Highlights

An Educational Experience in Ancient Rome to Evaluate the Impact of Virtual Reality on Human Learning Processes

Paolo Boffi,Monica Clerici,Alberto Gallace,Pier Luca Lanzi

- Development and testing of a Virtual Reality experience in an Ancient Roman Domus
- Preparation of a questionnaire to asses learning ratio and human factors
- The application showed to be most pleasant than traditional learning tools
- The application showed to be at least as effective in education as traditional learning tools

An Educational Experience in Ancient Rome to Evaluate the Impact of Virtual Reality on Human Learning Processes★

Paolo Boffi[a,1], Monica Clerici[b], Alberto Gallace[b] and Pier Luca Lanzi[a,*]

[a]*Dipartimento di Elettronica, Informazione e Bioingegneria, Polytechnic University of Milan, Milan, 20133, Italy*
[b]*Mind and Behavioral Technological Center - MiBTeC, Dipartimento di Psicologia, University of Milano-Bicocca, Milan, 20133, Italy*

ARTICLE INFO

Keywords:
Virtual Reality
Media in Education
Adult Learning
Human-Computer Interaction

ABSTRACT

Immersive Virtual Reality technology has recently gained significant attention and is expanding its applications to various fields. It also has many advantages in education, as it allows to both simplify the explanation of complex topics through their visualization, and explore lost or unreachable environments. To evaluate the impact of immersive experiences on learning outcomes we developed an educational experience that lets users visit an ancient Roman Domus and provides information about daily life in Roman times. We designed a between-subjects data collection to investigate learning ratio, user experience, and cybersickness of participants through anonymous questionnaires. We collected 76 responses of participants (18-35 y.o.) divided into three conditions: a Immersive Virtual Reality experience, a slide-based lecture and a 2D desktop-based experience. Our results show that the virtual reality experience is considered more engaging and as effective as more traditional 2D and slide-based experiences in terms of learning.

1. Introduction

Virtual Reality (VR), defined as "the sum of hardware and software systems that seek to perfect an all-inclusive, sensory illusion of being present in another environment" (Biocca and Delaney, 1995), started widely spreading over the last few years, thanks to technological improvements and the reduction of VR devices' prices. The diffusion of VR devices has brought new enthusiasm to several research fields. One is the potential educational value this technology can bring to different users. VR applications can be adopted as an educative tool for several learning fields (Bell and Fogler, 1995; Gonzalez Izard et al., 2017). The interactivity and multisensorial immersion this technology provides help the users better understand the topic and learn by doing or living the specific situation. Moreover, VR is suitable for explaining complex and abstract concepts and visualizing remote environments (in space or time) (Mortara et al., 2014). One educational field in which VR can bring several benefits is cultural heritage, with the development of virtual museums and virtual tours (Paíno Ambrosio and Rodríguez Fidalgo, 2020).

Even if the educational value of VR is generally acknowledged in the literature, it remains to be seen the relationship between the level of immersion and the learning outcome (Wu et al., 2020; Pellas et al., 2021). In this paper, we present a VR experience that we developed to teach people about the structure of an Ancient Roman Domus and how Ancient Romans lived in it and collect data regarding learning, user experience, and cybersickness. We also designed two experiences we used as control conditions: a desktop application that lets users navigate the same VR experience through the first-person view used in many video games and a set of PowerPoint slides that could be used in typical lecture scenarios. We compared the educational effectiveness of the three experiences using human subjects and an online questionnaire. Our aim was to determine if the VR experience could be as effective as traditional learning tools (i.e., PowerPoint slides) in conveying basic historical knowledge and to investigate any potential relationship between the level of immersion and learning outcomes. Our results suggest that VR could be as effective as conventional learning strategies, while also greatly enhancing the user experience, engagement, and motivation. Additionally, participants who experienced the VR application reported a higher level of presence than the desktop version while maintaining comparable learning outcomes.

★This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
*Corresponding Author
paolo.boffi@polimi.it (P. Boffi); monica.clerici@unimib.it (M. Clerici); alberto.gallace1@unimib.it (A. Gallace); pierluca.lanzi@polimi.it (P.L. Lanzi)
ORCID(s): 0000-0003-4579-2606 (P. Boffi); 0000-0002-1933-7717 (P.L. Lanzi)

The paper is organized as follows. In Section 2, we briefly overview the current state of the art of educational VR, focusing then on its use in the cultural heritage field and introducing the concept of Sense of Presence. In Section 3, we present our experience, starting from the requirements we identified, its overall structure, its main components, and the control conditions. In Section 4, we discuss the questionnaire we designed to evaluate our experience, its main modules, how we carried out the experimentation, and the feedback we received. In Section 5, we discuss the results of our experiments. In Section 6, we draw some conclusions and outline future research directions.

2. Related Works and Research Hypotheses

This section provides an overview of the theoretical approaches and relevant applications of VR in education that inform our study. We first introduce the main theories we employed to design our educational experiment and then highlight significant VR applications in education. For further in-depth information on this topic, we refer interested readers to recent reviews (Liu et al., 2017; Greenwald et al., 2017; Bekele et al., 2018; Kamińska et al., 2019; Radianti et al., 2020; Pellas et al., 2021; Villena-Taranilla et al., 2022). Additionally, we discuss the concept of Sense of Presence and its relevance to our work. Lastly, we outline the research questions and hypotheses guiding our study.

2.1. Learning Paradigms

There are several theoretical approaches we can follow in developing educational VR applications. We cite two of them: the *Constructivist Theory* (Liu et al., 2017) and the *Embodied Cognition Theory* (Shapiro, 2007).

Following the first one, we construct knowledge learning by doing and interacting with the environment through our senses. We build a world model based on the sensory inputs we receive due to the outcomes of our actions; we then update it by acting on the world and verifying the outcome of those actions through our senses. Experience plays a crucial role in learning. To explain complex or abstract concepts, the instructors often use metaphors and analogies, bringing those concepts to commonly observable reality (Christou, 2010).

The Embodied Cognition Theory is similar to but different from the previous one. It bases on two main concepts: cognition is not abstract and amodal and is not just about thinking but also involves perceiving and acting in the real world (Ionescu and Vasc, 2014). The critical point of this theory is the role of the body: learning does not involve only the mind, but there is a substantial contribution from our physical being. For instance, some studies demonstrate that hand movements play a role in solving or learning to solve different problems (Ionescu and Vasc, 2014).

Constructivism points attention to the sensory feedback we receive, while Embodied Cognition Theory focuses more on the body. However, they both agree that a practical interaction with the world can improve knowledge. For this reason, VR can strongly support learning activities, thanks to the interactivity it can bring. In this work, we based mainly on the Constructivist Theory: we strongly relied on the sensory feedback the user received from our experience (namely, the visual feedback of the reconstructed environment and the auditory feedback from the ambient sounds and the lecture contents).

2.2. VR in Education: What is Known so far

Several educational VR applications have been developed in the last few years. They can provide learning experiences in several fields, such as physics (Kumar et al., 2021; Dede et al., 1996), history and cultural heritage (Villena Taranilla et al., 2019; Liu et al., 2021; De Paolis et al., 2009; Häkkilä et al., 2019), math (Roussou, 2009), biology (Ou et al., 2021), and language learning (Pinto et al., 2019). They can also be used to increase people's awareness regarding specific topics, like climate changes (Fauville et al., 2020) or the space debris threat (Colombo et al., 2020).

Various studies have acknowledged the contribution of VR technology in increasing student engagement, motivation, and personal satisfaction compared to traditional learning tools (Allcoat and von Mühlenen, 2018; Mahmoud et al., 2020; Ip et al., 2019). However, the literature provides mixed feedback when comparing educational VR with other learning techniques. In one study, the authors of (Akbulut et al., 2018) developed a VR system to teach sorting algorithms to Computer Science bachelor students. They compared the results obtained by the students who used this tool with those who employed common PowerPoint presentations, noticing better results in the former group. On the other hand, Parong and Mayer (Parong and Mayer, 2018) found the opposite: they prepared a biology lecture in both VR and PowerPoint, with a higher learning outcome from the latter learning strategy. Even so, the students perceived the VR version as more engaging and motivating. Makransky and colleagues achieved similar results (Makransky et al., 2019): they investigated the effects of increased immersion in a virtual laboratory by comparing its VR and the 2D-desktop versions. Results showed a lower learning outcome in the VR experience but with a higher sense of presence. Generally speaking, there is still a lack of knowledge regarding the effectiveness of VR as an educational instrument with respect to other learning strategies (Wu et al., 2020), and further studies are needed in this direction. However, it is acknowledged its motivational and engaging value. For this reason, an educational VR experience can be considered effective in transmitting knowledge if it is at least as effective as traditional learning strategies (Pellas et al., 2021). We hypothesized that our application could reach this achievement, resulting as effective as other tools, particularly a 2D desktop-based experience and a PowerPoint presentation (H.1).

