



HAL
open science

Immersive learning of factual knowledge while assessing the influence of cognitive load and spatial abilities

Benjamin de Witte, Vincent Reynaert, Jérôme Hutain, Danny Kieken, Joseph Jabbour, Jalal Possik

► To cite this version:

Benjamin de Witte, Vincent Reynaert, Jérôme Hutain, Danny Kieken, Joseph Jabbour, et al.. Immersive learning of factual knowledge while assessing the influence of cognitive load and spatial abilities. *Computers & Education: X Reality*, 2024, 5, pp.100085. 10.1016/j.cexr.2024.100085 . hal-04727241

HAL Id: hal-04727241

<https://hal.science/hal-04727241v1>

Submitted on 9 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Immersive learning of factual knowledge while assessing the influence of cognitive load and spatial abilities

Benjamin De Witte^{a,e,*}, Vincent Reynaert^{a,e}, Jerome Hutain^{a,b}, Danny Kieken^c, Joseph Jabbour^{c,d}, Jalal Possik^c

^a Ethics Laboratory EA 7446, Lille catholic University, 14 Boulevard Vauban, 59000, Lille, France

^b Univ Rennes, LP3C (Psychology of Cognition, Behavior, Communication Laboratory), Rennes, France

^c LITL, Junia, Lille catholic University, Boulevard Vauban, 59000, Lille, France

^d Risk Sciences Laboratory, IMT Mines Alès, Alès, 30100, France

^e Piktura School 78 Boulevard Du Général Leclerc, 59100, Roubaix, France

ARTICLE INFO

Keywords:

Augmented and virtual reality
Media in education
Teaching/learning strategies
Improving classroom teaching

ABSTRACT

Immersive virtual reality (IVR) is a motivating new learning tool. However, discrepancies in results emerge regarding the type of knowledge, specifically factual knowledge. Moreover, to improve the instructional design, authors claimed to verify the impact of cognitive load and spatial abilities while learning in IVR. Sixty-one undergraduate students were recruited and divided into two groups in pre-post, test study design. Learning consisted of using “Google Earth” with a Head Mounted Display (HMD) or a desktop computer to improve general geographic knowledge. Pre-training and segmentation were included in the learning process, and levels of presence, spatial orientation and cognitive load were measured to evaluate their impact. Results show that the two groups significantly improved their performance between pre and post-test. HMD users felt more present than computer users. Overall cognitive load did not hinder the learning performance. The analysis of cognitive load dimensions reveals significant differences in the mental load and frustration levels between the two groups. Spatial orientation was not correlated to the learning process. The use of HMD is effective for learning geographic general knowledge. When managed overall cognitive load does not influence learning although the use of HMDs elicits high mental load levels. HMD users are experiencing less frustration indicating it may be suitable for long-distance learning. HMD technological features compensate for spatial orientation abilities. Finally, presence does not enhance learning, suggesting it should be combined with interactions eliciting feedback to be more effective.

1. Introduction

The integration of technology into education has significantly transformed traditional learning paradigms, with Immersive Virtual Reality (IVR) emerging as a particularly promising tool. IVR offers unique opportunities to enhance learning by immersing students in virtual environments that closely replicate real-world scenarios, thereby fostering deeper engagement and improved knowledge retention. While IVR has shown effectiveness across various domains, such as skills training and science education, its application in acquiring factual knowledge remains underexplored, with mixed results in existing studies. This study aims to address these gaps by investigating the impact of IVR on the acquisition of general geography knowledge, with a focus on the roles of cognitive processes, such as, cognitive load and

spatial abilities in the learning process.

To achieve this, we will employ a pre-post-test experimental design, comparing learning outcomes between an IVR group using Head-Mounted Displays (HMDs) and a Non-Immersive Virtual Reality (NIVR) group using desktop computers. Additionally, we will explore how instructional design principles, such as pre-training and segmentation, can mitigate cognitive load and enhance learning outcomes in IVR environments. By examining these factors, the study seeks to contribute to a deeper understanding of the instructional potential of IVR, particularly in the context of factual knowledge acquisition.

