQ5e), and the learning outcomes (RQ5f)?
- RQ6a-e. Does users' simulator sickness (Simulator Sickness) affect: Immersion/Presence (RQ6a), Positive Feelings (RQ6b), Motivation (RQ6c), Gains (RQ6d), and the learning outcomes (RQ6e)?
- RQ7a-d. Does users' immersion/presence affect: Positive Feelings (RQ7a), Motivation (RQ7b), Gains (RQ7c), and the learning outcomes (RQ7d)?
- RQ8a-c. Do users' positive feelings (Positive Feelings) affect: Motivation (RQ8a), Gains (RQ8b), and learning outcomes (RQ8c)?
- RQ9a-b. Does users' motivation to use the virtual environment and to learn (Motivation) affect: Gains (RQ9a) and the learning outcomes (RQ9b)?
- RQ10. Do the perceived knowledge gains (Gains) affect the learning outcomes?

Control variables should be included when developing a model because they usually affect the results. One can identify several such variables, but we considered essential the users' gender and age group. Consequently, we examined the following 14 research questions.

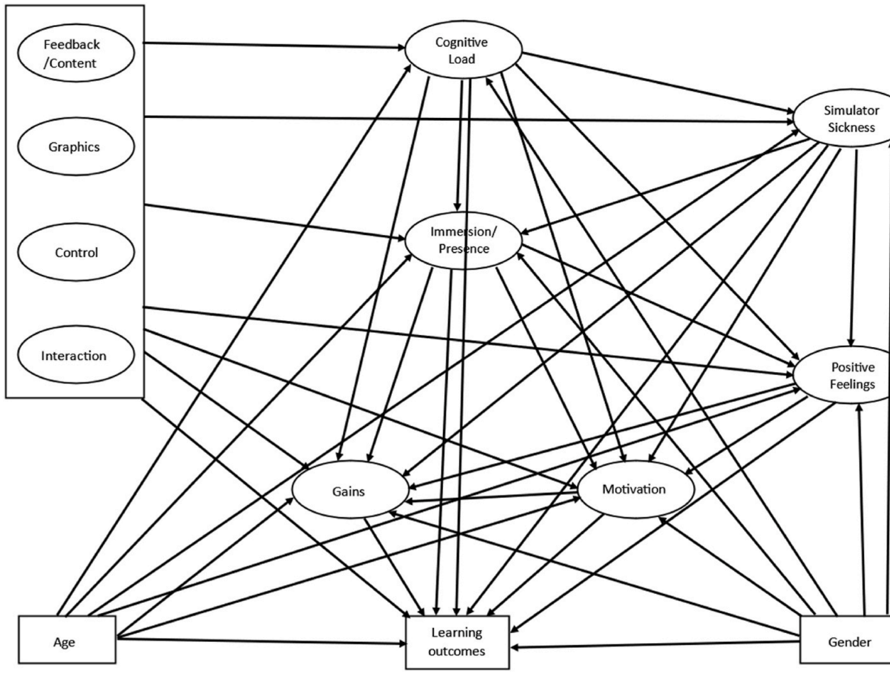

Fig. 1. The research model
Note. For the clarity of the presentation (i.e., the reduction of the number of arrows), we grouped the exogenous variables.

- RQ11a-g. Does users' age group affect: Cognitive Load (RQ11a), Simulator Sickness (RQ11b), Immersion/Presence (RQ11c), Positive Feelings (RQ11d), Motivation (RQ11e), Gains (RQ11f), and the learning outcomes (RQ11g)?
- RQ12a-g. Does users' gender affect: Cognitive Load (RQ12a), Simulator Sickness (RQ12b), Immersion/Presence (RQ12c), Positive Feelings (RQ12d), Motivation (RQ12e), Gains (RQ12f), and the learning outcomes (RQ12g)?

For model development purposes, we treated Feedback, Control, Graphics, and Interaction as exogenous variables, while we treated the remaining factors as endogenous ones.

4.2. Participants

Most users of IVR educational applications are expected to be either educators or students around 20 years old, given that most of the relevant applications are addressed to adolescents or young adults. Furthermore, most studies focused on university/college students, as indicated by recent reviews of the relevant literature (Alfaisal et al., 2022; Atsikpasi & Fokides, 2022; Tlili et al., 2022). On the basis of this assumption, we found it quite logical to target students studying at a department of primary education, as they belong to the above age group and, at the same time, are future educators. Due to logistical constraints that precluded the conduction of the experiment across varying geographical locations, we opted to employ a convenience sampling approach, thereby recruiting students exclusively from a single department of primary education. We posted a call for participation (presenting the study's objectives and procedures) on social media, addressed to students studying at the Department of Primary Education, University of the Aegean. There were no prerequisites for participation (e.g., experience in using HMDs). Three hundred and forty enrolled, in exchange for course credit. Although the study is part of an ongoing research project already reviewed and approved by the Department's Ethical Committee, the enrolled students were asked to give their informed consent as well.

The other issue we had to address was the sample size. One has to consider quite a lot of parameters, rendering it a rather hard-to-solve puzzle. Besides the number of items the instrument used for data collection has, the research settings, estimation method, data scaling, indicator reliability, and model complexity also have to be considered (Brown, 2015). However, we assumed that the 340 recruited students constituted a satisfactory sample size, as it was well above the 200 cases suggested by Kline (2016), within the suggestion of five to ten responders per item (Bentler & Chou, 1987), and above the suggestion of five observations per estimated parameter (Hair, Black, Babin, & Anderson, 2019).

4.2.1. Materials and apparatus

The educational application to use in the study was also a matter of consideration. It is true that there are several available; yet, few are in Greek, and even fewer are not too complex or too specialized (in terms of their learning subject) or do not need too much time to complete. It is also true that the application we developed for testing the MLES scale could have been used in this study as well, but it was rather important to test the scale in a different context. Taking together the above, we developed a new one, the theme of which was ancient Greek inventions, assuming that the topic might be of interest to future educators. We selected eight inventions, namely Heron's hovering sphere, Heron's aeolosphere, Heron's automatic holy water server with coin-collector, Heron's automatic opening of the temple gates after a sacrifice had taken place on its altar, Archimedes's steam cannon, Aeneas's hydraulic telegraph, Kleoxenos's and Dimokleitos's phryctoria (a telecommunication system using fire signals), and the flamethrower of the Boeotians (Fig. 2).

Since no relics of the inventions exist and no 3D models of these inventions were available, we developed them from scratch using Blender, which is a freely available 3D model developing software. Drawing from relevant ancient texts, published works, and museum exhibits, we made a concerted effort to accurately reconstruct them. The IVR application was developed using Unity. The models were fully functional, allowing users to interact with and observe their mechanisms in action. Their functionality, behaviors, as well as the interactions with them were implemented using Unity's integrated C# scripting/programming language, Unity's particle system, and Filo an addon for Unity developed by Virtual Method Studio. For instance, in the case of Heron's

Fig. 2. Screenshots from the IVR application. a: the flamethrower of the Boeotians, b: Heron's aeolosphere, c: Heron's automatic opening of the temple gates, d: Heron's automatic holy water server with coin-collector (transparent mode), e: Archimedes's steam cannon, and f: the components of an invention on a bench.

automatic opening of the temple gates, users could light the fire on the altar, witness the gates opening, and simultaneously, the mechanism became visible, enabling them to understand its functioning. We utilized Unity to develop the application. Sound effects (e.g., when the temple's doors were opening, when the steam cannon fired, and when the fire was lit) were also added, using freely available sounds.

In the landing/welcome area, users could select which invention they wanted to view. Within eight different spaces, each depicting an outdoor environment, we positioned the inventions. We incorporated information buttons to provide texts, images, and audio narrations explaining how a particular invention operated, along with details about its inventor and the historical context. Adjacent to each invention, we placed its components on a nearby bench, allowing users to examine them in detail.

Among the available HMDs, we chose the Meta Quest 2 due to its affordable price and commendable technical specifications, making it an attractive choice for the average consumer seeking an immersive VR experience. However, it is untethered (not connected to a computer). Although this results in fewer cables and greater freedom of movement, it comes at the cost of less impressive graphics. This limitation arises because the HMD handles image processing, and its processing power is constrained by the need to accommodate numerous electronics in a confined space. Given none of the participants were in possession of their own HMDs, an issue which will be elaborated further in section "4.4 Procedure," we provided them with several of these devices already in our ownership.

We conducted the experiment within a spacious office setting from which all furniture was removed, thus creating an unobstructed area approximately 40 m^2 in size. This decision served two purposes: (i) to prevent injuries that might have occurred if participants had bumped into furniture or walls, and (ii) to enable participants to walk instead of relying on the HMD's controllers to simulate movement, providing a more immersive experience.

4.3. Instruments

To gather data related to the learning outcomes, we devised eight short quizzes, corresponding to the number of inventions available for viewing. Depending on which invention a participant interacted with, they were given the corresponding quiz. Each quiz comprised of fifteen multiple-choice questions, all designed to assess declarative knowledge. The questions varied in difficulty and were related to the information presented directly (via images, texts, and audio narration) or indirectly (information that could be deduced from interacting with the inventions or their components). For instance, in the former case, participants were asked where and when an inventor lived, the purpose an invention served, the laws of physics and chemistry that governed the function of an invention. In the latter case, participants were asked to identify the materials used and/or the parts of an invention, estimate its dimensions. Participants were also tasked with identifying the correct sequence of steps for operating an invention, to estimate the speed an invention was functioning and/or rotating (e.g., for Heron's hovering sphere and Heron's aeolosphere), to estimate the distance a projectile could travel (e.g., for Archimedes's steam cannon and the flamethrower of the Boeotians), and to estimate how far the signal coming out of an invention could be seen (e.g., for Aeneas's hydraulic telegraph and Kleoxenos's and Dimokleitos's phryctoria).

For obtaining data concerning participants' views, attitudes, and feelings, we utilized the MLES scale, as presented in the "Introduction." This scale consists of a total of 43 items, examining the ten factors included in our research model: simulator sickness (Simulator Sickness, six items), perceived cognitive load (Cognitive Load, three items), motivation (Motivation, six items), perceived ease of use/control of the virtual environment (Control, three items), perceived knowledge gains (Gains, three items), positive feelings (Positive Feelings, four items), perceived quality of the virtual environment's graphics (Graphics, six items), perceived feedback and content quality (Feedback/Content, three items), perceived degree of interaction (Interaction, three items), and immersion/presence (Immersion/Presence, six items). In addition, we included two items for gathering data related to the two control variables in the model: participants' gender and age group.