In our study, we have focused on the educational use of VR to teach historical content, specifically the reconstruction of an ancient Roman Domus. Our decision to concentrate on this topic was driven by the scarcity of works exploring the use of VR for educational purposes in teaching history (Villena-Taranilla et al., 2022; Taranilla et al., 2022). While a few studies have investigated VR's potential in cultural heritage reconstruction (Deggim et al., 2017; Richards-Rissetto et al., 2013), there has been relatively little research on using VR for historical education. Moreover, the majority of the research in this field is focused on STEM subjects, leaving unknown the possible contribution of VR in historical teaching (Kavanagh et al., 2017). While there have been some works that tried to exploit this potential by, for instance, explaining concepts related to archaeology and heritage (Fabola and Miller, 2016) or teaching the Spanish life in the 16th century (Perez-Valle and Sagasti, 2012), they offered no proofs of this hypothesis through validated experimental design and data collections (Taranilla et al., 2022). Some VR experiences allow users to explore ancient Roman buildings (Fleury and Madeleine, 2012; Flyover Zone, 2018), while others are focused on ancient Roman Domus reconstruction (Gregl, 2020). However, these experiences have primarily been designed to offer visually stunning experiences rather than to assess their educational efficacy compared to traditional teaching methods rigorously. In opposition to them, we decided to focus our efforts on exploring the contribution that immersive VR can provide in explaining such a topic. We hypothesized that being immersed inside an ancient Roman Domus could elicit the users' interest, leading to a higher level of motivation with respect to other tools, such as PowerPoint presentations (H.2). Moreover, thanks to the motivational features of our application and the careful attention we paid to the design phase (refer to Section 3 for more details), we anticipate that the increased immersion of our VR experience will have no adverse effects on learning outcomes (H.4), in contrast to previously mentioned works. The authors of (Kyrlitsias et al., 2020) developed an immersive VR experience in the archaeological site of Choirokoitia. They compared the simulation with a desktop experience to collect data regarding Presence and learning ratio. However, their work focused on a specific location (the archaeological site of Choirokoitia) rather than a general topic like our study. As educational VR is commonly utilized for simple, specific applications, and not for general explanations (Pellas et al., 2021), we decided to focus our efforts in this direction. Moreover, they tested the user experience and sense of presence on very few items, and they did not include an evaluation of the level of cybersickness in their final questionnaire.

This last point, in particular, is very important: one of the main limitations of VR is the occurrence of simulator sickness, which manifests itself with symptoms like eye fatigue, disorientation, and nausea (LaViola, 2000). These uncomfortable feelings can significantly affect the user experience during the VR simulation (Saredakis et al., 2020; Chang et al., 2020). Despite the improvement in the quality of VR devices has reduced the occurrence of these unwanted effects (Lee et al., 2017; Kourtesis et al., 2019), users still report experiencing them (Dużmańska et al., 2018; Guna et al., 2019). Hence, it is important to consider the issue of cybersickness when designing and developing VR applications. Our work took into consideration certain precautions to minimize the occurrence of cybersickness (following mainly the findings of Saredakis et al., 2020 and Chang et al., 2020; refer to Section 3.5 for more details), and we hypothesized that users would only experience at most mild symptoms during the execution of our experience (H.3).

2.3. Immersion and Sense of Presence: Definitions and Measurements

As one of the aims of this work was to investigate whether the increased immersion of a VR experience can affect its learning outcome, we first need to clarify what we refer to by this term and the relationship between immersion and Sense of Presence. In this subsection, we will provide two main definitions of these concepts, one by Witmer and Singer (WS) and the other by Slater. Furthermore, we will present some of the most common methods used to measure the Sense of Presence in VR experiences.

For WS, Presence is "a psychological state of 'being there' mediated by an environment that engages our senses, captures our attention, and fosters our active involvement" (Witmer et al. (2005), p. 298). Hence, a Virtual Environment (VE) that provides a coherent set of sensory stimuli is linked to the user's Presence. WS associate Presence with the level of immersion, which they describe as a psychological state in which the user perceives themselves as fully included and interacting within the VE (Witmer and Singer, 1998). According to WS, a higher level of immersion leads to a higher level of Presence.

On the other hand, Slater defines Presence as encompassing three aspects: the sense of "being there" within the VE, the dominance of the VE as the user's environment, and the perception of the VE as having visited a place rather than just viewing computer-generated images (Slater, 1999). Slater also explains that Immersion refers to the objective technological properties of a system, while Presence is the subjective human experience of these properties (Slater, 2003). He notes that, given the same immersive system, different people may have varying levels of Presence.

The main difference between these two perspectives lies in their definitions of Immersion and its relationship to Presence. According to WS, the two concepts are closely related and cannot be easily differentiated. Slater, however, clearly separates the two and defines them as distinct yet related concepts.

The variety of definitions of Sense of Presence has led to several methods of measurement, which can be divided into three categories (Insko, 2003):

- *Subjective Measures*: self-reported measurements obtained through post-immersion questionnaires, including:
  - *Witmer & Singer Presence Questionnaire* (Witmer and Singer, 1998; Witmer et al., 2005)
  - *Slater-Usoh-Steed (SUS) Questionnaire* (Usoh et al., 2000)
  - *ITC-Sense of Presence Inventory* (Lessiter et al., 2001)
- *Behavioural Measures*: measurements of user behavior through techniques such as direct observation and self-monitoring. The greater the participant's feeling of presence in the VE, the more their responses will resemble those in the real world.
- *Physiological Measures*: objective body reactions, including heart rate, skin conductance, or temperature.

(Grassini and Laumann, 2020) analysed several literature works to compare the different techniques used to measure the Sense of Presence. Even if there seems to be no standard method for measuring Presence, the most used technique involves post-immersion questionnaires. Among them, the most popular is the WS one (Witmer and Singer, 1998; Witmer et al., 2005). We also chose to use the WS questionnaire, as it had more technical factors, useful for our multidisciplinary team.

2.4. Research Hypotheses

During our literature analysis, we stated some hypotheses regarding the possible results of our work. In this subsection, we list all of them:

H.1 In terms of learning outcomes, we expect our VR application to be at least as effective as the two control conditions, if not more effective, due to its immersive nature and the users being guided along a specific path; in particular, this last attention should limit potential distractions. Additionally, we anticipate that our VR application will be particularly effective in conveying spatial information about the Domus, thanks to the high level of immersion provided by the VR environment. These advantages should limit potential distractions and enhance learning outcomes.

H.2 We anticipate that users will be more engaged and motivated when using our VR application compared to the control conditions, as our application has been designed to be highly engaging. Additionally, we expect users to report a high level of satisfaction and acceptance of our application when compared to the two control conditions.

H.3 We expect the occurrence of Cybersickness to be infrequent or absent during the VR experience, and any symptoms experienced to be mild at most, due to the careful consideration given to its prevention during the application's design.

H.4 We anticipate a positive level of presence in the VR application, evidenced by higher scores compared to the control condition. Furthermore, we do not expect the sense of presence to impact the learning outcome negatively.

3. The Domus Romana VR Application

Our goal was to assess the effectiveness of VR as an educational tool in conveying basic historical concepts, such as the structure of an ancient Roman Domus, and to check the potential influence of the level of immersion on the learning outcome. For this reason, we developed an educational VR experience and compared it with two control conditions, a 2D desktop-based application and a slide-based lecture, in terms of learning outcome, user experience, and cybersickness.

Our research team, involving both VR engineers, cognitive neuroscientists, and psychologists, initially identified a list of requirements for this project. The first and foremost requirement was that the application should have two counterparts to compare the results the users achieve with the VR application with the ones of the control experiences. The VR experience and the control ones must then completely overlap in terms of informational content (but not in terms of interaction) so that the obtained results could be comparable. Still, there should be an essential distinction between the slide-based control condition and the "digital" ones: inside the applications (both the VR and the desktop version), there must be only three-dimensional elements, while only two-dimensional pictures should be displayed inside the standard lectures. The applications (both the VR and the desktop one) must implement two ways of use: one lecture-driven (called Lecture Mode), in which the user could follow only a pre-determined path with fixed explanation points, and another fully exploratory (called Exploration Mode), with the user who can freely explore the environment with no limitations. Finally, the user should be completely autonomous while using the applications; therefore, a tutorial has been developed for both the Lecture and the Exploration Mode.

Initially, the applications were structured with the main constraint that the user had to pass through all the Lecture Mode before enabling the Exploration Mode. Later on, we decided to split the two ways of use entirely, resulting in two different applications: this allowed us to avoid any problem regarding the Exploration Mode because we found no way to obtain an equivalent experience through a slide-based lecture. Moreover, we adapted the testing phase too: the central part of the experiment was the comparison between only the Lecture Mode (both in VR and with the desktop application) and the slides-based lecture. In addition, we wanted to test the Exploration Mode with a small number of users to collect some preliminary data (even if this was not our primary aim); however, we chose to focus only on the Lecture Mode testing and the comparison between them and the standard lecture, leaving other tests to future developments.

All the historical contents of the experience have been taken from Marco Vitruvio Pollione, 15 B.C.; Angela, 2007; Paoli, 1982; Flocchini et al., 2017, and reworked with the support of an expert in this field: a former history teacher that helped us also in designing the lecture structure. We kept the whole experience, as well as the questionnaire, in our language in order to reduce the cognitive load associated with the need to process a foreign language (differently from other works in literature, like Makransky et al., 2019) and not to add any linguistic bias to the results.

3.1. The Virtual Reality Experience

The Virtual Reality Experience (VRE) is designed as a guided tour of a 3D reconstruction of a Domus, with a fixed path that users must follow. To proceed, they must complete the current lecture section, and the experience is sequential, meaning users cannot skip any part to complete the journey. We expect these two features will help achieve our goal of preventing the potentially higher level of Sense of Presence from negatively impacting the learning outcome (H.4). The user is spawned before an introductory arc, with a starting menu showing that he/she is starting the guided tour. After a brief tutorial explaining how to move inside the environment and trigger the lecture audio, the experience begins. Movement is performed through teleportation, a common way to avoid motion sickness (Clifton and Palmisano, 2020): by pressing the trigger button of the controller, a teleporting ray will be shown, red if it is pointing to somewhere the user cannot teleport, green otherwise. Furthermore, to make the possibility of teleporting more intuitive, the controller will give haptic feedback to the user through a light vibration once the ray points to a valid area.