2. Literature review

To provide a comprehensive understanding of the role of IVR in

* Corresponding author. laboratoire Ethics Maison des chercheurs, 14 Boulevard Vauban, 59000, Lille, France.

E-mail address: benjamin.dewitte@univ-catholille.fr (B. De Witte).

education, this literature review is organized into four key sections. The first part discusses the fundamentals of immersive learning and the distinction between IVR and NIVR. The second part delves into the cognitive theories that underpin multimedia learning, particularly focusing on cognitive load in IVR environments. The third section explores the influence of spatial abilities on learning outcomes with IVR. Finally, the fourth section sets forth the objective of the study and the hypotheses that have been formulated.

2.1. Immersive learning

Education, like many other aspects of daily life, is influenced by technology (Raja & Nagasubramani, 2018). More precisely, virtual reality is emerging as a new learning tool. Indeed, Hamilton et al. (2021) noted that the decreasing cost of virtual reality devices and the increasing graphical fidelity of virtual environments are contributing to the proliferation of virtual reality in education. Virtual reality can be split into two categories; NIVR accessed through tools as a desktop computer providing virtual reality, and immersive virtual reality IVR accessed through tools as a cave automatic virtual environment (CAVE) or HMD (Buttussi & Chittaro, 2017). Learning with both, non-immersive and immersive virtual reality allows the user to be actively involved in the learning process (i.e. active learning) (Mayer et al., 2022). However, compared to non-immersive virtual reality, which allows to see the virtual world from outside, IVR provides the opportunity to “feel” inside the virtual world. The psychological construct of “feeling” in the virtual world, is called “sense of presence” and is made possible by the degree of immersion provided by the technological device (Makransky & Petersen, 2021). Likewise, when the technological device enables the feeling of generating and controlling actions, the user experiences a sense of “agency” (Moore & Fletcher, 2012) which reinforces the feeling of being in the virtual world. Both, presence and agency facilitate learning as they are general psychological affordances of learning in IVR (Makransky & Petersen, 2021). Improvements in the technological features of HMDs (e.g. field of view; frame rate; level of detail; controllers; tracking technology) can enhance the sense of presence and agency (Cummings & Bailenson, 2016; Jicol et al., 2023, pp. 1–17). In other words, modern HMDs are more likely to offer technological features that promote a greater sense of presence and agency (Cummings & Bailenson, 2016).

As said, immersive virtual reality can be accessed by HMDs or a CAVE. Given its lower price and higher availability, studies investigated more particularly HMDs as IVR learning tool (Wu et al., 2020). Indeed, the portability and size of HMDs, while offering high technological features (e.g. field, of view, frame rate, eye tracking) obviously facilitate experimentation. Results show for example that IVR is more efficient than traditional learning tools (e.g. 2D didactic presentations) for skills training in healthcare procedures (Ryan et al., 2022), safety training or science education (Wu et al., 2020). Indeed, providing opportunities to train or learn high hazard procedures without the fear of making a mistake, promotes learning (Ryan et al., 2022) and improves skills ability in the real world as there is evidence that virtual skill acquisition could be transferrable to the real world (Hamilton et al., 2021). In addition, IVR also offers the possibility to “immerse” in environments that would otherwise be inaccessible due to high costs (e.g. Greenland) (Makransky & Mayer, 2022) or inaccessibility (e.g. inside of the human body) (Klingenberg et al., 2023; Maresky et al., 2019). Giving the learner the opportunity to feel inside the environment they need to learn about is highly stimulating and motivating (Çeken & Taşkın, 2022). As a result, IVR offers the possibility of learning in numerous fields, be it geography (Bryan et al., 2018, pp. 1–9; Olmos et al., 2018, pp. 95–106), biology (Maresky et al., 2019), astronomy (Madden et al., 2018), architecture (Ummihusna & Zairul, 2022) or ecology (Makransky & Mayer, 2022).