4.4. Procedure

The students recruited for the study partook in the experiment on an individual basis. After welcoming each participant, we provided them with verbal instructions regarding what to anticipate and how to navigate and interact with the application. Subsequently, we handed them the HMDs, and after adjusting the straps and interpupillary distance to optimize image quality, they were given ten to 15 min to explore the welcoming space and menus available upon booting up the headset. We also briefed them on what to expect when nearing the boundaries of the "play" area and how to react. This initial procedure was crucial, given that none of the participants had any prior experience with HMDs.

The next phase involved participants launching the application and choosing which invention they wished to view from the landing space. They were provided with approximately 20–25 min, which we deemed sufficient for a comprehensive experience and the study of the relevant learning material. Participants encountering substantial navigation or interaction difficulties were given additional verbal instructions. Subsequently, participants completed the quiz, and we administered the MLES scale (which took 15–20 min to complete, and both were available online). The total time required for each participant slightly exceeded 1 h.

Anticipating potential cases of severe simulator sickness, we instructed participants to discontinue using the application, remove their headsets, and take a break. Conversely, if symptoms were mild, participants were given the choice to continue or halt the experiment at their discretion. We administered both the quiz and the scale to all participants, regardless of whether they terminated or completed the experiment, as we hypothesized that simulator sickness significantly influences one's learning experience and outcomes.

5. Results

5.1. Data preparation and assumptions checking

We obtained a score for the quizzes by allocating one to three points for each correct answer (depending on the question's difficulty) and then transforming it to a 100-point scale. We inserted the data from the quizzes and the MLES scale into SPSS 28 for conducting all the analyses presented in the coming sections. The first step was to check for missing data (none was found) and for unengaged responses (i.e., participants having no variance in their responses, $SD = 0.00$). We excluded six participants because of the latter issue. Thus, our final sample size was 334 participants, most of them being females ($N = 245$) and most belonging to the 20-24-year-old group ($N = 282$). We expected the above, as the participants were students studying at a department of primary education.

As multiple types of statistical analyses were to follow [i.e., confirmatory factor analysis (CFA) and structural equation modeling (SEM)], the second step was to check whether the data met the assumptions for these analyses. Normality was checked by inspecting the skewness and kurtosis of all the variables. We found all the items belonging to the factors labeled "simulator sickness" to be highly positively skewed (their skewness exceeded the recommended maximum value of $|2|$ (Finney & DiStefano, 2013). As a result, we decided to transform them using log (10) and recheck them. Despite their transformation, their skewness remained slightly above the recommended maximum. Although this can be considered a limitation, we can counter-argue that it is logical to have normality issues in this factor. That is because the vast majority of individuals did not face simulator sickness problems (see Table 5); therefore, their responses to the items examining this factor were mostly 1 (indicating no problems), resulting in positively skewed data. In light of the above, we did not exclude these items.

We checked linearity by conducting an OLS linear regression between each dependent-independent variable pair. Given that in all cases the p-value was less than 0.05, we concluded that their relationship was sufficiently linear. We checked multicollinearity by running a multivariate regression and inspecting the Variable Inflation Factor (VIF) for each independent variable. As the VIF in all cases was below the threshold of 10.00 (Vittinghoff, Glidden, Shiboski, & McCulloch, 2012), we concluded that there were no multicollinearity concerns. Finally, we checked heteroskedasticity by conducting the White Test for Heteroskedasticity, the Modified Breusch-Pagan Test for Heteroskedasticity, the Breusch-Pagan Test for Heteroskedasticity, and the F Test for Heteroskedasticity. The above tests confirmed that heteroskedasticity was not an issue ($p > .05$ in all cases).

5.2. Confirmatory factor analysis

To verify the factorial structure of the MLES scale, we conducted a CFA using AMOS 28. We chose Maximum Likelihood as the estimation method because it is known to be quite robust to moderate violations of normality (Matsunaga, 2010). We found the specified model that emerged to be unsatisfactory, as it exhibited mixed results concerning convergent and discriminant validity, cross-loadings, and model fit. To improve the model, we removed four items and re-ran the analysis. The resulting model demonstrated highly satisfactory properties. The standardized estimates were very good, ranging from 0.72 to 0.96 (Hair et al., 2019) (Table 1). All the R^2 values were above the 0.50 threshold. The only exception was SimSick5. Then again, as its value was close to the above threshold ($R^2 = 0.40$), we considered it acceptable. We calculated the model's composite reliability (CR) and the average

Table 1
Results of the CFA (retained items).

Item	Est.	SE	t	p	R^2	Item	Est.	SE	t	p	R^2
Motivation 2	.80	–	–	–	.63	Imm/Pre 6	.77	.08	12.98	<.001	.59
Motivation 1	.84	.06	19.16	<.001	.71	Imm/Pre 5	.87	.10	14.01	<.001	.75
Motivation 6	.78	.07	14.90	<.001	.61	Imm/Pre 3	.78	.06	17.28	<.001	.60
Motivation 3	.80	.05	18.95	<.001	.64	Imm/Pre 2	.71	–	–	–	.53
Motivation 4	.76	.07	14.63	<.001	.58	Pos. feel 4	.96	.06	21.44	<.001	.92
Motivation 5	.79	.07	14.77	<.001	.63	Pos. feel 2	.96	.05	24.16	<.001	.92
Sim. sick 2	.76	.07	14.01	<.001	.58	Pos. feel 1	.82	.06	19.51	<.001	.69
Sim. sick 4	.82	.07	12.32	<.001	.68	Pos feel 3	.86	–	–	–	.73
Sim. sick 3	.77	.06	12.26	<.001	.59	Graphics 1	.77	.06	15.45	<.001	.59
Sim. sick 1	.72	–	–	–	.52	Graphics 2	.80	.06	15.88	<.001	.64
Sim. sick 5	.64	.05	10.24	<.001	.40	Graphics 6	.80	.06	16.08	<.001	.64
Sim. sick 6	.72	.05	11.64	<.001	.52	Graphics 4	.73	.06	14.48	<.001	.53
Cog. load 2	.76	–	–	–	.58	Graphics 3	.83	–	–	–	.69
Cog. load 3	.87	.13	8.23	<.001	.76	Interaction 3	.78	.08	14.57	<.001	.60
Control 2	.85	–	–	–	.73	Interaction 1	.82	.07	15.44	<.001	.68
Control 3	.85	.07	17.86	<.001	.72	Interaction 2	.80	–	–	–	.65
Control 1	.81	.06	16.88	<.001	.65	Feed/Cont 2	.83	.07	14.90	<.001	.69
Gains 2	.96	–	–	–	.92	Feed/Cont 3	.80	.07	14.44	<.001	.65
Gains 3	.93	.03	29.08	<.001	.86	Feed/Cont 1	.77	–	–	–	.59
Gains 1	.75	.04	18.18	<.001	.56						

Notes. -: This value was fixed at 1.00 for model identification purposes; Est.: standardized estimate; SE: standard error.

variance extracted (AVE) to examine its reliability and convergent validity (Malhotra & Dash, 2011) (Table 2). As all the CR and AVE values were above the 0.70 and 0.50 thresholds respectively, we concluded that there were no concerns (Hair et al., 2019). We used the heterotrait-monotrait ratio of correlations (HTMT) technique for assessing its discriminant validity (Henseler, Ringle, & Sarstedt, 2015), which was confirmed given that all the HTMT values were below the -strict- threshold of 0.85 (Table 3). We also employed Cronbach's α as an auxiliary measure to assess the reliability of the scale. The cumulative α value was calculated to be 0.929. Detailed results for each individual factor can be found in Table 3.

For model fit assessment, we used several indices as presented in Table 4 (together with their respective thresholds, as provided by Hu & Bentler, 1999). Evidently, all indicated an excellent model fit. As a final note, we assessed the reliability of the scale's items and constructs using Cronbach's alpha. We found no issues as α ranged from 0.797 to 0.925, while the overall α was 0.929.

Summarizing the results of the CFA, we can argue that the MLES's validity and reliability were re-confirmed. We present the version that emerged during the previous analyses with the 39 retained items in Appendix II. The final step was to calculate ten more variables, representing the average score of the items in each factor (Table 5).

5.3. Structural equation modeling

As all the requirements for conducting SEM were met, we conducted the analysis using AMOS 28. We observed that the direct effects of several structural paths were statistically insignificant and/or their coefficients rather small, suggesting that they should be removed. To confirm which paths to retain, we run an analysis using the Specification Search Facility (available in AMOS). We applied the Bayesian Information Criterion (BIC) for selecting the final model, as BIC allows for the selection of the most robust, yet parsimonious models (Claeskins & Hkort, 2008). Tables 6 and 7, as well as Fig. 3, present details about the final model ($BIC = 0.00$).

We estimated all possible indirect effects in the model using an AMOS plugin developed by Gaskin and Lim (2018). Given that we expected all the values of the indirect path coefficients to be low (because a decimal number is multiplied by another decimal number), we present in Table 6 only the most notable ones ($\beta \geq 0.08$). The final model's fit indices were excellent (χ^2/df = 1.59, CFI = 0.95, SRMR = 0.05, RMSEA = 0.04, and PClose = 0.998).

We conducted the post-hoc power analysis for unsupported direct effects utilizing the method articulated by Soper (2016). For the twelve predictors of the dependent variable (including age and gender), an observed R^2 of 0.66 for this factor, a probability level of 0.05, and a sample size of 334, the statistical power was computed to be 1.00. As such, the model manifests robust prowess in the identification of significant effects. It's also worth noting that the non-significant effects that surfaced during the analysis were indeed statistically insignificant.

Table 2
Reliability and convergent validity.