There were eight lecture points inside the application (see Figure 2 for their placement inside the VE), each of them treating a different aspect of the Domus:

1. *Introduction*: a preliminary description of the Domus structure and spaces
2. *Atrium*: first main space of the Domus, with access to Cubicula and Tablinum
3. *Cubicula*: bedrooms of the Domus' inhabitants
4. *Impluvium*: pool in the middle of the Atrium that collects the rainwater

(a) Atrium

(b) Peristylium

Figure 1: The two main areas of the Domus: above there is the Atrium, seen from the entrance of the Domus itself, while below there is the Peristylium. The signboards that indicate the lecture points are visible

5. *Cartibulum*: table inside the Atrium, for fast businesses and meetings
6. *Tablinum*: Domus' office, for important businesses and meetings
7. *Peristylium*: second main space of the Domus, garden surrounded by columns
8. *Triclinium*: the banquet hall

Except for the Introduction, every lecture point is shown through green particles starting from the floor and a signboard, with the room's name that the audio will explain and one or two Latin sentences linked to it. The users can teleport themselves on the particles, making them disappear; then, they will be able to make the audio play using the primary button of the controller (A on the right and X on the left). During the execution of the lecture part, the users will not be allowed to move to remain focused on the lecture contents. Moreover, there is no way to stop the audio to avoid the loss of any information offered by the experience. After the end of the audio, the following lecture point will be shown. When the users teleport to the next lecture point, the previous signboard will disappear to keep the environment as immersive as possible. The Introduction had no lecture signboard; however, it had an additional feature: a three-dimensional map of the Domus, completely inspectable by the users from the height. Additionally,

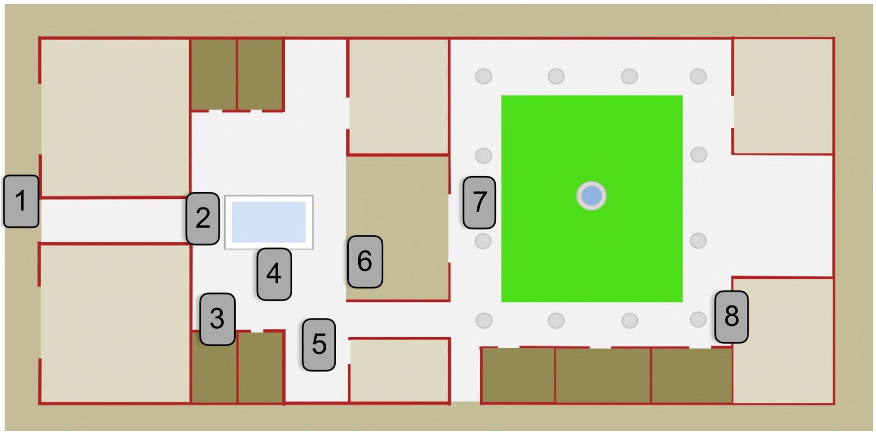


Figure 2: Map of the Domus, with the lecture points' placements

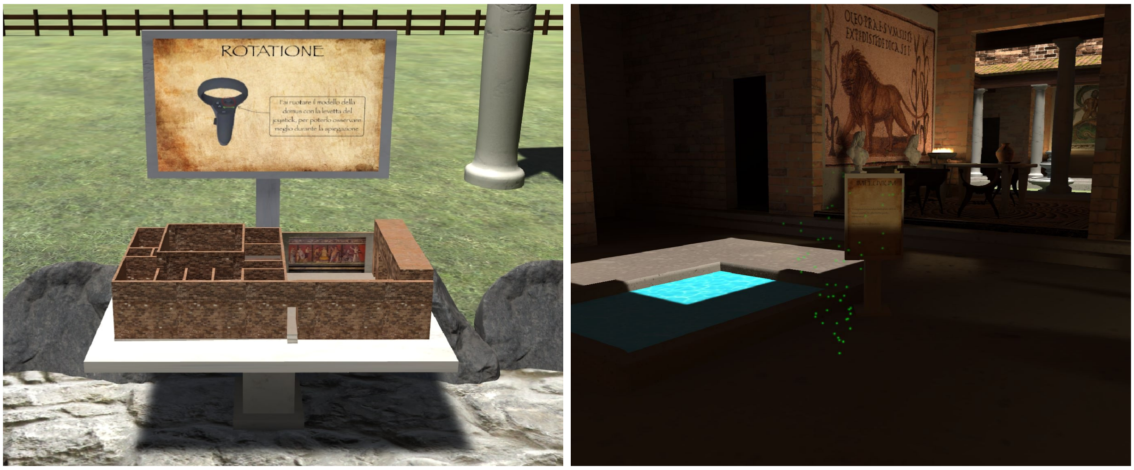


Figure 3: Two examples of lecture points: the Introduction (on the left) and the Impluvium (on the right)

users can rotate the map using the controller sticks, allowing for a more comprehensive examination of the internal environment during the explanatory phase.

3.2. The First Person Experience

The First Person Experience (FPE), developed for desktop PCs, reflects the VRE entirely. The main differences are how the user observes the virtual world (not through an Head-Mounted Display (HMD) but a 2D screen) and moves inside it (through mouse and keyboard). The tutorial is slightly changed (see Figure 5) to adequately explain the new input system.

3.3. The Multi-media Presentation Experience

The Multi-media Presentation Experience (MPE) is built using Microsoft PowerPoint. Its slides contain images taken from the VRE, brief textual descriptions, and audio tracks. The structure of the slides is made to reflect the design of the digital lectures thoroughly: the contents and the explanation audios are the same. Moreover, we kept the number and contents of written elements as limited as possible to not distract the users during the lecture audios execution and not add more information concerning the other two experiences.

Each lecture point presented in the digital experiences corresponds precisely to one slide of the standard lecture, except for the Introduction, Atrium, and Cartibulum ones. The first one was split into four slides (shown in Figure 4),

Figure 4: Introduction slides of the MPE, corresponding to the single introduction point of the VRE and FPE

each showing different aspects of the general structure of the Domus. In contrast, the other two were joined together in a single slide because of the brief content of the explanation.

3.4. Tutorial

To guarantee the complete autonomy of the users during the experience, we prepared two tutorials for both the VRE and the FPE.

To keep the environment as immersive as possible, the tutorial stages have been represented as signboards (similar to the ones for the lectures), with a heading word in Latin and images and sentences to describe how to act inside the environment. We placed the tutorial signboards on the road that connects the introductory arc to the Domus to let the users become familiar with the application controls while proceeding toward the central area of the experience. There were three tutorial signboards:

- *Motus*: explains how to move inside the environment and the meaning of the color of the teleport rays
- *Lectio*: explains where to teleport to hear the lectures and how to start the execution of the audio
- *Rotatione*: explains how to rotate the 3D map of the Domus

3.5. Avoiding Cybersickness: The Adopted Strategies

As discussed in Section 2.2, preventing Cybersickness is essential for providing a positive VR experience. In our application, this was a key consideration, given that studies have suggested a correlation between the realism of the simulation and the severity of sickness experienced by users (Chang et al., 2020). Realistic environments can create higher immersion, leading to an expectation of vestibular feedback closer to real-world ones. As we aimed to have a

Figure 5: Comparison between one of the tutorial signboards in the VRE (on the left) and in the FPE (on the right)

representation of the environment near to reality as much as possible, we were mindful of this issue throughout the development of our VR application and incorporated well-known strategies to minimize this unwanted effect. Firstly, we limited movement to teleportation, a commonly used technique to prevent mismatches between the visual feedback of the VR experience and the steady vestibular feedback of the user's body (Clifton and Palmisano, 2020; Chang et al., 2020).

Another important consideration during the development of our VR application was the need for frame-rate stability. We aimed to create a realistic and stable simulation without frame drops (Lee et al., 2017). One of the main challenges in achieving this goal was the lights: the level of fidelity of a 3D reconstruction depends not only on the models that compose the environment but also on the lights (Yu et al., 2012). Moreover, lightning plays a crucial role in building the atmosphere of the virtual environment (Edensor, 2015) and may be a critical factor in increasing the users' immersion (Wang, 2022). However, rendering realistic lighting can be computationally intensive, impacting performance. To address this issue, we decided to pre-compute the majority of the lights and bake their respective lightmaps at compile-time. This allowed us to achieve higher levels of realism without sacrificing performance.

Finally, the potential impact of extended VR exposure on users is worth considering. Research has shown that experiences lasting more than 10 minutes can increase cybersickness and negative outcomes (Saredakis et al., 2020; Chang et al., 2020). Therefore, we deliberately decided to limit the overall duration of our VR experience to approximately 10 minutes to minimize the risk of negative effects on users. By taking these precautions, we aim to reduce as much as possible the level of cybersickness and achieve our goal as stated in hypothesis H.3.

4. Questionnaire, Experimental Design, and Results

4.1. Questionnaire Outline

To perform the experimental phase of this work, we needed a questionnaire to collect data and evaluate users' learning ratio, user experience, and cybersickness level. For this reason, we created a questionnaire, relying on validated and customized sets of items during the development. To test *Personal Innovativeness*, *Technology Acceptance*, *User Experience (UX)*, *Presence* and *Cybersickness*, we relied upon validated measures provided by the literature (see Table A.3) while regarding familiarity with the technology, learning ratio and the overall experience we used prototypical items. To select and structure the variables of our questionnaire, we primarily followed two similar case studies: (Kyrlitsias et al., 2020) tested *Presence*, *UX* and learning performances on a VR reproduction of a hardly accessible archaeological site, while (Sagnier et al., 2020) tested *Technology Acceptance*, *UX*, *Cybersickness*, *Presence* and *Personal Innovativeness* of a VE for aeronautical training.