The aforementioned literature extensively demonstrates the value of IVR in learning, as evidenced by numerous systematic reviews on the

subject (Cao et al., 2023; Conrad et al., 2024; Radianti et al., 2020). These studies consistently highlight IVR’s potential to enhance educational experiences by providing immersive, interactive environments that foster deeper engagement and improved knowledge retention. The collective findings underscore the growing recognition of IVR as a powerful tool in educational settings, capable of transforming traditional learning approaches by immersing students in realistic scenarios that bridge the gap between theoretical knowledge and practical application. Despite promising results, learning with IVR still has some gaps, which have been highlighted in several papers (Hamilton et al., 2021; Makransky & Petersen, 2021; Mayer et al., 2022). For example, it is important to note that IVR learning does not yet seem to be the ideal tool for acquiring all types of knowledge (Makransky & Petersen, 2021). In fact, studies aimed at the acquisition of procedural knowledge have more positive results compared to studies aimed at factual knowledge, which have contrasting results. Procedural knowledge can be defined as the way of “how to do something” (Krathwohl, 2002). One reason why IVR seems to be more appropriate for procedural knowledge is that it provides optimal conditions for rehearsing and training high-risk procedures (Ryan et al., 2022; Wu et al., 2020; Xin et al., 2019). Factual knowledge can be defined as the basic elements that students must know to be acquainted with a discipline or to solve problems in it (Krathwohl, 2002). In other words, knowledge of specific details, elements, or facts. The learning of the latter has been studied, for example, by Parong and Mayer (2018). In this study, the authors compared the learning of factual knowledge in IVR with a slideshow presentation on a desktop computer. The results show that the slideshow group outperformed the IVR group in learning outcomes. Another study comparing the learning of factual knowledge through different means (i.e. video versus HMD) found no significant difference between the group that followed the lesson through the HMD and the group that followed the video lesson (Meyer et al., 2019). Other studies that combined factual knowledge and conceptual knowledge (defined as the relationships between basic elements within a larger structure that enable them to function together (Krathwohl, 2002),) into declarative knowledge also found mixed results. In fact, the study by Webster (2016) found positive results for the IVR group compared to a lecture-based group, the study by Makransky, Borre-Gude, and Mayer (2019) found no differences between an IVR group and a booklet group and finally the IVR group was outperformed by a desktop computer group in the study by Makransky, Terkildsen, and Mayer (2019). A systematic review confirms this discrepancy in findings, with half of the 24 studies analysed showing learning benefits, while the other half show no benefits or are detrimental to learning when using IVR (Hamilton et al., 2021). Taken together, these studies suggest that the instructional design for learning factual knowledge with IVR tools is not entirely clear, and further research is needed.

2.2. Cognitive load and cognitive theory of multimedia learning

Cognitive load theory suggests that learning is facilitated when conditions match the human cognitive architecture (Sweller et al., 1998). Indeed, learning requires working memory resources to process new information and is limited in the amount of new information that can be processed simultaneously. Therefore, instructional methods should avoid exceeding this capacity in order to prevent cognitive overload (Sweller et al., 2019).

The theory of multimedia learning is concerned with the cognitive load of multimedia learning (Mayer & Moreno, 1998). To reduce the cognitive load, recommendations have been provided (Mayer & Moreno, 2003). It has been suggested that applying these recommendations could be effective when learning with IVR. Indeed, studies have found that the use of IVR tools results in high levels of cognitive load (Han et al., 2021; Makransky, Terkildsen, & Mayer, 2019; Srivastava et al., 2019) as learners are faced with a variety of stimuli in an environment that may be novel (Concannon et al., 2019). Incorporating recommendations could help to reduce the cognitive load and facilitate learning,

especially since principles such as segmentation and pre-training have been shown to be effective in multimedia learning (Mayer & Moreno, 2003). Pre-training involves introducing key concepts and main characteristics before the learning session (Çeken & Taşkın, 2022). Studies have evaluated the impact of pre-training on IVR and have shown an increase in learning (Meyer et al., 2019; Petersen et al., 2020). Similarly, segmentation, which consists in dividing a lesson into smaller chunks, has recently been shown to enhance knowledge transfer during IVR learning (Klingenberg et al., 2023; Mayer & Pilegard, 2005, pp. 169–182). In both cases, the authors suggested that pre-training and segmentation may have reduced the cognitive load, thereby facilitating learning. However, cognitive load was not measured. Therefore, the authors recommended measuring cognitive load during IVR learning while applying recommendations to manage cognitive load, which would provide a better understanding of its underlying mechanisms in IVR. This suggestion adds to a systematic review that invites future research to gain a deeper understanding of this interaction (Pellas et al., 2021).