Factor	CR	AVE	α	Factor	CR	AVE	α
Motivation	0.91	0.63	.915	Positive Feelings	0.95	0.81	.872
Simulator sickness	0.88	0.55	.870	Control	0.88	0.70	.904
Immersion/ Presence	0.87	0.62	.878	Interaction	0.84	0.64	.839
Cognitive Load	0.80	0.67	.797	Gains	0.91	0.78	.925
Graphics	0.89	0.62	.891	Feedback/ Content	0.84	0.64	.843

6. Discussion

6.1. General comments

By examining Table 5, the first thing to note is that the participants achieved a rather good mean score on the quizzes [$M = 77.20$ (max. = 100), $SD = 15.30$]. Considering that we expected participants not to have any prior knowledge about the inventions we presented to them, this might be an indication that IVR's application had a positive impact on learning as many others suggested (Atsikpasi & Fokides, 2022; Tlili et al., 2022). In addition, although the application's graphics were not so impressive, because of technology restrictions, the perceived graphics' quality factor had the highest mean amongst all ($M = 4.46$, $SD = 0.62$). Given the rather high mean values, we can support the view that participants assessed the degree of perceived interaction and perceived quality of feedback and content a being quite high ($M = 4.12$, $SD = 0.75$ and $M = 4.43$, $SD = 0.61$ respectively). We can also suggest that they found their experience motivating and that it generated positive feelings ($M = 4.30$, $SD = 0.71$ and $M = 4.21$, $SD = 0.87$ respectively). In this respect, we can confirm the credibility of previous studies with similar findings (e.g., Butt et al., 2018; Caro et al., 2018; Go et al., 2021; Jeon & Jung, 2021, pp. 361–368; Rupp et al., 2019). Moreover, they found

Table 3
HTMT analysis.

Factor	Motivation	Simulator sickness	Immersion /Presence	Cognitive Load	Graphics	Positive Feelings	Control	Interaction	Gains	Feedback/ Content
Motivation										
Simulator sickness	0.18									
Immersion /Presence	0.40	0.01								
Cognitive Load	0.16	0.32	0.03							
Graphics	0.60	0.26	0.50	0.16						
Positive Feelings	0.60	0.20	0.28	0.15	0.63					
Control	0.56	0.25	0.34	0.19	0.47	0.46				
Interaction	0.57	0.11	0.51	0.13	0.62	0.50	0.59			
Gains	0.61	0.20	0.31	0.32	0.42	0.41	0.38	0.42		
Feedback/ Content	0.63	0.08	0.48	0.24	0.61	0.53	0.47	0.61	0.61	

Table 4
Model fit assessment.

Measure	Estimates	Thresholds
χ^2	1011.95	–
DF	642.00	–
χ^2/df	1.58	Between 1 and 3
CFI	0.96	>.95
SRMR	0.05	<.08
RMSEA	0.04	<.06
PClose	0.998	>.050

Notes. CFI: comparative fit index, SRMR: standardized root mean square residual, RMSEA: root mean square error of approximation, χ^2/df: minimum discrepancy divided by its degrees of freedom.

Table 5
Descriptive statistics for the quiz and scale's factors.

Variable ($N = 334$)	*Min.*	*Max.*	*Mean*	*SD*
Learning Outcomes (evaluation quizzes scores)	30.00	100.00	77.20	15.30
Graphics	2.20	5.00	4.46	0.62
Cognitive Load	1.00	5.00	1.69	0.84
Control	1.00	5.00	3.93	0.87
Immersion/Presence	1.00	5.00	3.67	1.07
Feedback/Content	2.67	5.00	4.43	0.61
Interaction	2.00	5.00	4.12	0.75
Motivation	1.67	5.00	4.30	0.71
Gains	1.00	5.00	4.07	0.85
Positive Feelings	1.00	5.00	4.21	0.87
Simulator Sickness	0.00	0.67	0.08	0.12

Note. Simulator sickness ranged from 0.00 to 1.00 because it was transformed.

cognitive load to be rather low ($M = 1.69$, $SD = 0.84$), as others noted (e.g., Cecotti et al., 2020, pp. 16–23). On the negative side, it seems that they did not become so immersed ($M = 3.67$, $SD = 1.07$). As for the perceived ease of use/control of the virtual environment, the mean fell slightly below the value of 4.00 ($M = 3.93$, $SD = 0.87$). This finding suggests that the participants did not face significant usability problems, but some work needs to be done in this direction.

Chin (1998) suggested that for a model to have meaningful predictive power, it has to be able to explain a significant percentage of the variance in its dependent variables (as indicated by their R^2s). For social sciences, R^2 values greater than 0.50 can be considered strong (Hair & Alamer, 2022). Furthermore, the values of the model's structural paths have to be substantial (β preferably close to 0.20 and ideally above 0.30), although smaller values cannot be ignored when they are statistically significant (Chin, Marcolin, & Newsted, 2003).

As our objective was to develop a model for explaining learning outcomes, we can assume that its predictive power regarding this factor is more than satisfactory ($R^2 = 0.660$). The perceived feedback and content quality had a strong direct positive impact on this factor ($\beta = 0.28$), as well as an indirect negative one through the perceived cognitive load and simulator sickness ($\beta = -0.08$). We noted both the direct and indirect impact (through positive feelings and motivation) of the perceived graphics quality ($\beta = 0.17$ and $\beta = 0.12$ respectively). The degree of interaction had a lesser direct impact ($\beta = 0.14$). Both users' immersion/presence and their motivation had a positive impact as well ($\beta = 0.13$ and $\beta = 0.16$ respectively). As expected, the perceived cognitive load had a negative effect ($\beta = -0.14$), and the same applied for simulator sickness, although the effect was stronger ($\beta = -0.20$).

We also explained motivation rather well ($R^2 = 0.580$). The perceived feedback and content quality had a strong impact on this factor ($\beta = 0.28$). The perceived graphics' quality had both a direct positive effect ($\beta = 0.18$) and an indirect one through positive feelings ($\beta = 0.12$). The perceived ease of use/control had a noticeable positive impact ($\beta = 0.23$) and the same applied for user' positive feelings ($\beta = 0.26$).

We moderately explained immersion/presence ($R^2 = 0.388$). The perceived feedback and content quality, as well as the graphics quality, and degree of interaction, had substantial positive impacts ($\beta = 0.23$, $\beta = 0.25$, and $\beta = 0.24$ respectively). What is interesting is that we found that perceived cognitive load had a positive effect, although not that strong ($\beta = 0.17$). One's gender also played a role ($\beta = 0.14$).

As for perceived knowledge gains, we moderately explained this factor ($R^2 = 0.427$). The perceived quality of feedback and content had considerable direct ($\beta = 0.28$) and indirect positive effects (through motivation, $\beta = 0.10$). The perceived graphics' quality had an indirect positive effect through users' positive feelings and motivation ($\beta = 0.12$). Indirect were the positive effects of the perceived ease of use/control and one's positive feelings (through motivation, $\beta = 0.08$ and $\beta = 0.09$ respectively). The perceived cognitive load had a negative effect ($\beta = -0.20$), while users' motivation had a rather impressive direct positive effect ($\beta = 0.36$).

We moderately explained participants' positive feelings ($R^2 = 0.477$). The perceived feedback and content quality had a noticeable positive effect ($\beta = 0.27$), while the path linking the perceived graphics' quality to this factor was not only positive but also the most impressive one among all the other paths present in our model ($\beta = 0.48$). The perceived ease of use/control had a positive impact ($\beta = 0.16$). What we consider as an unexpected finding is that immersion/presence had a negative impact ($\beta = -0.14$).

Also, we modestly explained the variance in simulator sickness ($R^2 = 0.206$). Again, an unexpected finding was that the perceived quality of feedback and content had a positive impact on this factor and a rather strong one ($\beta = 0.28$). Then again, it had an indirect negative impact through the perceived cognitive load ($\beta = -0.08$). The perceived graphics' quality and ease of use/control had both strong negative

Table 6
The results of SEM, direct and indirect effects.

RQ	Factor	To	Factor	*Est.*	*SE*	*t*	*P*	*β*
RQ1a	Feedback/Content	→	Cognitive Load	−.33	.10	−3.51	<.001	−.23
RQ1b	Feedback/Content	→	Simulator Sickness	.36	.11	3.22	.001	.28
–	Feedback/Content– > Cognitive Load– > Simulator Sickness						.006	−.08
RQ1c	Feedback/Content	→	Immersion/Presence	.38	.14	2.80	.005	.23
RQ1d	Feedback/Content	→	Positive Feelings	.39	.10	3.87	<.001	.27
RQ1e	Feedback/Content	→	Motivation	.34	.08	4.36	<.001	.28
RQ1f	Feedback/Content	→	Gains	.47	.12	3.96	<.001	.28
–	Feedback/Content– > Motivation– > Gains						<.001	.10
RQ1g	Feedback/Content	→	Learning Outcomes	.77	.18	4.24	<.001	.28
–	Feedback/Content– > Cognitive Load– > Simulator Sickness– > Learning Outcomes						.007	−.08
RQ2a	Graphics	→	Cognitive Load	–	–	–	–	–
RQ2b	Graphics	→	Simulator Sickness	−.31	.10	−3.24	.001	−.26
RQ2c	Graphics	→	Immersion/Presence	.37	.12	3.18	.001	.25
RQ2d	Graphics	→	Positive Feelings	.63	.09	6.76	<.001	.48
RQ2e	Graphics	→	Motivation	.19	.07	2.70	.007	.18
–	Graphics– > Positive Feelings– > Motivation						.007	.12
–	Graphics– > Positive Feelings– > Motivation– > Gains						.005	.12
RQ2f	Graphics	→	Gains	–	–	–	–	–
RQ2g	Graphics	→	Learning	.42	.15	2.87	.004	.17
–	Graphics– > Positive Feelings– > Motivation– > Learning Outcomes						.009	.12
RQ3a	Control	→	Cognitive Load	–	–	–	–	–
RQ3b	Control	→	Simulator Sickness	−.19	.07	−2.96	.003	−.21
RQ3c	Control	→	Immersion/Presence	–	–	–	–	–
RQ3d	Control	→	Positive Feelings	.16	.06	2.73	.006	.16
RQ3e	Control	→	Motivation	.19	.05	4.11	<.001	.23
–	Control– > Motivation– > Gains						<.001	.08
RQ3f	Control	→	Gains	–	–	–	–	–
RQ3g	Control	→	Learning Outcomes	–	–	–	–	–
RQ4a	Interaction	→	Cognitive Load	–	–	–	–	–
RQ4b	Interaction	→	Simulator Sickness	–	–	–	–	–
RQ4c	Interaction	→	Immersion/Presence	.32	.12	2.78	.005	.24
RQ4d	Interaction	→	Positive Feelings	–	–	–	–	–
RQ4e	Interaction	→	Motivation	–	–	–	–	–
RQ4f	Interaction	→	Gains	–	–	–	–	–
RQ4g	Interaction	→	Learning Outcomes	.32	.14	2.38	.017	.14
RQ5a	Cognitive Load	→	Simulator Sickness	.30	.06	4.75	<.001	.33
RQ5b	Cognitive Load	→	Immersion/Presence	.20	.07	2.97	.003	.17
RQ5c	Cognitive Load	→	Positive Feelings	–	–	–	–	–
RQ5d	Cognitive Load	→	Motivation	–	–	–	–	–
RQ5e	Cognitive Load	→	Gains	−.24	.06	−3.83	<.001	−.20
RQ5f	Cognitive Load	→	Learning Outcomes	−.27	.09	−3.13	.002	−.14
RQ6a	Simulator Sickness	→	Immersion/Presence	–	–	–	–	–
RQ6b	Simulator Sickness	→	Positive Feelings	–	–	–	–	–
RQ6c	Simulator Sickness	→	Motivation	–	–	–	–	–
RQ6d	Simulator Sickness	→	Gains	–	–	–	–	–
RQ6e	Simulator Sickness	→	Learning Outcomes	−.43	.09	−4.61	<.001	−.20
RQ7a	Immersion/Presence	→	Positive Feelings	−.13	.05	−2.40	.016	−.14
RQ7b	Immersion/Presence	→	Motivation	–	–	–	–	–
RQ7c	Immersion/Presence	→	Gains	–	–	–	–	–
RQ7d	Immersion/Presence	→	Learning Outcomes	.21	.08	2.59	.010	.13
RQ8a	Positive Feelings	→	Motivation	.21	.05	4.21	<.001	.26
–	Positive Feelings– > Motivation– > Gains						.007	.09
RQ8b	Positive Feelings	→	Gains	–	–	–	–	–
RQ8c	Positive Feelings	→	Learning Outcomes	–	–	–	–	–
RQ9a	Motivation	→	Gains	.51	.10	5.35	<.001	.36
RQ9b	Motivation	→	Learning Outcomes	.38	.13	2.96	.003	.16
RQ10	Gains	→	Learning Outcomes	–	–	–	–	–
RQ12c	Gender	→	Immersion/Presence	.28	.10	2.79	.005	.14