We came up with eight main blocks (outlined in Figure 6). Each of them has been taken from a different source, and, in some cases, we selected only a portion of the total number of items (see Table A.3 for more details). The main

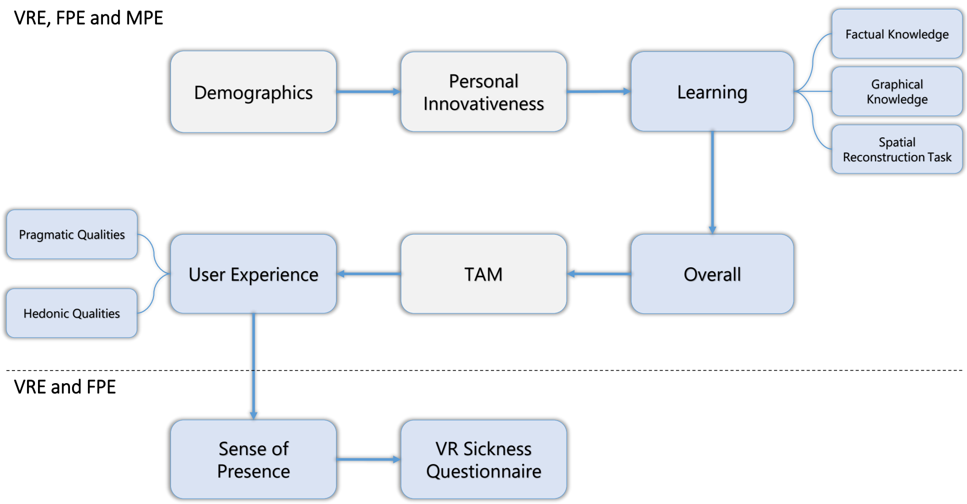


Figure 6: Questionnaire outline; some blocks of the questionnaire have been used only for the VRE and the FPE

structure of the questionnaire was the same for both the VRE and the FPE. However, for the MPE, we proposed a modified version without the *Presence* and *Cybersickness* blocks; moreover, some of the questions of the other blocks have been adapted to fit the new condition.

The questionnaire started with introductory questions regarding Demographics, Personal Innovativeness, and Technology Familiarity. It followed a series of **Learning** questions, aimed at testing the overall learning ratio of the users, divided in:

- *Factual Knowledge*: multiple-choice and true/false questions about the topics introduced during the experience
- *Graphical Knowledge*: multiple-choice questions related to a focused observation of the environment
- *Spatial Reconstruction* tasks:
  - *Room Name*: assign the correct room name to each area of the experience
  - *Object Location*: locate a pool of selected objects in their proper location inside the VE

The Overall Opinion block was a set of multiple-choice, open, and semantic differential items aimed at measuring the overall experience from a qualitative point of view. After that, we had the **Technology Acceptance Model (TAM)** block, a set of items to measure the level of acceptance of VR as a learning tool and divided into *Perceived Ease of Use*, *Perceived Usefulness*, *Attitude Toward Use*, and *Behavioural Intention*. However, we decided to exclude the analyses coming from the TAM section of our questionnaire. This is due to two reasons. Firstly, it exceeds the intended scope of this paper. Secondly, the limited sample size falls short of meeting the requirements for thoroughly analysing those variables.

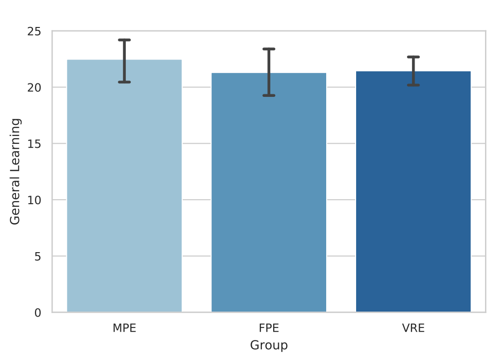
Another major block conceived **UX**, further divided into two groups of semantic differentials: *Pragmatic Qualities (PQ)* and *Hedonic Qualities (HQ)*.

To measure the presence the user experienced, we added the **Presence** block, divided into *Involvement*, *Sensory Fidelity*, *Adaptation/Immersion*, and *Interface Quality*.

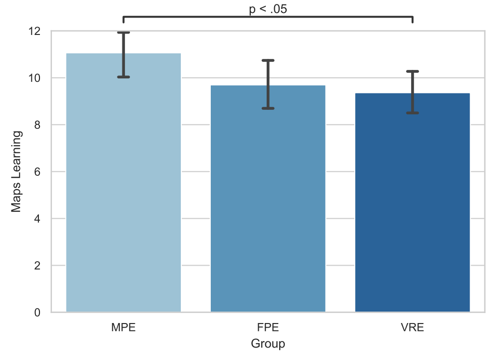
Finally, we dedicated the last block of the questionnaire to **Cybersickness**, with a set of symptoms that participants were required to notify. We decided to test two main groups of symptoms (*Oculomotor* and *Disorientation*), with an additional external item for nausea.

4.2. Materials and Methods

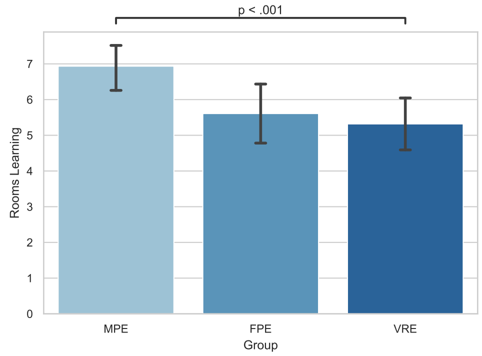
To test our application's efficiency, we went for a three-conditions between-subject design. Participants have been randomly assigned to one of the three previously-described conditions: VRE, FPE, or MPE. The participants should



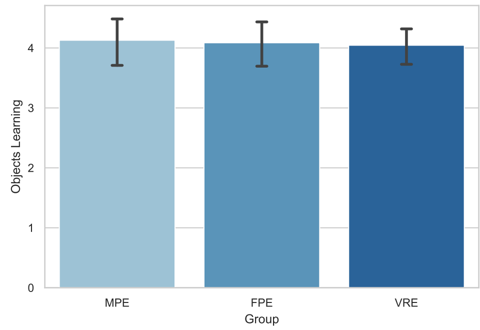
(a) *General Learning* results



(b) *Maps Learning* results



(c) *Rooms Learning* results



(d) *Objects Learning results*

Figure 7: Learning results for the participants of the three conditions

first take part the respective experience and thencomplete our survey, realized on Qualtrics[1]. The complete duration of the experiment was between 30 and 45 minutes. Initially, we intended to conduct the entire experiment and collect data in our laboratory. However, due to COVID-19 pandemic regulations, we needed to find an alternative solution that would allow us to gather a sufficient number of participants without putting them in danger. As a result, we opted for a hybrid approach, where participants assigned to the VRE condition were asked to visit our laboratory. In contrast, the FPE application and MPE presentation were distributed remotely to participants assigned to these conditions.

We implemented the VRE and the FPE on Unity3D 2019, using the Unity XR Interaction Toolkit v. 0.1. The application has been designed for VR HMDs connected to a computer, and in particular for the Oculus Rift[2] and Rift S[3] headsets. Participants used these headsets (and relative controllers) connected to a PC (dual NVIDIA® GeForce® RTX 2080 Ti, Intel® CoreTM i X-series Processors) with a Full HD monitor (24.5′, 144Hz refresh rate).

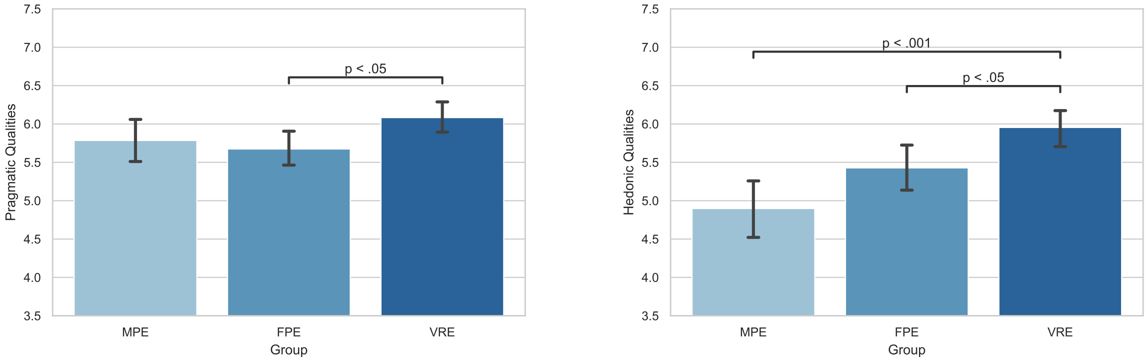
4.3. Results

The experimentation involved 76 participants, 36 males and 40 females, aged between 18 and 35 (see Table A.2 for more details). In the following, we report the main results; all the numerical data are listed in Appendix A.

Learning Ratio

To understand the efficacy of the three experiences in transmitting basic historical knowledge, we computed the *Learning* variable as the sum of all the correct answers given by each subject in the learning section of our questionnaire. According to our sample, no differences occurred in the three conditions ($F(2, 47) = 0.453$, $p = 0.638$), reflecting what we expected. To make a further analysis, we computed in the same way also *General Learning* (Factual and Spatial

[1]https://www.qualtrics.com/it/
[2]https://www.oculus.com/rift
[3]https://www.oculus.com/rift-s



(a) *Pragmatic Qualities* results

(b) *Hedonic Qualities* results

Figure 8: User Experience results for the participants of the three conditions

Knowledge) and *Spatial Task* (map reconstruction tasks, called *Maps Learning*). No difference were found also for *General Learning*; however, we had a significant difference in the *Spatial Task* ($F(2, 47) = 3.457$, $p = 0.04$) between the VRE and the MPE ($M = 1.701$, $t(51) = 2.564$, $p = 0.035$). As this went against our preliminary hypothesis H.1, we considered two sub-components: *Room Name* and *Object Location*. In particular, the former one highlighted a significant difference ($F(2, 46) = 5.555$, $p = 0.007$) between the two mentioned groups, with a better performance recorded in the MPE ($M = 1.617$, $t(48) = 3.156$, $p = 0.008$) (check Figure 7 for more details).