2.3. Spatial abilities

As stated in the previous section, exploring cognitive processes during learning is thus particularly important for comprehending learning outcomes whether utilising multimedia (Scheiter et al., 2017) or IVR (Parong & Mayer, 2021). Therefore, authors called for investigating the impact of spatial abilities in IVR learning (Makransky & Mayer, 2022; Makransky & Petersen, 2021). Spatial abilities may be defined as a personal ability to generate, retain, retrieve, and transform well-structured visual images (Lohman, 2013). Spatial abilities have been found to have a positive correlation with numerous academic and professional fields. Therefore, the acquisition of skills in subjects such as science, technology, engineering, mathematics (Stieff & Uttal, 2015) or anatomy (Berney et al., 2015) is positively influenced by spatial abilities. Spatial abilities also positively influence performance in professional fields such as architecture (Berkowitz et al., 2021), surgery (Wanzel et al., 2002) and art (Haanstra, 1996).

Although there is a correlation between learning and spatial abilities, it is important to note that discrepancies in results do exist (Kühl et al., 2022). Specifically, in multimedia learning, it seems that the results vary depending on whether the content is presented in an animated or static way (Heo & Toomey, 2020). Two existing theories provide an explanation. On one side, the “ability as enhancer hypothesis” states that dynamic and animated resources are more complex because they contain procedural information and thus require high spatial abilities to be assimilated. In opposition, the “ability as compensator hypothesis” states for its part that animated and dynamic content can compensate for learners’ lack of spatial abilities thereby facilitating the learning process (Heo & Toomey, 2020). As with multimedia learning, questions have been raised about the impact of spatial skills on learning through IVR. Studies have highlighted the need to assess the influence of spatial abilities on immersive learning (Makransky & Petersen, 2021; Mayer et al., 2022). While it has been demonstrated that use of virtual reality tools can improve spatial abilities (Di & Zheng, 2022; Molina-Carmona et al., 2018), it is not yet clear whether spatial abilities can enhance learning in IVR. This question is complicated by the existence of multiple spatial abilities and tests, which can create uncertainty when selecting which spatial ability to evaluate and which test to use (Yılmaz, 2009).

2.4. The present study and hypotheses

To summarise the above addressed sections. IVR tools provide the opportunity to users to feel as if they were “inside” the content to be learned and are therefore innovative new learning tools. In addition, IVR is perceived by its users as highly motivating for learning, which increases its effectiveness (Çeken & Taşkın, 2022; Makransky & Petersen, 2021). However, as explained, further research is necessary to address

the gaps regarding the learning of cognitive knowledge (i.e. factual, or conceptual knowledge) with IVR tools. Moreover, recent studies have called for measuring and investigating the effects of cognitive load (Klingenberg et al., 2023) and spatial abilities in IVR learning (Mayer et al., 2022). Answering the above points will also improve the understanding of instructional design for learning with IVR, as an analysis of the relationship between learning factual knowledge in IVR, cognitive load and spatial abilities will be attempted.

Specifically, we will assess participants’ learning of factual knowledge (i.e. general geography knowledge) in a pre-post-test experimental design. Participants will be assigned to either an immersive virtual reality group (i.e. HMD) or a non-immersive virtual reality group (i.e. desktop computer). Instructional design principles such as pre-training and segmentation, aim to manage cognitive load during the learning process (Mayer & Moreno, 2003). Hence, the study will evaluate spatial ability, cognitive load, presence, as covariates while comparing levels of motivation regarding the learning tool. Following the current research, this study proposes the subsequent hypotheses:

1. **Hypothesis 1 (H1):** Participants in the IVR group report a higher sense of presence and higher levels of motivation than the non-immersive virtual reality group.