Notes. *Est.*: standardized estimate; SE: standard error β: path coefficient; for the indirect effects the analysis calculated only the p and β values; the highlighted rows indicate the rejected paths; for clearance of presentation, we included in the table just the only significant path of the control variables, as all the paths suggested by RQ11 and RQ12a, b, and d-g were rejected.

effects ($\beta = -0.26$ and $\beta = -0.21$ respectively). The positive impact of the perceived cognitive load was rather strong ($\beta = 0.33$).

Finally, we failed to explain the variance in perceived cognitive load ($R^2 = 0.057$), probably because the perceived quality of feedback and content was the only one to affect this factor, having a substantial negative impact ($\beta = -0.23$).

From the above, we can conclude that our model was able to explain a significant percentage of the variance in several factors and that several of our model's structural paths had significant values, further solidifying its credibility.

6.2. *Comments on the confirmed and unconfirmed paths*

RQ1, the effects of the perceived feedback and content quality. We found that the perceived quality of feedback and content affected all the factors we hypothesized it might affect. Specifically, it had an impact on the perceived cognitive load ($\beta = -0.23$), simulator sickness (directly, $\beta = 0.28$ and indirectly through the perceived cognitive load $\beta = -0.08$), immersion/presence ($\beta = 0.23$), positive feelings ($\beta = 0.27$), motivation ($\beta = 0.28$), perceived knowledge gains (directly, $\beta = 0.28$ and indirectly through motivation, $\beta = 0.10$), and the learning outcomes (directly $\beta =$

Table 7
The results of SEM, squared multiple correlations (R^2).

Factor	R^2	Short interpretation
Cognitive Load	.057	Weak explanation of the variance
Simulator Sickness	.206	Modest explanation
Immersion/Presence	.388	Moderate explanation
Gains	.427	Moderate explanation
Positive Feelings	.477	Moderate explanation
Motivation	.580	Strong explanation
Learning Outcomes	.660	Strong explanation

Note. Thresholds and interpretations suggested by Hair and Alamer (2022).

0.28 and indirectly through perceived cognitive load and simulator sickness, $\beta = -0.08$).

The negative impact on cognitive load can serve as an indicator that we adequately designed and presented the learning content, to achieve a minimal impact on the participants working memory as others advised (Leppink et al., 2014; Novak et al., 2018). Indeed, the participants assessed its quality as being rather good (M = 4.43, SD = 0.61). Then again, the perceived feedback and content quality increased users' symptoms of simulator sickness. We suspect that this happened not because of the quality, but because of the learning content *per se*. On the one hand, the participants had to put some effort into studying the learning material. On the other hand, they had to put some effort into coping with the virtual environment (e.g., navigating and interacting). So, they were probably disorientated or overwhelmed at some point, allowing for simulator sickness symptoms to emerge. The fact that we found an indirect negative effect through the perceived cognitive load, may serve as evidence for the validity of our assumption. That is because when users considered the quality of feedback and content as being good and, at the same time, they assessed the cognitive load as being low, this decreased the chances of them having simulator sickness.

As for the remaining effects of this factor, they give further support to the findings of previous studies indicating that the sense of realism is subject to the design of feedback cues (Faeth & Harding, 2014), that these cues can positively affect user experience and task accuracy (Faeth & Harding, 2014; Kobayashi et al., 2016), and that the quality of teaching content had an impact on learning (Portman et al., 2015; Potkonjak et al., 2016).

RQ2, the effects of the perceived quality of the virtual environment's graphics. This factor was the second most influential one in our model, as there were just two missing paths. It affected simulator sickness ($\beta = -0.26$), immersion/presence ($\beta = 0.25$), positive feelings ($\beta = 0.48$), motivation (directly, $\beta = 0.18$ and indirectly through positive feelings, $\beta = 0.12$), perceived knowledge gains (indirectly through positive feelings and motivation, $\beta = 0.12$), and the learning outcomes (directly $\beta = 0.17$ and indirectly through positive feelings and motivation, β = 12). We expected these multiple effects, as there is literature suggesting that visual fidelity results in decreased symptoms of simulator sickness (de Winkel et al., 2022), increased immersion (Mystakidis, 2022; Parong & Mayer, 2018), increased motivation (Lee et al., 2010), and increased performance (e.g., Harrington, 2012; Kim & Ahn, 2021). To the above, we can add that the graphics' quality greatly increases one's enjoyment and positive feelings (given that the β value of this path was the most impressive one).

On the other hand, we found no direct or indirect effects on cognitive load and perceived knowledge gains, although some suggested that representation fidelity affected users perceived cognitive benefits (Makransky & Petersen, 2019). As for the cognitive load, we assumed that if the graphics are good, meaning that the 3D objects accurately represent the real ones (in our case the ancient Greek inventions), this will help students to have less trouble understanding their function,

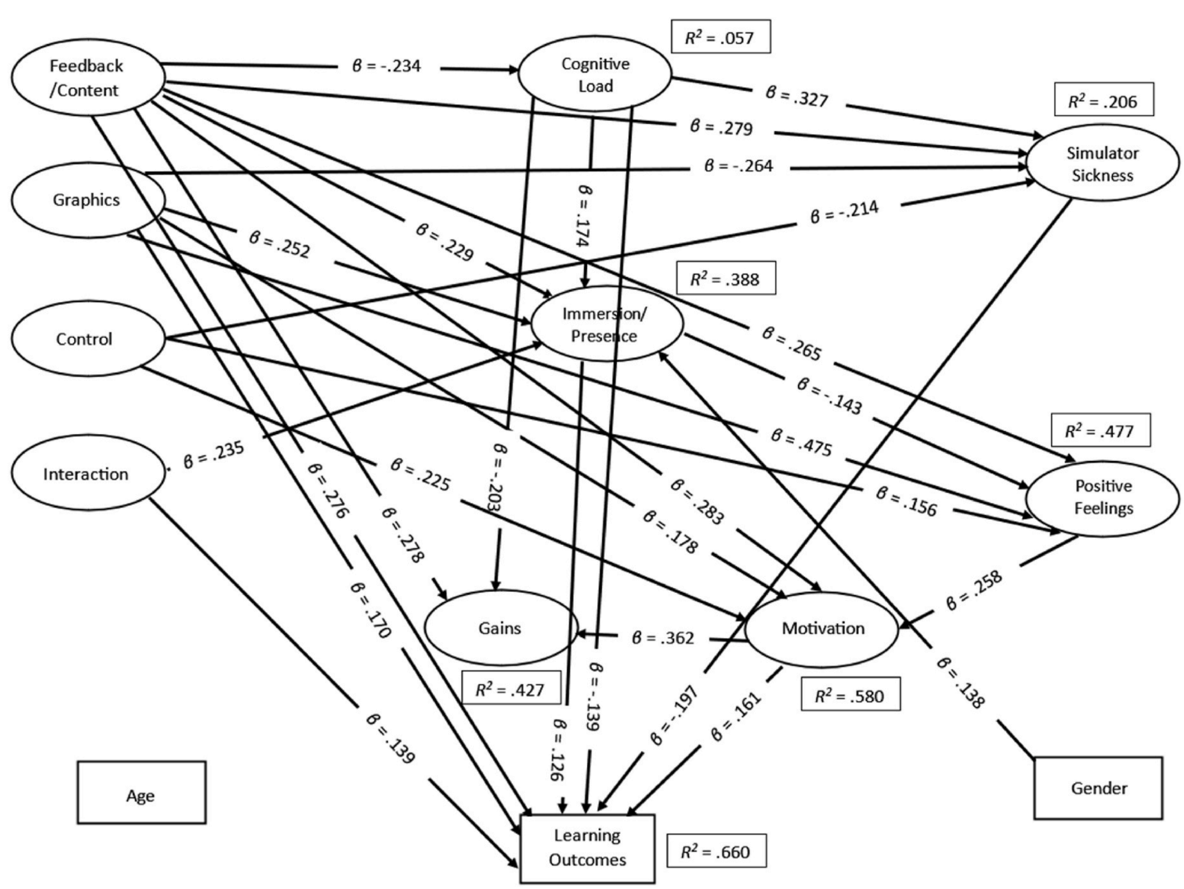

Fig. 3. The study's final model presenting the variables' significant paths and explained variances.

which, in turn, will help to alleviate their cognitive load. Alas, our assumption was not confirmed, and we have to conclude that the quality of the graphics does not affect cognitive load (neither positively nor negatively).

RQ3, *the effects of perceived ease of use/control*. We noticed a negative impact of this factor on simulator sickness ($\beta = -0.21$). We have to remind readers that instead of using controllers for their movement, the participants could walk. We suspect that this allowed them to find the control of the application more natural and easier and, at the same time, reduced their chances of having simulator sickness. We also noted a positive impact on users' positive feelings ($\beta = 0.16$). In a broader sense, this factor reflects users' attitudes, which, according to the TAM (Davis et al., 1989), is, indeed, affected by ease of use. As for motivation, we found that ease of use/control positively affected it ($\beta = 0.23$).