User Experience

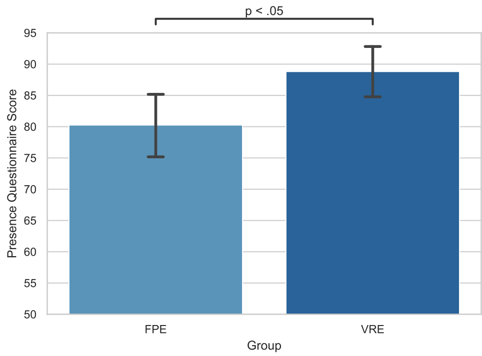
We measured UX through an adaptation by Sagnier et al., 2020 of the *AttrakDiff2* semantic differential, composed of couples of adjectives to be evaluated through a 7-point Likert scale. Moreover, we grouped the items into two main components: *PQ* and *HQ*. We found significant differences regarding the PQ ($F(2, 49) = 3.68$, $p = 0.033$). On average, participants of the VRE condition evaluated more positively the Pragmatic characteristics of the experience in respect of the FPE ($M = 0.409$, $t(42) = 2.62$, $p = 0.0232$). In contrast, no difference has been recorded with the MPE. Regarding the HQ, we found significant differences too ($F(2, 48) = 10.34$, $p < .001$): on average, participants evaluated more positively the VRE condition than the other two (FPE: $M = 0.527$, $t(43) = 2.62$, $p = 0.032$; MPE: $M = 1.057$, $t(50) = 4.47$, $p < .001$) (check Figure 8 for more details).

Sense of Presence

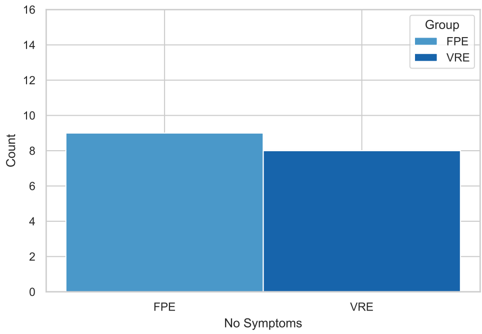
Through a shortened version of the Presence Questionnaire (Witmer et al., 2005), we evaluated the overall *Presence Score (PS)* and four related constructs (*Involvement*, *Sensory Fidelity*, *Immersion*, *Interface Quality*), on a 7-point Likert scale. We found a significant difference in PS between the two conditions ($t(43) = 2.57$, $p = 0.014$); furthermore, the VRE reached higher scores ($M = 88.8$, $SD = 10.1$) with respect to the FPE ($M = 80.3$, $SD = 12.1$) (check Figure 9a for more details). Regarding the other components, we found significant differences between the VRE and the FPE only concerning Involvement ($t(37) = 4.06$, $p < .001$) and Immersion ($t(42) = 2.05$, $p = 0.047$). In contrast, the other two did not differ in the two conditions.

Cybersickness

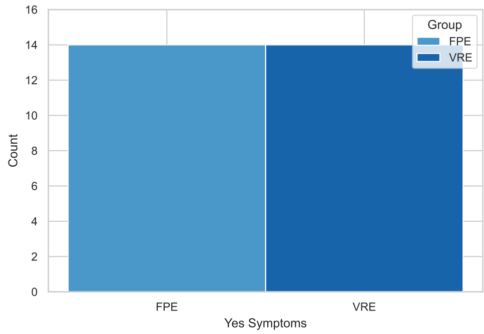
Cybersickness scores were registered through a 4-point Likert scale, with the following labels: 0 (no symptoms), 1 (light symptoms), 2 (moderate symptoms), and 3 (severe symptoms). We proposed a list of symptoms to be evaluated, following the indications of Kim et al., 2018 (more details on symptoms and how to compute the values in Table A.1). We computed then a main variable called *Cybersickness* and two sub-components, *Oculomotor* and *Disorientation*. During our analysis, we found out that a relevant portion of our participants scored 0: the mode of Cybersickness and its relative factors was 0; therefore, our distribution was strongly unbalanced. For this reason, we decided to transform Cybersickness into a categorical variable: we assigned the label *no symptoms* to participants who scored 0 and *yes symptoms* to those who scored greater than 0. After that, we computed a Chi-Square Test of Independence



(a) Presence Questionnaire Score



(b) *No Symptoms* answers



(c) *Yes Symptoms* answers

Figure 9: Presence Questionnaire Score (above) and Cybersickness results (below) for the participants of the VRE and FPE conditions

($\chi^2 = 0.036$, $p = 0.848$), which confirmed that there seem to be no relationships between the experimental condition and Cybersickness (Figures 9b and 9c).

5. Discussion and Limitations

5.1. Discussion on the Obtained Results

Our research focused on developing and assessing Human Factors and educative implications involved in a VR application reproducing an ancient Roman Domus. Our results show no differences in the three groups regarding the *General Learning* variable (a finding consistent with H.1), suggesting that VRE can effectively transmit basic historical knowledge like the familiar FPE and MPE tools. Virtual and desktop experiences show a learning performance comparable to traditional PowerPoint presentations (probably the most widely used educational tool), and this can be considered a relevant outcome (Pellas et al., 2021). Surprisingly, in contrast with our initial hypothesis, MPE participants outperformed VRE participants in the Spatial Reconstruction task when participants were asked to name each room on a map. At the same time, we note that in MPE, each room was introduced with a slide specifying the room name; however, this was infeasible in the VRE and FPE experiences due to the limitations of the first-person view perspective. Indeed, we introduced signs with the room names nearby each lecture point; however, the solution could not solve the issue. Furthermore, 2D maps are more familiar tools to navigate environments; accordingly, we believe that a more familiar interaction favored participants in MPE compared to VRE and MPE participants who only interacted with 3D models and environments. Even so, thanks to the overall results, we argue that our experience seems efficient as widely credited types of complementary educational tools, particularly in teaching basic knowledge about ancient roman houses.

Figure 10: Two participants testing the VRE

Regarding UX, our results show that the PQ score is significantly higher for VRE than FPE; however, we recorded no significant difference for MPE. As PQ sustains the *do-goals* of the individual in using a product/technology (Hassenzahl, 2008), our results suggest that VRE might be considered more efficient in explaining the structure of an Ancient Roman Domus than a desktop experience. Still, this difference could not be detected for a more familiar educational tool like MPE. Our results show significant differences in the HQ score between all the conditions. VRE obtained the highest positive score, suggesting it could efficiently support the *be-goals* (Hassenzahl, 2008). Analyzing these two results together, we can argue that participants found VRE more enjoyable and at least equally efficient than the other experiences. This is an encouraging outcome, as users perceived VR as equally effective as more conventional tools in conveying historical knowledge, even though it was a new instrument they were unfamiliar with. Based on these findings, we hypothesize that with broader adoption and use of VR technology, its efficacy will continue to improve, making it an increasingly valuable support tool in history classes.

Regarding the sense of presence, VRE participants outperformed FPE participants, which aligns with our initial hypothesis H.4. Furthermore, the VRE group scored higher in Involvement and Immersion, suggesting that VRE participants felt more engaged and immersed than FPE participants. Combined with the analysis of the *General Learning* variable, these results are promising as they suggest that the increased level of sense of presence provided by VR technology did not hurt the learning outcome, as we also hypothesized (H.4). While this does not definitively establish the effectiveness of VR as a learning tool, it provides a solid starting point for further investigation.

Motion sickness is a critical aspect of VR applications that can dramatically affect the users' experience in the virtual world. Accordingly, we designed the VRE experience to avoid this issue (Section 3.5). Our results show that, in line with H.3, we were successful in our endeavor. In VRE, we recorded a negligible level of *Cybersickness* comparable to the one experienced by the users in FPE. Thus, our application did not cause significant motion sickness; most importantly, participants classified the occurred symptoms mainly as "light" when this happened. By measuring this construct, we have also ruled out possible confounding variables that could impact the educational outcome. We consider this a significant achievement, demonstrating the effectiveness of our strategies in preventing the adverse side effects of VR technology.

5.2. Limitations of the Experiment

Our work seems to be not exempt from limitations. One of the problems involves both our sample size and its age range. Regarding the first issue, we managed to test our work with a limited amount of participants due to current regulations related to the COVID-19 pandemic that prevented us from making more people access the laboratory. Moreover, we had an unbalanced testing group: the MPE was tested by almost ten participants more than the other two conditions. This issue, caused by the regulations related to the COVID-19 pandemic too, impacted the types of analyses we could perform: having unbalanced samples limited our choices for the statistical analyses.

We distributed two of the three testing conditions telematically: due to the regulations related to the COVID-19 pandemic, we were forced to find a solution for collecting data in a safe environment, asking the participant to move from their houses if, and only if, there was no alternative (as in the case of the VRE testing). This is a limitation of our research: we had low control over contextual variables that could have influenced our results. Moreover, finding participants for the telematically-distributed conditions was easier, which brought an unbalanced sample size for the MPE condition. Moreover, we had first to ask the participants if they could move to the laboratory, which may have influenced the randomness of the users' condition assignment. However, the randomness is fully guaranteed between the FPE and the MPE.

The last two limitations directly involved our VE. As explained in Section 3, while introducing the overall structure of the Domus, we implemented two different strategies: a 2D map with tags for the MPE and a rotating 3D map for the FPE and MPE. In our opinion, this choice influenced our results in the Room Name section, favoring the MPE condition with a more intuitive and familiar presentation of information. Another issue we faced in our study regards the tutorial presented at the beginning of the VRE. Some participants requested help from the experimenters to master the use of the application. For this reason, we believe that an upgrade of our initial tutorial is needed. This could be characterized by a virtual room where participants can try all the commands before entering the "real" experience.