2. **Hypothesis 2 (H2):** As HMDs induce high levels of cognitive load, it is expected that participants in the non-immersive virtual reality group will have lower overall cognitive load scores than the IVR group. An interaction between overall cognitive load scores and learning progress for the IVR group is expected.

3. **Hypothesis 3 (H3):** It is expected that the sense of presence will have a positive impact on learning. Therefore, it is anticipated that the IVR group will outperform the NIVR group in terms of learning.

4. **Hypothesis 4 (H4):** We expect an interaction between spatial abilities and progress on the general geography knowledge variable. The following results are possible:

- Given the spatial ability as enhancer theory and that the displayed environment is dynamic and complex, participants with higher spatial abilities will learn better than participants with low spatial abilities.
- Considering the spatial ability as compensator theory and that the displayed environment is dynamic and complex, no spatial abilities are needed to progress.

3. Material & methods

3.1. Participants and inclusion criteria

A total of 61 participants were recruited (36 women and 25 men) between the ages of 17 and 23 ($M = 19.48$, $SD = 1.26$). All the participants are undergraduate students at the faculties of the Catholic University of Lille in northern France. For obvious reasons, participants studying “Geography”, “Digital Learning” or those who used an HMD and Google Earth on a regular basis were not selected. Conversely, participants who had “tried on” an HMD at some point in their lives were included. The only compensation the participants received for their participation was free refreshments (i.e. free drinks and food throughout the process). All participants agreed to participate of their own free will and gave written informed consent in accordance with the Declaration of Helsinki.

3.2. Variables

We assessed general geography knowledge, spatial abilities, cognitive load, motivation, and presence. In addition, participants using an HMD were asked to indicate whether they experienced any symptoms of Cybersickness.

3.2.1. General geography knowledge

General geography knowledge was assessed using a multiple-choice-questionnaire. This type of assessment was chosen because the literature indicates that multiple-choice-questionnaires are a reliable mean of assessing factual knowledge (Al-Rukban, 2006; Palmer & Devitt, 2006). The use of a multiple choice questionnaire also reduces the testing effect compared to open-ended questions (Wiklund-Hörnqvist et al., 2014). The testing effect can be defined as the improvement in subsequent performance after taking one or more tests (Karpicke & Roediger III, 2008). The questionnaire was designed to assess participants' general knowledge of four global geographic topics:

- Major cities and capitals around the world (e.g. Budapest is located in?)
- Renowned landmarks and natural wonders (e.g. The Potala Palace is located in?)
- Seas, rivers, and oceans (e.g. The Danube flows into?)
- Well-known mountain ranges and forests (e.g. The Highlands are in?)

Once the questions elaborated, to ensure that the learning content was appropriate to the level of the participants, three independent experts (i.e. 2 high school "geography teachers" and 1 university "geography" professor reviewed the questions and were asked to rate the relevance of each question to the expected knowledge level of the participants on a 5-point Likert scale (from "really not relevant" to "extremely relevant"). This review took place several weeks before the start of the process, and questions deemed not relevant were excluded from the pool of questions.

The multiple-choice questionnaire was designed using the "UNITY®" software (Unity technologies; <https://unity.com/>). The questionnaire provided multiple options with only one being the correct answer (Fig. 1). To avoid the ceiling effect (Selzer et al., 2019) and the guessing effect (Baghaei & Amrahi, 2011), seven answer choices are offered, comprising the correct answer and one "don't know" option. During the questionnaire, the selected item flashes green if the answer is correct and red if the answer is incorrect to provide feedback to the participant (Fig. 1). If the answer is incorrect, the correct answer is not displayed to prevent the participant from remembering it. Similarly, at no point in the process is the participant given a score (i.e. summative feedback) as it can lead to disengagement due to overconfidence if they have performed well or loss of confidence, shame, or embarrassment if they have performed poorly. Indeed, studies have shown that summative feedback can have different effects on learner motivation (Harlen et al., 2002). Scores were computed by displaying the correct answers, incorrect answers and the unanswered questions and were automatically stored in an Excel spreadsheet for data analysis.