What we did not find were links between the perceived ease of use/control and the perceived cognitive load, immersion/presence, perceived knowledge gains, motivation, and learning outcomes. We based these assumptions on past research indicating that the perceived ease of use had an impact on the perceived cognitive load (Ibili & Billinghurst, 2019), motivation (together with other VR features, Makransky & Lilleholt, 2018), and the learning outcomes (e.g., Asad et al., 2022; Fokides, 2017; Fokides & Atsikpasi, 2018; Lee et al., 2010). Yet, we found just an indirect effect on perceived knowledge gains through motivation ($\beta = 0.08$). Given that, we conclude that the perceived ease of use and control of the virtual environment is a motivating factor, but without any actual impact on learning. This conclusion is in line with Dalgarno, Hedberg, and Harper (2002) suggestion that control over the virtual environment is not enough for one to achieve conceptual understanding because the tasks' design is more important. As for immersion/presence, on the basis of our results, we are inclined to believe that this relationship does not exist, although it seemed logical to assume that when users find a device or application easy to use this will help them to become more immersed in the virtual environment.

RQ4, the effects of interaction. We expected the increased levels of the perceived interaction to have an effect on multiple factors. Yet, we found no direct or indirect effects on the perceived cognitive load, simulator sickness, positive feelings, and perceived knowledge gains. We also did not find any relationship between perceived interaction and motivation, although others indicated that there is such a link (e.g., Diaz et al., 2020; Lee et al., 2010). Given that, we conclude that the degree of the perceived interaction with the virtual environment does not act as a motivating factor and that it is unrelated to simulator sickness.

Although it did not motivate participants, their perceived degree of interaction had a positive impact on their learning ($\beta = .14$), as indicated by other studies as well (Diaz et al., 2020; Portman et al., 2015; Potkonjak et al., 2016), probably because it allowed students to become active learners (Diaz et al., 2020; Mystakidis, 2022). In addition, the degree of interaction had a noticeable effect on immersion/presence ($\beta = 0.24$). In this respect, we believe that interaction adds to the believability of the virtual environment, allowing users to become more immersed, as Diaz et al. (2020) also noted.

RQ5, the effects of cognitive load. We noted that the perceived cognitive load positively affected simulator sickness ($\beta = 0.33$). This means that when users considered their cognitive load as being high, the worse were their symptoms of simulator sickness. As we are not aware of literature suggesting this link and given that the effect was considerable, our best guess is that there is a strong mechanism connecting one's effort to understand the learning content with the manifestation of simulator sickness symptoms, worth further investigation. We also found that the perceived cognitive load had a positive impact on immersion/presence ($\beta = 0.17$), meaning that the higher the cognitive load the higher the users' immersion/presence. This is also a hard-to-explain finding, as we expected a negative impact. Quite interestingly, Makransky et al. (2017) in their model examined the exact opposite direction and concluded that immersion led to increased cognitive load. As cognitive load and immersion/presence require mental effort, we can assume that users' brains process them similarly and have a two-way relationship.

Then again, the perceived cognitive load did not affect users' positive feelings and motivation. In line with past research (Cecotti et al., 2020, pp. 16–23), the participants in our study assessed their cognitive load as being low ($M = 1.69$, $SD = 0.84$). Given that, we think that it is logical not to have affected these two factors, although there are studies suggesting that it can affect both motivation (Johnson, 2005) and engagement (Atsikpasi & Fokides, 2022; Maraj et al., 2017). Nevertheless, as expected and in line with past research (e.g., Huang et al., 2022; Makransky et al., 2017), this factor negatively impacted one's perceived knowledge gains ($\beta = -0.20$) and learning outcomes ($\beta = -0.14$).

RQ6, the effects of simulator sickness. To be honest, we expected simulator sickness to have multiple direct and indirect negative effects. That is because there is quite rich literature suggesting that it affects several factors such as motivation (Johnson, 2005), learning, presence, and engagement (e.g., Atsikpasi & Fokides, 2022; Maraj et al., 2017), even if the symptoms are mild (Hsin et al., 2022). On the other hand, in our study, there were just a few cases of severe and mild simulator sickness, as indicated by the mean in this factor [$M = 0.08$ (max. $=$ 1.00), $SD = 0.12$]. This was probably due to the way we designed the application. For example, users could walk instead of using the controllers for their movement; therefore, no conflicting signals were coming to their brains from their motor, visual, and vestibular systems. Moreover, we believe that the reduced number of symptoms did not allow the negative effects of this factor to emerge. On the other hand, simulator sickness had a negative impact on the learning outcomes ($\beta = -0.20$), as others noted (Atsikpasi & Fokides, 2022; Hsin et al., 2022; Maraj et al., 2017). Given that, we are inclined to believe that even mild symptoms can negatively impact learning outcomes, as suggested by Hsin et al. (2022).

RQ7, the effects of immersion/presence. From the analyses of our data, two effects emerged regarding this factor; one on participants' positive feelings and another on their learning. We find the result regarding the impact on the former factor a bit troubling, as it was a negative one ($\beta = -0.14$). This means that the higher the participants' immersion/presence, the less their positive feelings. This finding contradicts the ones of previous studies (e.g., Lehikko, 2021; Parong & Mayer, 2021), which suggested that users immersive/interactive experiences led to more enjoyable learning. This finding is also odd as the mean in the positive feelings factor was quite good ($M = 4.21$, $SD = 0.87$), indicating that the participants found their experience enjoyable. Moreover, we found that immersion/presence did not have any effect on one's motivation and perceived knowledge gains. Therefore, we cannot confirm the findings of previous studies indicating that immersion is a motivating factor or that it can lead to higher perceived learning outcomes, although studies suggested that (e.g., Diaz et al., 2020; Huang et al., 2022; Makransky & Lilleholt, 2018; Makransky & Petersen, 2019).

We can only speculate why we had these findings. First, the participants' immersion in the virtual environment was not that high ($M = 3.67$, $SD = 1.07$). Second, we suspect that users were immersed in a virtual environment they did not find enjoyable because they were not so interested in what it presented. In fact, the application was about ancient technology; by itself, the topic although educational, cannot be considered entertaining. As a result, users' mental and emotional engagement with the virtual environment, because of immersion/presence, was not enough to generate positive feelings, motivate, or make them believe that they learned.

Despite all these, immersion/presence had a positive impact on the learning outcomes as several others suggested (e.g., Atsikpasi & Fokides, 2022; Jeon & Jung, 2021, pp. 361–368; Kim & Ahn, 2021; Lee et al., 2010; Maas & Hughes, 2020), although the effect was not a prominent one ($\beta = 0.13$). Taking together all the above, we can conclude that even if immersion/presence does not act as a motivating factor and even if it does not make the learning experience enjoyable, it can still impact learning, as it probably helps individuals to better understand the learning content.

RQ8, the effects of positive feelings. We found that what can be characterized as positive feelings (i.e., the items included in the MLES scale, namely enjoyment, satisfaction, enthusiasm, and excitement) acted as a quite significant motivational factor ($\beta = 0.26$). Such a link was suggested by others (Barry et al., 2015) who assumed that enjoyable VR projects enhanced students' motivation. Then again, contrary to what the literature suggested (e.g., Kim & Ahn, 2021; Lehikko, 2021; Parong & Mayer, 2021), we did not find any relationship between users' positive feelings and learning or perceived knowledge gains, though we found an indirect effect on the latter through motivation ($\beta = .09$). Therefore, we are inclined to believe that one's positive feelings act more as a motivating factor and less as a direct learning facilitator.

RQ9, the effects of motivation. We noted a rather strong effect on users' perceived knowledge gains ($\beta = 0.36$) and a weaker one on their learning ($\beta = 0.16$). The link between motivation and learning is theoretically established (e.g., the interest theory, Renninger & Hidi, 2017). It has also been empirically established in the context of VR (e.g., Fokides, 2017; Fokides & Atsikpasi, 2018), although others did not find this link (Makransky & Petersen, 2019). There is also evidence for the link between motivation and perceived knowledge gains (Makransky & Lilleholt, 2018). Taking together the above, the fact that we explained the variance in motivation rather well, and the fact that the participants found their experience motivating, we can conclude that motivation is a factor affecting both the perceived and actual learning.

RQ10, *the effect of the perceived knowledge gains on the learning outcomes*. We found that there is neither a positive nor a negative effect. Consequently, we conclude that there is a discrepancy between what learners actually learn (or do not learn) and what they think they learned.

RQ11 and RQ12, the effects of age and gender. The participants' age did not seem to have any effect whatsoever. We do not consider this finding surprising. First, our sample was rather homogeneous in terms of age, as the participants were university students. Second, age differences among Greek educators regarding the use of ICT tools in general, seem to be diminishing as indicated in a recent study (Fokides & Kapetangiorgi, 2022).

As for gender, it had an effect only in immersion ($\beta = 0.14$), with females being more subjective to this feeling. Although the effect was not that strong and our sample was unbalanced in terms of gender distribution, we think that this finding calls for further research. Indeed, while gender and immersion have been studied in other contexts (e.g., in multiplayer online games, Stavropoulos, Rennie, Morcos, Gomez, & Griffiths, 2021) and while there is an indication that there are differences between males and females regarding the feeling of presence (Annetta, Klesath, & Meyer, 2009), it seems that few examined gender differences regarding immersion in immersive VR environments (e.g., Selcuk, 2021).

6.3. *Implications for research and practice*

Our findings have implications for individuals involved in the industry, researchers, as well as education stakeholders. While a moderate portion of the variance in immersion was explained, a significant percentage remained unaccounted for. This suggests that additional factors, beyond those we identified, also influence immersion. Therefore, researchers can explore other factors that impact this aspect, which can assist those in the VR industry in delivering more immersive experiences.

We observed a significant direct effect of the perceived degree of interaction and graphics quality on learning outcomes, along with an indirect effect of perceived ease of use. Furthermore, these factors had notable impacts on other variables. All three factors can be categorized as technological features. As a result, software developers and computer engineers can explore methods to enhance these features, not only to augment their impact on learning but also to facilitate the widespread adoption of related technologies. Additionally, perceived feedback and content quality had substantial effects on various factors beyond learning outcomes. Given their central role, both software developers and educators must find ways to present learning content more effectively and provide adequate feedback to learners.