6. Conclusions and Future Research Directions

In this work, we described the design and development of a research project focused on understanding if and how VR could be a suitable technological support for educational purposes. We prepared three interactive learning experiences: one for VR HMDs (in particular, Oculus Rift and Rift S), one for desktop PCs, and a more traditional slide-based lecture in PowerPoint. All the experiences explained the structure of an ancient Roman Domus by means of a digital reconstruction of the environment and recorded audio descriptions. We collected data from our participants through a questionnaire adapted for all three conditions and composed of previously-validated and novel items. The analysis of the results was promising, showing that the VRE was rated as more pleasant and at least equally efficient than the other two conditions to explain the chosen topic. Furthermore, the higher level of Sense of Presence provided by the VRE compared to the FPE did not negatively impact the learning outcome. However, some limitations in the sample balancing (mainly due to the current regional regulations related to dealing with the pandemic emergency) and the application design reduce the extent of our findings. Moreover, due to the already-stated limitations, this work is hardly generalizable. However, we think this work is worth being published, as it offers a good perspective on designing and testing VR applications for cultural heritage. Future research directions should include better control of the sample size (both in the number of participants and the age range) and more attention to some design issues we identified. Moreover, to further increase the users' immersion in the experience and test the modality of Exploration in both the VRE and the FPE, we aim to add the implementation of the VRE on the Cyberith's Virtualizer[4]. We would also like to assess the value of our application in a more ecological setting, like inside a museum exhibit or a classroom, to check to a greater extent its educational value. Finally, we would like to test also the Exploration Mode, adding more interaction and simple tasks to complete, to test if the integration between movement freedom, knowledge points, and interactive elements could increase the results, leading to better and more performant learning experiences.

Acknowledgments

We are grateful to Maria Teresa Belardinelli, a former high school history and Latin teacher, for assisting us in locating historical sources and developing the educational content. Our appreciation also extends to Albert Gregl, the author of Gregl (2020), for generously sharing the Domus model used in our project.

Statements on Open Data and Ethics

The study was conducted following the ethical guidelines of the University of Milano Bicocca's ethics committee. The study received approval from the committee (code $RM - 2021 - 402$), and all participants provided informed consent before participating. The data collected were kept secure to ensure confidentiality and anonymity, and the study was designed to minimize any potential risks or discomfort for the participants.

Declaration of Competing Interest

None of the authors identified a conflict of interest or personal relationships that could have appeared to influence the work reported in this paper.

[4] https://www.cyberith.com

Acronyms

VR	Virtual Reality	1
WS	Witmer and Singer	3
VE	Virtual Environment	3
SUS	Slater-Usoh-Steed	4
VRE	Virtual Reality Experience	5
FPE	First Person Experience	7
MPE	Multi-media Presentation Experience	7
UX	User Experience	9
TAM	Technology Acceptance Model	10
PQ	Pragmatic Qualities	10
HQ	Hedonic Qualities	10
HMD	Head-Mounted Display	7
PS	Presence Score	12

References

Agarwal, R., Prasad, J., 1998. A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. Information Systems Research 9, 204–215. doi:https://doi.org/10.1287/isre.9.2.204.

Akbulut, A., Catal, C., Yıldız, B., 2018. On the effectiveness of virtual reality in the education of software engineering. Computer Applications in Engineering Education 26, 918–927. doi:10.1002/cae.21935.

Allcoat, D., von Mühlenen, A., 2018. Learning in virtual reality: Effects on performance, emotion and engagement. Research in Learning Technology 26. doi:10.25304/rlt.v26.2140.

Angela, A., 2007. Una Giornata nell'Antica Roma. Mondadori.

Bekele, M.K., Pierdicca, R., Frontoni, E., Malinverni, E.S., Gain, J., 2018. A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage. Journal on Computing and Cultural Heritage 11. doi:https://doi.org/10.1145/3145534.

Bell, J., Fogler, H.S., 1995. The Investigation and Application of Virtual Reality as an Educational Tool, in: Proceedings of the American Society for Engineering Education Annual Conference, pp. 1718–1728. [Online]. Available **here**.

Biocca, F., Delaney, B., 1995. Immersive Virtual Reality Technology. L. Erlbaum Associates Inc. pp. 57–124.

Chang, E., Kim, H.T., Yoo, B., 2020. Virtual Reality Sickness: A Review of Causes and Measurements. International Journal of Human-Computer Interaction , 1658–1682doi:10.1080/10447318.2020.1778351.

Christou, C., 2010. Virtual Reality in Education. chapter 12. pp. 228–243. doi:10.4018/978-1-60566-940-3.ch012.

Clifton, J., Palmisano, S., 2020. Effects of Steering Locomotion and Teleporting on Cybersickness and Presence in HMD-based Virtual Reality. Virtual Reality 24, 453–468. doi:https://doi.org/10.1007/s10055-019-00407-8.

Colombo, C., Di Blas, N., Gkolias, I., Lanzi, P.L., Loiacono, D., Stella, E., 2020. An educational experience to raise awareness about space debris. IEEE Access 8, 85162–85178. doi:10.1109/ACCESS.2020.2992327.

Davis, F.D., Bagozzi, R.P., Warshaw, P.R., 1989. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science 35, 982–1003. doi:https://doi.org/10.1287/mnsc.35.8.982.

De Paolis, L., Aloisio, G., Celentano, M., Oliva, L., Vecchio, P., 2009. Design and Development of a Virtual Reality Application for Edutainment in Cultural Heritage. International Conference on Virtual Systems and MultiMedia , 80–84doi:10.1109/VSMM.2009.18.

Dede, C., Salzman, M., Bowen Loftin, R., 1996. ScienceSpace: virtual realities for learning complex and abstract scientific concepts, in: Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium, pp. 246–252. doi:10.1109/VRAIS.1996.490534.

Deggim, S., Kersten, T.P., Tschirschwitz, F., Hinrichsen, N., 2017. Segeberg 1600 - Reconstructing a Historic Town for Virtual Reality Visualisation as an Immersive Experience. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-2/W8, 87–94. doi:10.5194/isprs-archives-XLII-2-W8-87-2017.

Dużmańska, N., Strojny, P., Strojny, A., 2018. Can Simulator Sickness Be Avoided? A Review on Temporal Aspects of Simulator Sickness. Frontiers in Psychology 9. doi:10.3389/fpsyg.2018.02132.

Edensor, T., 2015. Light design and atmosphere. Visual Communication 14, 331–350. doi:10.1177/1470357215579975.

Fabola, A., Miller, A., 2016. Virtual reality for early education: A study, pp. 59–72. doi:10.1007/978-3-319-41769-1_5.

Fauville, G., Muller Queiroz, A.C., Bailenson, J.N., 2020. Virtual Reality as a Promising Tool to Promote Climate Change Awareness, in: Kim, J., Song, H. (Eds.), Technology and Health. Academic Press, pp. 91–108. doi:https://doi.org/10.1016/B978-0-12-816958-2.00005-8.

Fleury, P., Madeleine, S., 2012. Reviving Ancient Rome: Virtual Reality at the Service of Cultural Heritage, in: Ioannides, M., Fritsch, D., Leissner, J., Davies, R., Remondino, F., Caffo, R. (Eds.), Progress in Cultural Heritage Preservation, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 159–169. doi:10.1007/978-3-642-34234-9_16.

Flocchini, N., Bacci, P.G., Moscio, M., Sampietro, M., Lamagna, P., 2017. Lingua e Cultura Latina. Bompiani.

Flyover Zone, 2018. Rome Reborn. https://www.romereborn.org. [Online]. Accessed: 2021-11-15.

Gonzalez Izard, S., Juanes, J., Ruisoto, P., 2017. Virtual Reality Educational Tool for Human Anatomy. Journal of Medical Systems 41. doi:10.1007/s10916-017-0723-6.

Grassini, S., Laumann, K., 2020. Questionnaire measures and physiological correlates of presence: A systematic review. Frontiers in Psychology 11. doi:10.3389/fpsyg.2020.00349.

Greenwald, S.W., Kulik, A., Kunert, A., Beck, S., Frohlich, B., Cobb, S., Parsons, S., Newbutt, N., Gouveia, C., Cook, C., Snyder, A., Payne, S., Holland, J., Buessing, S., Fields, G., Corning, W., Lee, V., Xia, L., Maes, P., 2017. Technology and Applications for Collaborative Learning in Virtual Reality, in: Proc. 12th Int. Conf. Comput. Supported Collaborative Learn. (CSCL), pp. 719–726. [Online]. Available: https://repository.isls.org/handle/1/210.

Gregl, A., 2020. Domus Romana. https://www.albertgregl.com/domusromana/. [Online]. Accessed: 2021-02-23.

Guna, J., Geršak, G., Humar, I., Song, J., Drnovšek, J., Pogačnik, M., 2019. Influence of video content type on users' virtual reality sickness perception and physiological response. Future Generation Computer Systems 91, 263–276. doi:https://doi.org/10.1016/j.future.2018.08.049.

Häkkilä, J., Hannula, P., Luiro, E., Launne, E., Mustonen, S., Westerlund, T., Colley, A., 2019. Visiting a Virtual Graveyard: Designing Virtual Reality Cultural Heritage Experiences, in: Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia, Association for Computing Machinery. pp. 1–4. doi:10.1145/3365610.3368425.

Hassenzahl, M., 2004. The Thing and I: Understanding the Relationship Between User and Product. Springer Netherlands, Dordrecht. chapter 3. pp. 31–42. doi:10.1007/1-4020-2967-5_4.