3.2.2. Spatial ability test

As previously stated, different spatial abilities and thus different tests exists. As the learning sessions consisted of orienting oneself in Google Earth in order to visit the world's most famous natural and architectural cites, we decided to use the spatial orientation test (SO) (Hegarty & Waller, 2004). The SO tests the participants' ability to imagine different perspectives or orientations in space. The SO test consisted of a booklet with a separate page for each of the test items. Each item shows a picture of an array of objects (e.g. house, flower, tree, car) and an "arrow circle" with a question about the direction between some of the objects. For the question on each page, participants should imagine themselves standing at one object in the array (e.g. the car; the word "car" will be named in the centre of the arrow circle) and facing another object (e.g. traffic light; the word "traffic light" will be named at the top of the arrow circle). The task was to draw an arrow from the centre object (the car) showing the direction to a third object (e.g. the stop sign) from this facing orientation. Here is an example of questions: "Imagine you are standing at the car and facing the traffic light. Point to the stop sign." Each participant had to complete 12 different items within a 5-min period. Physically rotating the booklet was forbidden. The final performance includes the errors (in degrees) between the indicated location and the correct location of the object.

3.2.3. Cognitive load test

To evaluate cognitive load, the NASA-Task Load Index (TLX) was used (Hart & Staveland, 1988). The NASA-TLX has been shown to be a reliable tool for assessing cognitive load (Pasaribu et al., 2022; Rubio et al., 2004). It is a 2-part evaluation procedure. First it rates the contribution of the following 6 factors to the learning task workload: (1) mental demand, (2) physical demand, (3) time constraint, (4) performance, (5) frustration level, and (6) general effort. Each item provides information about the nature of the workload imposed by the laparoscopic task. The second part consisted of rating the relative contribution of each factor to the whole workload by comparing each item to the 5 remaining others. Each participant performed 15 pairwise comparisons, thus selecting the factor contributing the most to the overall workload. The number of times each factor is selected is noted and then tallied; finally, its relative contribution to the whole load is rated with a single coefficient. Thus, the overall estimated workload results from multiplying each single rating by the weight given to each of the 6 factors. This method ensures a reliable internal consistency.

3.2.4. Presence questionnaire

There are several reliable tests for assessing "sense of presence" (Grassini & Laumann, 2020). The most commonly used are the "Presence Questionnaire" (Witmer & Singer, 1998) and the Slater-Usuh-Steed



Fig. 1. The provided image displays the multiple-choice questionnaire with feedback, respectively green if correct (A) and red if incorrect (B). Translation: A) where is located Saint Peter's Basilica? Possible choices from top to down: Prague/Italy/Vatican/Lisbon/San Marino/Paris/I don't know Translation: B) where is located Helsinki? Possible choices from top to down: Norway/Sweden/Denmark/Finland/Lithuania/Latvia/I don't know. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(SUS) questionnaire (Slater & Steed, 2000). Due to time constraints and the desire not to burden participants excessively, the SUS was selected over the Presence Questionnaire due to its brevity (Usoh et al., 2000). Hence, the questionnaire comprised seven questions assessing three dimensions of presence:

- The impression of being present in the virtual environment
- The extent to which the virtual environment becomes the dominant reality.
- The extent to which the virtual environment is remembered as a “place”.

As in previous studies (Alem et al., 2006; Usoh et al., 2000), we adapted the terminology of the questions to the virtual environment presented in our study. Participants had to rate their feelings on a 7-point Likert scale. Questions were for instance, “I had a sense of being there in the visited places?” 1. Not at all ... 7. Very much. Total score was calculated by adding scores of each item.

3.2.5. Motivation questionnaire

Learners’ intrinsic motivation to use a computer desktop or an HMD was assessed through an 8-item questionnaire inclusive of a 7-point Likert style items developed by Isen and Reeve (2005) (e.g. acquisition of knowledge through IVR piques my curiosity). The participants were asked to rate how motivating and interesting learning with IVR, or computer desktop is (1 = strongly disagree, 7 = strongly agree). Total score was calculated by adding scores of each item.