Regarding simulator sickness, we found that although it was not highly pronounced, it did have a noticeable impact on learning outcomes. As mentioned in a previous section, simulator sickness can be considered an undesirable consequence of the technology used. Various methods for reducing simulator sickness can be employed. For instance, one commonly used method is teleportation from one point to another. However, this technique may negatively affect immersion when movement in the virtual world is not swift (Griffin & Folmer, 2019). Therefore, researchers and engineers should explore and test alternative methods.

During the implementation of our project, we observed that none of the participants had prior experience using HMDs, and few had an understanding of what IVR is. It is important to note that our sample consisted of students in a department of education. Education stakeholders can take several steps to rectify this issue. At an academic level, courses related to the educational applications of technologies and IVR-related applications should be included. In-service training programs can also be beneficial. Additionally, research aimed at identifying effective teaching methods that leverage IVR's potential is necessary.

6.4. *Limitations and future studies*

Our study is subject to several limitations. A larger sample size would have provided us with greater confidence in our results and the model's reliability. Due to our utilization of a convenience sampling method, there may be associated concerns regarding the generalizability of the results in different populations. As always, concerns about the responders' trustworthiness arise. Moreover, the MLES scale, although retested in this study, is newly developed and not established. We examined ten factors (excluding the learning outcomes) and two control variables. We undoubtedly omitted certain factors, rendering the model somewhat incomplete in terms of inclusivity. Our data collection method was primarily based upon the subjective experiences of our participants. We centered our methods around users' emotions, perceptions, and attitudes, which can potentially be influenced by a countless array of factors that fall beyond our scope. Therefore, a more objective approach -perhaps implementing specialized equipment for data collection (e.g., eye-tracking devices and electroencephalograms)- could potentially allowed for more reliable and definitive conclusions. Our focus was on university students and a specific learning subject, thus questioning the generalizability of the results. One might argue that the time allocated for each participant to interact with the application was insufficient. However, managing more than 300 participants was challenging, making it impossible to extend the interaction time or sessions.

Future studies should take these limitations into account. Diverse target populations, both in terms of age and education level, can help identify similarities or differences compared to the present model. This holds true for various learning domains and types of applications. Collaboration is certainly a factor to consider when testing multiuser applications. Examining self-efficacy becomes relevant when groups with different IVR device-related skills are included in the sample. The addition of other factors is also worth considering, but we advise caution. While adding factors may enhance the explanation of variance in other factors and learning outcomes, it can exponentially increase a model's complexity, potentially compromising its stability and result clarity. Comparative studies necessitating the contrast of learning outcomes from IVR usage with those attained from other technologies are of significant importance to gain an in-depth understanding of IVR's actual potential. Additionally, assessing varying age groups and learning domains is also crucial for the same purpose. Furthermore, given that our study did not incorporate any form of structured instruction in conjunction with the use of the IVR application, examining the potential impact on the interactions of various factors and learning outcomes should IVR be integrated into a meticulously planned teaching scenario is a subject worth of further investigation. Finally, we believe that an in-depth understanding of IVR's educational potential can be achieved through longitudinal studies. Studies with a limited number of sessions are susceptible to the "wow effect," wherein users exhibit enthusiasm when encountering a technological artifact they have never previously experienced (Kamstrupp, 2016), which may influence the results.

7. Conclusion

IVR, as the next step in the evolution of VR technologies, has the potential to expand the boundaries of human activities facilitated by current ICT tools, including education. Given this, one must understand how learners acquire knowledge in immersive environments, such as the one utilized in this project. This constituted our primary objective, resulting in a model that represents the interactions of factors we deemed crucial in shaping one's learning. Data were amassed via a validated scale, designed to assess ten subjective factors and users' attitudes and feelings postulated to influence learning outcomes in IVR educational applications. The choice and subsequent integration of these elements stemmed from a comprehensive review of literature. This review focused on studies with a primary objective of evaluating the factors that influence learning within both IVR and VR environments in general. In summary, our findings indicated that the perceived quality of graphics, feedback, and content, in conjunction with increased perceived interaction, motivation, and the immersive experience offered by IVR applications, positively influenced learning outcomes. We also observed a negative impact from perceived cognitive load and simulator sickness. Interestingly, both of these factors did not affect users' positive emotions and motivation. Stakeholders in education, research, and the computer/software industry may find our study's results valuable when designing and implementing IVR applications.

Funding

The study received no funding.

Data availability

The data can be provided upon request by contacting the corresponding author.

Ethical approval

The study's participants volunteered, understood that they could withdraw from the experiment at any time, and provided their informed consent. They were protected by hiding their personal information in this study. The study is part of an ongoing research project for which approval from the Department's Ethical Committee has been granted.

Ethical statement

We hereby declare that this manuscript is the result of our independent creation under the reviewers' comments. Except for the quoted contents, this manuscript does not contain any research achievements that have been published or written by other individuals or groups, or by AI tools; we are the only authors of the manuscript. The legal responsibility of this statement shall be borne by us.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix I. List of factors examined in past research and their grouping

Initial factor	Grouped factor	Initial factor	Grouped factor
Realism	Perceived quality of the virtual environment's graphics	Presence	Immersion/Presence
Visual aesthetics		Spatial presence	
Attractiveness		Social presence	
Aesthetics		Physical presence	
Audiovisual appeal		Core self-presence	
System quality		Extended self-presence	
System attributes		Proto self-presence	
Visual aesthetic		Immersion	
Environment		Attention	
Pragmatic quality		Involvement	
Media richness		Sensory immersion	
Variety		Flow	
Cognitive load	Cognitive load	Distraction	
Extrinsic cognitive load		Absorption	
Information overload		Concentration	
Perceived complexity		Empathy	
Control	Ease of use/Control of the virtual environment	Dissociation	
Autonomy		Non-mediation	
Mastery		Action awareness merging	
Ownership		Loss of self	
Misuse		Focused immersion	
Usability		Perspicuity	
Ease of use		Extended self	
Ease of control		Focus	
Long learning phase		Engagement	
Navigation		Internal/external correspondence	
Pragmatic quality		illusion	
Controls		Imagination	
Operator		Empathic concerns	
Consistency		Autotelic experience	
No bugs/errors		Fantasy	

(*continued on next page*)

(*continued*)

Initial factor	Grouped factor	Initial factor	Grouped factor
Understandability		Action awareness	
System naturalness		Loss of self-consciousness	
System responsiveness		Relatedness	
Playability		Plausibility	
Camera		Focused attention	
Gameplay		Autotelic focus	
Gaming		External correspondence	
Enjoyment	Positive feelings	Temporal dissociation	
Playfulness		Behavioral engagement	
Stimulation		Cognitive engagement	
Pleasure		Spatial awareness	
Fun		Core self	
Hedonic quality		Emotional attachment	
Novelty		Transformation of time	
Satisfaction		Emotional engagement	
Challenge		Reality judgment	
Positive affection		Social experience	
Pride		Possible action	Perceived degree of interaction
Increase status		Freedom	
Happy		Creative freedom	
Captivation		Simplicity	
Hope		Paradox of control	
Relief		Unusual action	
Trust		Ability to act	
Confidence		Active experimentation	
Personal gratification		Ability to examine	
Creativity		Likelihood to recommend	Motivation
Excitement		Personal innovativeness	
Delightfulness		Frequent use	
Accomplishment		Motivation	
Anticipation		Reuse	
Appreciation		Commitment	
Emotions (in general)		Intention to use	
Curiosity		Word-of-mouth Intention	
Anxiety	Negative feelings	Hedonic motivation	
Boredom		Nausea	Simulator sickness
Frustration		Oculomotor	
State anxiety		Disorientation	
Anger		Simulator sickness	
Tension		Social interaction	Collaboration/Social interactions
Trait anxiety		Learner interaction	
Hopelessness		Communication place	
Distress		Perspective-taking	Uncategorized
Tiredness		Facilitators	
Shame		Fictional	
Competence	Self-efficacy	Dependability	
Skills balance		Learn friends	
Skills (advanced, mainframe, beginning)		Social influence/subjective norm	
Efficacy		Performance avoidance	
Self-efficacy		Performance approach	
Efficiency		Perceived behavioral control	
Perceived usefulness	Perceived knowledge gains	Organizational factor	
Knowledge improvement		Mastery-avoidance	
Competition		Mastery-approach	
Discovery		Hedonic quality-stimulation	
Narrative understanding		Future perception	
Comprehension		Goal orientation	
Performance expectancy		Experimental fidelity	
Perceived Learning		Regulatory uncertainty	
Self-assessment of performance		Reflective observation	
Deep learning		Perceived triability	
Clear goals	Feedback and content quality	Perceived observability	
Feedback		Perceived compatibility	
Guidance		Facilitating conditions	
Content		Effort expectancy	
Settings			
Menus			
Help			
Narratives			
Multimodality			
Interface quality			
Information quality			
Abstract conceptualization			

Appendix II. The study's version of the MLES scale

Factor	Item
Perceived quality of the virtual environment's graphics	The app was aesthetically pleasing
	I enjoyed the app's graphics
	The app was visually appealing
	I was satisfied with the app's graphics
	The graphics of the app were attractive
Perceived cognitive load	The presentation of too much information prevented the memorization of what was important*
	The effort to study the information that the application presented to me, was mentally tiring*
Perceived ease of use/control of the virtual environment	I used/controlled the app with ease
	I had full control over what I did
	When using the app, I had no problems doing whatever I wanted
Immersion/Presence	I forgot/ignored everything around me
	I lost the sense of where I am
	I lost the sense of time
	I felt like I was living in another place and time
Perceived feedback and content quality	Overall, the learning content was well presented
	The app gave me useful feedback regarding what I had to do
	The information provided by the app (e.g., objectives, help messages, images, texts, and audio) was clear and understandable
Perceived degree of interaction	I could interact a lot with the virtual world
	The virtual world responded well to my actions
	The interactions with the virtual objects were similar to the interactions with real objects
Motivation	I want to know more about what I saw in the app
	I enjoyed the content so much that I would like to know more about this topic
	The content had things that triggered my curiosity
	I feel motivated to keep using the app
	I was intrigued to see what was in the app
	I wanted to explore the app more
Perceived knowledge gains	I understood the basic ideas/issues presented to me within the app
	I learned through the app
	The content increased my knowledge and understanding of the subject presented by the app
Simulator sickness (To what degree you felt …)	Dizziness?*
	Your head being "heavy"?*
	Vertigo?*
	A general discomfort?*
	Nausea?*
	Headache?*
Positive feelings (To what degree you felt …)	Joy?
	Satisfaction?
	Enthusiasm?
	Excitement?