Hassenzahl, M., 2008. User Experience (UX): Towards an Experiential Perspective on Product Quality, in: Proceedings of the 20th Conference on l'Interaction Homme-Machine, Association for Computing Machinery. pp. 11–15. doi:10.1145/1512714.1512717.

Insko, B.E., 2003. Measuring presence: Subjective, behavioral and physiological methods, in: Being there: Concepts, effects and measurements of user presence in synthetic environments, pp. 109–119.

Ionescu, T., Vasc, D., 2014. Embodied Cognition: Challenges for Psychology and Education. Procedia - Social and Behavioral Sciences 128, 275–280. doi:https://doi.org/10.1016/j.sbspro.2014.03.156.

Ip, H.H.S., Li, C., Leoni, S., Chen, Y., Ma, K.F., Wong, C.H., Li, Q., 2019. Design and Evaluate Immersive Learning Experience for Massive Open Online Courses (MOOCs). IEEE Transactions on Learning Technologies 12, 503–515. doi:10.1109/TLT.2018.2878700.

Kamińska, D., Sapiński, T., Wiak, S., Tikk, T., Haamer, R.E., Avots, E., Helmi, A., Ozcinar, C., Anbarjafari, G., 2019. Virtual Reality and Its Applications in Education: Survey. Information 10. doi:10.3390/info10100318.

Kavanagh, S., Luxton-Reilly, A., Wuensche, B., Plimmer, B., 2017. A Systematic Review of Virtual Reality in Education. Themes in Science and Technology Education 10, 85–119.

Kim, H.K., Park, J., Choi, Y., Choe, M., 2018. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. Applied Ergonomics 69, 66–73. doi:https://doi.org/10.1016/j.apergo.2017.12.016.

Kourtesis, P., Collina, S., Doumas, L.A.A., MacPherson, S.E., 2019. Technological competence is a pre-condition for effective implementation of virtual reality head mounted displays in human neuroscience: A technological review and meta-analysis. Frontiers in Human Neuroscience 13. doi:10.3389/fnhum.2019.00342.

Kumar, V., Gulati, S., Deka, B., Sarma, H., 2021. Teaching and Learning Crystal structures through Virtual Reality based systems. Advanced Engineering Informatics 50, 101362. doi:https://doi.org/10.1016/j.aei.2021.101362.

Kyrlitsias, C., Christofi, M., Michael-Grigoriou, D., Banakou, D., Ioannou, A., 2020. A Virtual Tour of a Hardly Accessible Archaeological Site: The Effect of Immersive Virtual Reality on User Experience, Learning and Attitude Change. Frontiers in Computer Science 2, 23. doi:10.3389/fcomp.2020.00023.

LaViola, J.J., 2000. A discussion of cybersickness in virtual environments. SIGCHI Bull. 32, 47–56. doi:10.1145/333329.333344.

Lee, J., Kim, M., Kim, J., 2017. A study on immersion and vr sickness in walking interaction for immersive virtual reality applications. Symmetry 9. doi:10.3390/sym9050078.

Lessiter, J., Freeman, J., Keogh, E., Davidoff, J., 2001. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. Presence: Teleoperators and Virtual Environments 10, 282–297. doi:10.1162/105474601300343612.

Liu, D., Bhagat, K.K., Gao, Y., Chang, T.W., Huang, R., 2017. The Potentials and Trends of Virtual Reality in Education. Springer Singapore. chapter 7. pp. 105–130. doi:10.1007/978-981-10-5490-7_7.

Liu, Y., Lin, Y., Shi, R., Luo, Y., Liang, H.N., 2021. RelicVR: A Virtual Reality Game for Active Exploration of Archaeological Relics, in: Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play, Association for Computing Machinery, New York, NY, USA. pp. 326–332. doi:10.1145/3450337.3483507.

Mahmoud, K., Harris, I., Yassin, H., Hurkxkens, T.J., Matar, O.K., Bhatia, N., Kalkanis, I., 2020. Does Immersive VR Increase Learning Gain When Compared to a Non-immersive VR Learning Experience?, in: Zaphiris, P., Ioannou, A. (Eds.), Learning and Collaboration Technologies. Human and Technology Ecosystems, Springer International Publishing. pp. 480–498. doi:10.1007/978-3-030-50506-6_33.

Makransky, G., Terkildsen, T.S., Mayer, R.E., 2019. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learning and Instruction 60, 225–236. doi:https://doi.org/10.1016/j.learninstruc.2017.12.007.

Marco Vitruvio Pollione, 15 B.C. De Architectura. volume 6.

Mortara, M., Catalano, C.E., Bellotti, F., Fiucci, G., Houry-Panchetti, M., Petridis, P., 2014. Learning cultural heritage by serious games. Journal of Cultural Heritage 15, 318–325. doi:10.1016/j.culher.2013.04.004.

Ou, K.L., Chu, S.T., Tarng, W., 2021. Development of a Virtual Wetland Ecological System Using VR 360° Panoramic Technology for Environmental Education. Land 10. doi:10.3390/land10080829.

Paíno Ambrosio, A., Rodríguez Fidalgo, M.I., 2020. Past, present and future of Virtual Reality: Analysis of its technological variables and definitions. Culture & History Digital Journal , 10–2253doi:https://doi.org/10.3989/chdj.2020.010.

Paoli, U.E., 1982. Vita Romana. Mondadori.

Parong, J., Mayer, R., 2018. Learning science in immersive virtual reality. Journal of Educational Psychology 110. doi:10.1037/edu0000241.

Pellas, N., Mystakidis, S., Kazanidis, I., 2021. Immersive Virtual Reality in K-12 and Higher Education: A systematic review of the last decade scientific literature. Virtual Reality 25, 835–861. doi:10.1007/s10055-020-00489-9.

Perez-Valle, A., Sagasti, D., 2012. A novel approach for tourism and education through virtual vitoria-gasteiz in the 16th century, in: 2012 18th International Conference on Virtual Systems and Multimedia, pp. 615–618. doi:10.1109/VSMM.2012.6365991.

Pinto, D., Peixoto, B., Krassmann, A., Melo, M., Cabral, L., Bessa, M., 2019. Virtual Reality in Education: Learning a Foreign Language, in: Rocha, Á., Adeli, H., Reis, L.P., Costanzo, S. (Eds.), New Knowledge in Information Systems and Technologies, Springer International Publishing. pp. 589–597. doi:10.1007/978-3-030-16187-3_57.

Radianti, J., Majchrzak, T.A., Fromm, J., Wohlgenannt, I., 2020. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. Computers & Education 147. doi:https://doi.org/10.1016/j.compedu.2019.103778.

Rese, A., Baier, D., Geyer-Schulz, A., Schreiber, S., 2017. How augmented reality apps are accepted by consumers: A comparative analysis using scales and opinions. Technological Forecasting and Social Change 124, 306–319. doi:https://doi.org/10.1016/j.techfore.2016.10.010.

Richards-Rissetto, H., Robertsson, J., von Schwerin, J., Agugiaro, G., Remondino, F., Girardi, G., 2013. Geospatial Virtual Heritage: A Gesture-Based 3D GIS to Engage the Public with Ancient Maya Archaeology. Amsterdam University Press. chapter 12. pp. 118–130. doi:https://doi.org/10.2307/j.ctt6wp7kg.

Roussou, M., 2009. A VR Playground for Learning Abstract Mathematics Concepts. IEEE Computer Graphics and Applications 29, 82–85. doi:10.1109/MCG.2009.1.

Sagnier, C., Loup-Escande, E., Lourdeaux, D., Thouvenin, I., Valléry, G., 2020. User Acceptance of Virtual Reality: An Extended Technology Acceptance Model. International Journal of Human-Computer Interaction 36, 993–1007. doi:10.1080/10447318.2019.1708612.

Salloum, S.A., Qasim Mohammad Alhamad, A., Al-Emran, M., Abdel Monem, A., Shaalan, K., 2019. Exploring Students' Acceptance of E-Learning Through the Development of a Comprehensive Technology Acceptance Model. IEEE Access 7, 128445–128462. doi:10.1109/ACCESS.2019.2939467.

Saredakis, D., Szpak, A., Birckhead, B., Keage, H.A.D., Rizzo, A., Loetscher, T., 2020. Factors associated with virtual reality sickness in head-mounted displays: A systematic review and meta-analysis. Frontiers in Human Neuroscience 14. doi:10.3389/fnhum.2020.00096.

Shapiro, L., 2007. The Embodied Cognition Research Programme. Philosophy Compass 2, 338–346. doi:https://doi.org/10.1111/j.1747-9991.2007.00064.x.

Slater, M., 1999. Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. Presence: Teleoperators and Virtual Environments 8, 560–565. doi:10.1162/105474699566477.

Slater, M., 2003. A note on presence terminology. [Online]. Available at: https://tinyurl.com/slater2003.

Taranilla, R.V., Cózar-Gutiérrez, R., González-Calero, J.A., Cirugeda, I.L., 2022. Strolling through a city of the roman empire: an analysis of the potential of virtual reality to teach history in primary education. Interactive Learning Environments 30, 608–618. doi:10.1080/10494820.2019.1674886.

Usoh, M., Catena, E., Arman, S., Slater, M., 2000. Using presence questionnaires in reality. Presence: Teleoperators and Virtual Environments 9, 497–503. doi:10.1162/105474600566989.

Villena Taranilla, R., Cozar-Gutierrez, R., Gonzalez-Calero, J.A., Lopez Cirugeda, I., 2019. Strolling through a city of the Roman Empire: an analysis of the potential of virtual reality to teach history in Primary Education. Interactive Learning Environments , 1–11doi:10.1080/10494820.2019.1674886.

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

Wang, L., 2022. How does light shape the environment and affect human perception in the virtual environment, in: Proceedings of the 2022 International Conference on Social Sciences and Humanities and Arts (SSHA 2022), pp. 217–221. doi:10.2991/assehr.k.220401.044.

Witmer, B.G., Jerome, C.J., Singer, M.J., 2005. The Factor Structure of the Presence Questionnaire. Presence 14, 298–312. doi:10.1162/105474605323384654.

Witmer, B.G., Singer, M.J., 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence: Teleoperators and Virtual Environments 7, 225–240. doi:10.1162/105474698565686.

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, 1991–2005. doi:https://doi.org/10.1111/bjet.13023.

Yu, I., Mortensen, J., Khanna, P., Spanlang, B., Slater, M., 2012. Visual realism enhances realistic response in an immersive virtual environment - part 2. IEEE Computer Graphics and Applications 32, 36–45. doi:10.1109/MCG.2012.121.

A. Tables and Numerical Results

VRSQ symptom	Oculomotor [1]	Disorientation [2]
General Discomfort	o	
Fatigue	o	
Eyestrain	o	
Difficulty focusing	o	
Headache		o
Fullness of head		o
Blurred vision		o
Dizzy (eyes closed)		o
Vertigo		o
SSQ components	**Computation**	
Oculomotor [3]	$([1]/12) * 100$	
Disorientation [4]	$([2]/15) * 100$	
Total	$([3] + [4])/2$	

Table A.1
Virtual Reality Sickness Questionnaire (VRSQ) and its computation score (Kim et al., 2018)

	VRE	**FPE**	**MPE**	**Full sample**	
	n	*n*	*n*	*n*	%
Gender					
Female	12	10	18	40	52.6
Male	10	13	13	36	47.4
Total	22	23	31	76	100
Educational level					
High School	8	5	6	19	25.0
B.Sc.	13	9	19	41	53.9
M.Sc.	1	8	6	15	19.7
Ph.D.	0	1	0	1	1.3
Age					
Mean	23.6	24.6	24	24.1	
Median	24	25	23	24	

Table A.2
Participants demographic information

CONSTRUCT & SOURCE	SELECTED ITEMS	SCALE	EXAMPLES
Personal Innovativeness (PI) (Agarwal and Prasad, 1998; Sagnier et al., 2020)	All	Likert from 1 (totally disagree) to 7 (completely agree)	Among my peers, I am usually the first to try out new technology
Technology Acceptance Model (TAM) (Davis et al., 1989; Rese et al., 2017; Salloum et al., 2019)	Rese et al., 2017: PEOU2, PEOU3, PU1, PU3, AT2, BI2, BI3, BI4. Salloum et al., 2019: PEOU3, PU3, ATT1, ATT3	Likert from 1 (totally disagree) to 7 (completely agree)	This simulation was intuitive to use
User Experience (UX) (Hassenzahl, 2004; Sagnier et al., 2020)	All	Likert from 1 to 7	Cumbersome - Straightforward
Sense of Presence (PS) (Witmer and Singer, 1998; Witmer et al., 2005)	4, 5, 8, 10, 11, 14, 16, 17, 19, 20, 21, 22, 23, 25, 30, 32	Likert from 1 to 7 (not at all - somewhat - completely)	How much delay did you experience between your actions and expected outcome?
Cybersickness (VRSQ) (Kim et al., 2018)	All	Likert with values 0 (not at all), 1 (slightly), 2 (moderate), 3 (very)	General discomfort, Fatigue, Eyestrain

Table A.3
Overview of the questionnaire components taken from the literature; we did not consider the TAM items in our analyses because they were beyond the scope of this paper

	VRE		FPE		MPE		F	η^2
	M	SD	M	SD	M	SD		
Learning	21.45	3.203	21.30	4.875	22.48	5.452	0.453	
General Learning	12.09	2.091	11.61	3.159	11.42	3.149	0.473	
Spatial Task	9.36	2.083	9.70	2.687	11.06	2.774	3.457	0.085
Room Name	5.32	1.756	5.61	2.061	6.94	1.948	5.555	0.130
Object Location	4.05	0.722	4.09	0.9	4.13	1.088	0.056	

Table A.4
Means, Standard Deviations and Welch's One-Way Analyses of Variance of Learning

	VRE		FPE		MPE		F	η^2	*Cronbach's* α
	M	SD	M	SD	M	SD			
Pragmatic Quality (PQ)	6.08	0.479	5.67	0.567	5.78	0.782	3.68	0.063	0.648
Hedonic Quality (HQ)	5.95	0.635	5.43	0.716	4.90	1.078	10.34	0.210	0.847

Table A.5
Means, Standard Deviations and Welch's One-Way Analyses of Variance of User Experience

	VRE		**FPE**			
	M	SD	M	SD	$t(43)$	*Cohen's d*
Presence Score (PS)	88.8	10.1	80.3	12.1	2.57	0.766

Table A.6
Means, Standard Deviations and Student's Independent Samples T-test of the Presence Questionnaire Score

	VRE		**FPE**				
	M	SD	M	SD	$t(43)$	*Cohen's d*	*Cronbach's* α
Involvement	5.56	0.578	4.63	0.937	4.06	1.204	0.644
Sensory Fidelity	4.98	1.568	5.12	1.157	0.318		0.716
Immersion	5.68	0.716	5.19	0.892	2.05	0.609	0.732
Interface Quality	5.95	0.792	5.81	1.184	0.478		0.624

Table A.7
Means, Standard Deviations and Student's Independent Samples T-test of the Sense of Presence components

	VRE		**FPE**		Total
	Obs	Exp	Obs	Exp	
No Symptoms	9	8.69	8	8.31	17
Yes Symptoms	14	14.31	14	13.69	28
Total	22		23		

Table A.8
Contingency table of Cybersickness divided in conditions and relative Chi-Square statistic

	VRE				**FPE**			
	M	Mdn	$Mode$	SE	M	Mdn	$Mode$	SE
Cybersickness	7.58	7.08	0.00	2.07	6.16	4.17	0.00	1.9
Oculomotor	5.95	0.635	5.43	0.716	4.90	1.078	10.34	
Disorientation	5.95	0.635	5.43	0.716	4.90	1.078	10.34	

Table A.9
Cybersickness and relative components' descriptives

	Group	INVOLVEMENT	SENSORY FIDELITY	IMMERSION	INTERFACE QUALITY	PS
N	FPE	23	23	23	23	23
	VRE	22	22	22	22	22
Missing	FPE	0	0	0	0	0
	VRE	0	0	0	0	0
Mean	FPE	4.63	5.12	5.19	5.81	80.3
	VRE	5.56	4.98	5.68	5.95	88.8
Median	FPE	4.60	5.33	5.33	6.33	84.0
	VRE	5.50	5.50	5.67	6.17	88.0
Standard Deviation	FPE	0.937	1.16	0.892	1.18	12.1
	VRE	0.578	1.57	0.716	0.792	10.1
Minimum	FPE	2.80	3.00	2.67	3.00	55.00
	VRE	4.60	1.67	4.67	4.00	73.0
Maximum	FPE	6.00	7.00	6.33	7.00	101
	VRE	7.00	7.00	7.00	7.00	112

Table A.10
Presence Questionnaire complete descriptives

	Group	OCULOMOTOR	DISORIENTATION	CYBERSICKNESS
N	FPE	23	23	23
	VRE	22	22	22
Missing	FPE	0	0	0
	VRE	0	0	0
Mean	FPE	9.42	2.90	6.16
	VRE	9.09	6.06	7.58
Median	FPE	8.33	0.00	4.17
	VRE	8.33	0.00	7.08
Standard Deviation	FPE	13.1	6.61	9.09
	VRE	12.8	9.41	9.69
Minimum	FPE	0.00	0.00	0.00
	VRE	0.00	0.00	0.00
Maximum	FPE	50.0	26.7	31.7
	VRE	58.3	33.3	42.5

Table A.11
Cybersickness complete descriptives

	Group	Age	Learning	PEU	PU	Attitude	Behavior	Pragmatic	Hedonic
N	MPE	31	31	31	31	31	31	31	31
	FPE	23	23	23	23	23	23	23	23
	VRE	22	22	22	22	22	22	22	22
Missing	MPE	0	0	0	0	0	0	0	0
	FPE	0	0	0	0	0	0	0	0
	VRE	0	0	0	0	0	0	0	0
Mean	MPE	24.0	22.5	5.67	5.45	5.42	4.88	5.78	4.90
	FPE	24.6	21.3	5.77	5.46	5.84	5.36	5.67	5.43
	VRE	23.6	21.5	5.76	6.20	6.11	5.56	6.08	5.95
Median	MPE	23	24.0	6.00	5.33	5.67	4.67	5.83	5.00
	FPE	25	22.0	6.00	5.67	6.00	5.67	5.67	5.50
	VRE	24	21.0	5.67	6.33	6.33	5.67	6.00	6.08
Standard Deviation	MPE	2.94	5.45	1.16	1.05	1.01	1.30	0.782	1.08
	FPE	3.94	4.88	0.901	1.02	0.926	1.23	0.567	0.716
	VRE	1.79	3.20	1.06	0.746	0.678	1.06	0.479	0.635
Minimum	MPE	19	7.00	2.33	2.00	2.67	1.67	3.83	1.33
	FPE	18	11.0	4.00	3.33	4.00	2.67	4.67	3.83
	VRE	20	16.0	2.67	4.33	5.00	4.00	5.17	4.33
Maximum	MPE	32	28.0	7.00	7.00	7.00	7.00	7.00	6.50
	FPE	33	28.0	7.00	7.00	7.00	7.00	6.83	6.67
	VRE	28	27.0	7.00	7.00	7.00	7.00	7.00	6.67

Table A.12
Learning, TAM, and UX complete descriptives