Notes. * = although these items are negatively worded, their reverse coding is not necessary as there are no positively worded items belonging to the same factor; all items were presented using a five-point Likert-type scale.

References

Abd Majid, F., & Mohd Shamsudin, N. (2019). Identifying factors affecting acceptance of virtual reality in classrooms based on technology acceptance model (TAM). *Asian Journal of University Education, 15*(2), 1–10. https://doi.org/10.24191/ajue.v15i2.7556

Abichandani, P., Mcintyre, W., Fligor, W., & Lobo, D. (2019). Solar energy education through a cloud-based desktop virtual reality system. *IEEE Access, 7*, 147081–147093. https://doi.org/10.1109/ACCESS.2019.2945700

Alawadhi, M., Alhumaid, K., Almarzooqi, S., Aljasmi, S., Aburayya, A., Salloum, S. A., et al. (2022). Factors affecting medical students' acceptance of the metaverse system in medical training in the United Arab Emirates. *South eastern European Journal of public health (SEEJPH)*. https://www.seejph.com/index.php/seejph/article/view/310.

Alfaisal, R., Hashim, H., & Azizan, U. H. (2022). Metaverse system adoption in education: A systematic literature review. *Journal of Computers in Education, 2022*, 1–45. https://doi.org/10.1007/s40692-022-00256-6

Almarzouqi, A., Aburayya, A., & Salloum, S. A. (2022). Prediction of user's intention to use metaverse system in medical education: A hybrid SEM-ML learning approach. *IEEE Access, 10*, 43421–43434. https://doi.org/10.1109/ACCESS.2022.3169285

Annetta, L., Klesath, M., & Meyer, J. (2009). Taking science online: Evaluating presence and immersion through a laboratory experience in a virtual learning environment for entomology students. *Journal of College Science Teaching, 39*(1), 27–34.

Armougum, A., Orriols, E., Gaston-Bellegarde, A., Joie-La Marle, C., & Piolino, P. (2019). Virtual reality: A new method to investigate cognitive load during navigation. *Journal of Environmental Psychology, 65*, Article 101338. https://doi.org/10.1016/j.jenvp.2019.101338

Asad, M. M., Naz, A., Churi, P., Guerrero, A. J. M., & Salameh, A. A. (2022). Mix method approach of measuring VR as a pedagogical tool to enhance experimental learning: Motivation from literature survey of previous study. *Educational Research International*. https://doi.org/10.1155/2022/8262304, 2022.

Atsikpasi, P., & Fokides, E. (2022). A scoping review of the educational uses of 6 DoF HMDs. *Virtual Reality, 26*(1), 205–222. https://doi.org/10.1007/s10055-021-00556-9

Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., & Rey, B. (2004). Immersion and emotion: Their impact on the sense of presence. *CyberPsychology and Behavior, 7*(6), 734–741. https://doi.org/10.1089/cpb.2004.7.734

Barry, D. M., Ogawa, N., Dharmawansa, A., Kanematsu, H., Fukumura, Y., Shirai, T., et al. (2015). Evaluation for students' learning manner using eye blinking system in Metaverse. *Procedia Computer Science, 60*, 1195–1204. https://doi.org/10.1016/j.procs.2015.08.181

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16*, 78–117. https://doi.org/10.1177/0049124187016001004

Bertrand, J., Bhargava, A., Madathil, K. C., Gramopadhye, A., & Babu, S. V. (2017). The effects of presentation method and simulation fidelity on psychomotor education in a bimanual metrology training simulation. *Proceedings of the 2017 IEEE Symposium on 3D User Interfaces, 3DUI, 2017*, 59–68. https://doi.org/10.1109/3DUI.2017.7893318

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.

Brown, S. A., & Venkatesh, V. (2005). A model of adoption of technology in the household: A baseline model test and extension incorporating household life cycle. *Management Information Systems Quarterly, 29*(3), 11. https://doi.org/10.2307/25148690

Butt, A. L., Kardong-Edgren, S., & Ellertson, A. (2018). Using game-based virtual reality with haptics for skill acquisition. *Clinical Simulation in Nursing, 16*, 25–32. https://doi.org/10.1016/j.ecns.2017.09.010

Caro, V., Carter, B., Dagli, S., Schissler, M., & Millunchick, J. (2018). Can virtual reality enhance learning: A case study in materials science. *Proceedings of the 2018 IEEE Frontiers in education conference (FIE), 1-4*. IEEE. https://doi.org/10.1109/fie.2018.8659267

Cecotti, H., Day-Scott, Z., Huisinga, L., & Gordo-Pelaez, L. (2020). Virtual reality for immersive learning in art history. *Proceedings of the 2020 6th international conference* *of the immersive learning research network (iLRN)*. IEEE. https://doi.org/10.23919/iLRN47897.2020.9155108

Chin, W. W. (1998). Commentary: Issues and opinion on structural equation modeling. *MIS quarterly*. vii-xvi.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research, 14*(2), 189–217. https://doi.org/10.1287/isre.14.2.189.16018

Claeskins, G., & Hkort, N. (2008). *Model selection and model averaging*. Cambridge University Press. https://doi.org/10.1017/CBO9780511790485

Coban, M., Bolat, Y. I., & Goksu, I. (2022). The potential of immersive virtual reality to enhance learning: A meta-analysis. *Educational Research Review, 36*, Article 100452. https://doi.org/10.1016/j.edurev.2022.100452

Dalgarno, B., Hedberg, J., & Harper, B. (2002). The contributions of 3D environments to conceptual understanding. *Proceedings of the 19th Annual Conference of the Australasian society for Computers in Learning in Tertiary Education*, 149–158 (UniTech).

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35*(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982

Diaz, J., Saldaña, C., & Avila, C. (2020). Virtual world as a resource for hybrid education. In *International Journal of emerging Technologies in learning (iJET)* (pp. 94–109). https://doi.org/10.3991/ijet.v15i22.14393, 15(15.

Duan, H., Li, J., Fan, S., Lin, Z., Wu, X., & Cai, W. (2021). Metaverse for social good: A university campus prototype. *Proceedings of the 29th ACM international Conference on multimedia, 153-161*. ACM. https://doi.org/10.1145/3474085.3479238

Erturk, E., & Reynolds, G. B. (2020). The expanding role of immersive media in education. *Proceedings of the international Conference on E-learning, 191-194*. International Association for Development of the Information Society.

Faeth, A., & Harding, C. (2014). Emergent effects in multimodal feedback from virtual buttons. *ACM Transactions on Computer-Human Interaction, 21*(1), 1–23. https://doi.org/10.1145/2535923

Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 269–314). IAP.

Fokides, E. (2017). A model for explaining primary school students' learning outcomes when they use multi-user virtual environments. *Journal of Computers in Education, 4*(3), 225–250. https://doi.org/10.1007/s40692-017-0080-y

Fokides, E. (2023). Development and testing of a scale for examining factors affecting the learning experience in the Metaverse. *Computers & Education: X Reality, 2023*(2), 100025. https://doi.org/10.1016/j.cexr.2023.100025.

Fokides, E., & Atsikpasi, P. (2018). Development of a model for explaining the learning outcomes when using 3D virtual environments in informal learning settings. *Education and Information Technologies, 25*(3), 2265–2287. https://doi.org/10.1007/s10639-018-9719-1

Fokides, E., & Kapetangiorgi, D.-M. (2022). The use of computers by Greek educators. Did the COVID-19 pandemic change anything? *Journal of Information Technology Education: Research, 21*, 217–244. https://doi.org/10.28945/4975

Fokides, E., & Zampouli, C. (2017). Content and language integrated learning in OpenSimulator project. Results of a pilot implementation in Greece. *Education and Information Technologies, 22*(4), 1479–1496. https://doi.org/10.1007/s10639-016-9503-z

Fussell, S. G., & Truong, D. (2022). Using virtual reality for dynamic learning: An extended technology acceptance model. *Virtual Reality, 26*, 249–267. https://doi.org/10.1007/s10055-021-00554-x

Gallagher, M., & Ferrè, E. R. (2018). Cybersickness: A multisensory integration perspective. *Multisensory Research, 31*(7), 645–674. https://doi.org/10.1163/22134808-20181293

Gaskin, J., & Lim, J. (2018). *Indirect Effects* [AMOS Plugin] http://statwiki.gaskination.com/index.php?title=Main_Page.

Go, S. Y., Jeong, H. G., Kim, J. I., & Sin, Y. T. (2021). 메타버스의 개념과 발전 방향 [Concept and developmental direction of metaverse]. *Korea Information Processing Society Review, 28*, 7–16.

Griffin, N. N., & Folmer, E. (2019). Out-of-body locomotion: Vectionless navigation with a continuous avatar representation. In *Proceedings of the 25th ACM symposium on virtual reality software and technology* (pp. 1–8). https://doi.org/10.1145/3359996.3364243. ACM.

Hair, J., & Alamer, A. (2022). Partial Least Squares Structural Equation Modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics, 1*(3), Article 100027. https://doi.org/10.1016/j.rmal.2022.100027

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis*. Cengage Learning, 8th ed.

Harrington, M. C. (2012). The virtual trillium trail and the empirical effects of freedom and fidelity on discovery-based learning. *Virtual Reality, 16*(2), 105–120. https://doi.org/10.1007/s10055-011-0189-7

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135.

Hsin, L. J., Chao, Y. P., Chuang, H. H., Kyo, T. B. J., Yang, C. C. H., Huang, C. G., et al. (2022). Mild simulator sickness can alter heart rate variability, mental workload, and learning outcomes in a 360° virtual reality application for medical education: A post hoc analysis of a randomized controlled trial. *Virtual Reality, 2022*, 1–17. https://doi.org/10.1007/s10055-022-00688-6

Huang, W., Roscoe, R. D., Craig, S. D., & Johnson-Glenberg, M. C. (2022). Extending the cognitive-affective theory of learning with media in virtual reality learning: A structural equation modeling approach. *Journal of Educational Computing Research, 60*(4), 807–842. https://doi.org/10.1177/07356331211053630

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huifen, G., Yan, M., Aizhu, W., Zhu, X., Lijiao, D., & Wenxiang, F. (2021). Who can benefit from immersive virtual reality in education? Effectiveness of IVR in teaching using meta-analysis. *Proceedings of the 2021 tenth international conference of educational innovation through technology (EITT)*. IEEE. https://doi.org/10.1109/EITT53287.2021.00069

Ibili, E., & Billinghurst, M. (2019). Assessing the relationship between cognitive load and the usability of a mobile augmented reality tutorial system: A study of gender effects. *International Journal of Assessment Tools in Education, 6*(3), 378–395. https://doi.org/10.21449/ijate.594749

Jeon, J., & Jung, S. K. (2021). *Exploring the educational applicability of Metaverse-based platforms. 한국정보교육학회: 학술대회논문집* [Korean Society for Information Education: Conference Proceedings] https://koreascience.kr/article/CFKO202130548299122.page.

Johnson, D. M. (2005). Introduction to and review of simulator sickness research. *Research report 1832*. U.S. Army Research Institute for the Behavioral and Social Sciences. https://doi.org/10.1037/e456932006-001

Kamstrupp, A. K. (2016). The wow-effect in science teacher education. *Cultural Studies of Science Education, 11*, 879–897. https://doi.org/10.1007/s11422-015-9684-6

Kim, Y. J., & Ahn, S. Y. (2021). Factors influencing nursing students' immersive virtual reality media technology-based learning. *Sensors, 21*(23), 8088. https://doi.org/10.3390/s21238088

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford Press.

Kobayashi, A., Aoki, R., Kitagawa, N., Kimura, T., Takashima, Y., & Yamada, T. (2016). Towards enhancing force-input interaction by visual-auditory feedback as an introduction of first use. *Proceedings of the Human-Computer Interaction. Interaction Platforms and Techniques: 18th International Conference, HCI International*. https://doi.org/10.1007/978-3-319-39516-6_17, 2016, Proceedings, Part II 18 (pp. 180–191). Springer International Publishing.

Kye, B., Han, N., Kim, E., Park, Y., & Jo, S. (2021). Educational applications of metaverse: Possibilities and limitations. *Journal of Educational Evaluation for Health Professions, 18*, 1–13. https://doi.org/10.3352/jeehp.2021.18.32

Lee, E. A. L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education, 55*(4), 1424–1442. https://doi.org/10.1016/j.compedu.2010.06.006

Lehikko, A. (2021). Measuring self-efficacy in immersive virtual learning environments: A systematic literature review. *Journal of Interactive Learning Research, 32*(2), 125–146.

Leppink, J., Paas, F., Van Gog, T., van Der Vleuten, C. P., & Van Merrienboer, J. J. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction, 30*, 32–42. https://doi.org/10.1016/j.learninstruc.2013.12.001

Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education, 95*, 174–187. https://doi.org/10.1016/j.compedu.2016.01.001

Lin, H., Wan, S., Gan, W., Chen, J., & Chao, H. C. (2022). *Metaverse in education: Vision, opportunities, and challenges*. https://doi.org/10.1109/BigData55660.2022.10021004. ***arXiv preprint*** arXiv:2211.14951.

Maas, M. J., & Hughes, J. M. (2020). Virtual, augmented and mixed reality in K-12 education: A review of the literature. *Technology, Pedagogy and Education, 29*(2), 231–249. https://doi.org/10.1080/1475939X.2020.1737210

Maereg, A. T., Nagar, A., Reid, D., & Secco, E. L. (2017). Wearable vibrotactile haptic device for stiffness discrimination during virtual interactions. *Frontiers in Robotics and AI, 4*, 1–9. https://doi.org/10.3389/frobt.2017.00042

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research & Development, 66*(5), 1141–1164. https://doi.org/10.1007/s11423-018-9581-2

Makransky, G., & Petersen, G. B. (2019). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. *Computers & Education, 134*, 15–30. https://doi.org/10.1016/j.compedu.2019.02.002

Makransky, G., & Petersen, G. B. (). The cognitive affective model of immersive learning (camil): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review, 33*, 937–958. https://doi.org/10.1007/s10648-020-09586-2, 2021.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2017). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction, 60*, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Malhotra, N. K., & Dash, S. (2011). *Marketing research an applied orientation*. Pearson Publishing.

Maraj, C. S., Badillo-Urquiola, K. A., Martinez, S. G., Stevens, J. A., & Maxwell, D. B. (2017). Exploring the impact of simulator sickness on the virtual world experience. In *Advances in human factors, business management, training and education* (pp. 635–643). Springer. https://doi.org/10.1007/978-3-319-42070-7_59.

Matsunaga, M. (2010). How to factor-analyze your data right: do's, don'ts, and how-to's. *International Journal of Psychological Research, 3*(1), 97–110. https://doi.org/10.21500/20112084.854

Moriuchi, E., Landers, V. M., Colton, D., & Hair, N. (2021). Engagement with chatbots versus augmented reality interactive technology in e-commerce. *Journal of Strategic Marketing, 29*(5), 375–389. https://doi.org/10.1080/0965254X.2020.1740766

Mystakidis, S. (2022). *Metaverse. Encyclopedia, 2*(1), 486–497. https://doi.org/10.3390/encyclopedia2010031. MDPI.

Novak, E., Daday, J., & McDaniel, K. (2018). Assessing intrinsic and extraneous cognitive complexity of e-textbook learning. *Interacting with Computers, 30*(2), 150–161. https://doi.org/10.1093/iwc/iwy001

Pagano, A., Palombini, A., Bozzelli, G., De Nino, M., Cerato, I., & Ricciardi, S. (2020). Arkaevision vr game: User experience research between real and virtual paestum. *Applied Sciences, 10*(9), 3182. https://doi.org/10.3390/app10093182

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology, 110*(6), 785–797. https://doi.org/10.1037/edu0000241

Parong, J., & Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning, 37*(1), 226–241. https://doi.org/10.1111/jcal.12482

Portman, M. E., Natapov, A., & Fisher-Gewirtzman, D. (2015). To go where no man has gone before: Virtual reality in architecture, landscape architecture and environmental planning. *Computers, Environment and Urban Systems, 54*, 376–384. https://doi.org/10.1016/j.compenvurbsys.2015.05.001

Potkonjak, V., Gardner, M., Callaghan, V., Mattila, P., Guetl, C., Petrović, V. M., et al. (2016). Virtual laboratories for education in science, technology, and engineering: A review. *Computers & Education, 95*, 309–327. https://doi.org/10.1016/j.compedu.2016.02.002

Queiroz, A. C. M., Nascimento, A. M., Tori, R., & da Silva Leme, M. I. (2018). Using HMD-based immersive virtual environments in primary/K-12 education. In *Proceedings of the international conference on immersive learning* (Vols. 160–173). Springer. https://doi.org/10.1007/978-3-319-93596-6_11

Ramírez-Correa, P., Rondán-Cataluña, F. J., Arenas-Gaitán, J., & Martín-Velicia, F. (2019). Analysing the acceptation of online games in mobile devices: An application of UTAUT2. *Journal of Retailing and Consumer Services, 50*, 85–93. https://doi.org/10.1016/j.jretconser.2019.04.018

Rasimah, C. M. Y., Nurazean, M., Salwani, M. D., Norziha, M. Z., & Roslina, I. (2015). A systematic literature review of factors influencing acceptance on Mixed Reality technology. *ARPN Journal of Engineering and Applied Sciences, 10*(23), 18239–18246.

Renninger, K. A., & Hidi, S. E. (2017). *The power of interest for motivation and engagement*. Routledge. https://doi.org/10.4324/9781315771045

Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective experiences in 360-degree videos. *Computers & Education, 128*, 256–268. https://doi.org/10.1016/j.compedu.2018.09.015

Selcuk, C. (2021). *A study on immersion and emotions' influence on impulse buying in virtual environments [Master's thesis*. Brock University. https://dr.library.brocku.ca/bitstream/handle/10464/15593/Brock_Selcuk_Cem_2022.pdf?sequence=1.

Shen, S., Xu, K., Sotiriadis, M., & Wang, Y. (2022). Exploring the factors influencing the adoption and usage of Augmented Reality and Virtual Reality applications in tourism education within the context of COVID-19 pandemic. *Journal of Hospitality, Leisure, Sports and Tourism Education, 30*, Article 100373. https://doi.org/10.1016/j.jhlste.2022.100373

Soper, D. S. (2016). In *Post-hoc statistical power calculator for multiple regression [Software]*. Retrieved from http://www.danielsoper.com/statcalc.

Stavropoulos, V., Rennie, J., Morcos, M., Gomez, R., & Griffiths, M. D. (2021). Understanding the relationship between the Proteus effect, immersion, and gender among World of Warcraft players: An empirical survey study. *Behaviour & Information Technology, 40*(8), 821–836. https://doi.org/10.1080/0144929X.2020.1729240

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research & Development, 68*(1), 1–16. https://doi.org/10.1007/s11423-019-09701-3

Tlili, A., Huang, R., Shehata, B., Liu, D., Zhao, J., Metwally, A. H. S., et al. (2022). Is metaverse in education a blessing or a curse: A combined content and bibliometric analysis. *Smart Learning Environments, 9*(1), 1–31. https://doi.org/10.1186/s40561-022-00205-x

Tzanavari, A., Charalambous-Darden, N., Herakleous, K., & Poullis, C. (2015). Effectiveness of an Immersive Virtual Environment (CAVE) for teaching pedestrian crossing to children with PDD-NOS. *Proceedings of the 2015 IEEE 15th international conference on advanced learning technologies*. IEEE. https://doi.org/10.1109/ICALT.2015.85

Villena-Taranilla, R., Cózar-Gutiérrez, R., González-Calero, J. A., & Diago, P. D. (2023). An extended technology acceptance model on immersive virtual reality use with primary school students. *Technology, Pedagogy and Education, 32*(3), 367–388. https://doi.org/10.1080/1475939X.2023.2196281

Villena-Taranilla, R., Tirado-Olivares, S., Cozar-Gutierrez, R., & González-Calero, J. A. (2022). Effects of virtual reality on learning outcomes in K-6 education: A meta-analysis. *Educational Research Review, 35*, Article 100434. https://doi.org/10.1016/j.edurev.2022.100434

Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2012). *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models*. Springer Science & Business Media.

de Winkel, K. N., Talsma, T. M. W., & Happee, R. (2022). A meta-analysis of simulator sickness as a function of simulator fidelity. *Experimental Brain Research, 240*, 3089–3105. https://doi.org/10.1007/s00221-022-06485-6

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The factor structure of the presence questionnaire. *Presence: Teleoperators and Virtual Environments, 14*(3), 298–312. https://doi.org/10.1162/105474605323384654

Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology, 51*(6), 1991–2005. https://doi.org/10.1111/bjet.13023