3.2.6. Cybersickness symptoms survey

To assess cybersickness, most studies use the Simulator Sickness Questionnaire developed by Kennedy et al. (1993), Stone III (2017). This test was developed to measure sickness in the context of simulation and although it is exhaustive, we did not want to take too much time to remain relevant. Therefore, in this study, we opted to inquire about participants’ symptoms they were experiencing rather than using the questionnaire, which would have been longer. Consequently, a simple yes or no form was used to assess whether participants experienced nausea, vomiting, headaches, stomach pain, dizziness, or eye strain.

3.3. Material

For the learning sessions of both groups, a Dell® desktop computer with 12th Gen Intel® Core™i9-12900 2.40 GHz processor was used to access Google Earth. In the non-immersive virtual reality group, participants navigated through the application using a keyboard and mouse. For the IVR group, the participants made use of the PICO 4 enterprise® headset which was connected to the desktop computer. This model offers a high-resolution display of 1200 PPI resolution and a large field of view of 105°. For additional information on product specifications, please refer to: <https://www.picoxr.com/global/products/pico4e/specs>.

3.4. Procedure

The study was conducted at the Catholic University of Lille from the end of September 2023 to mid-November 2023. Participants were invited to attend experimental sessions every Thursday within the specified period at 10 a.m., 12am or 2pm, depending on their availability. Before starting, some clarifications were given to avoid any misunderstandings. For example, it was clarified that the aim was to investigate the efficiency of IVR in learning, not to judge the level of their knowledge. This was particularly important as it prevented the participants from feeling embarrassed during the tests. It was also recalled that the participants were free to withdraw from the study at any time. Finally, to avoid both unexpected additional cognitive load and bias it was pointed out that participants could not access to their

mobile phones or talk to each other throughout the experiment.

The steps of the procedure are detailed below and, in a flowchart, inserted at the end of this section (see Fig. 2).

- A) Participants completed the spatial orientation test. Meanwhile the participants were randomly divided (i.e. draw) into two groups, a non-immersive virtual reality group (NIVR) using a desktop computer and an immersive virtual reality (IVR) group using an HMD.
- B) All the participants had then to answer to 40 multiple choice questions which were randomly selected from a pool of 80 questions on a desktop computer application. It should be noted that each participant’s 40 questions were different, as they were randomly selected by the computer. Participants were given 20 s to answer a question, with a countdown timer displayed on each screen. Participants were instructed to respond only if they were completely certain, otherwise they could choose the “I don’t know” option. Once the questionnaire was completed, participants were unable to modify their selections and instructed to remain in front of their desktop computer until the subsequent stage.
- C) Both groups were then invited to attend a pre-training session to familiarise themselves with the respective HMD and desktop computer. This pre-training consisted of an explanation of the necessary functionalities of the HMD that would be required to follow the upcoming learning sessions without hindrance. The pre-training also aimed to prevent learners from focusing on “how to use an HMD” rather than focusing on the content to be learned during the learning sessions. As all participants were naïve to this model and version of the HMD used, the use of a new tool in learning can increase the “extraneous processing” and overwhelm participants with content that does not support the instructional goal (Mayer et al., 2022; Wu et al., 2020). The pre-training session lasted 10 min and was scripted to ensure that the instructions were identical for each participant. If there were any blockages or misunderstandings, an expert was available to help the participant.
- D) After a brief break of 5 min, the participants commenced their first learning session lasting 15 min. The IVR group utilized an HMD connected to a desktop computer that generated the virtual environment whereas the non-immersive virtual reality group used the same desktop computer model in a conventional manner (see Fig. 3). To guide participants’ learning, it was recalled that the multiple-choice questions were all related to global geographical topics (i.e. major cities and capitals around the world; renowned landmarks and natural wonders; seas rivers and oceans; well-known mountain ranges and forests). The learning session entailed utilising “Google Earth” in virtual reality mode to travel around the globe and gain insights into the previously mentioned global topics. The participants in the HMD group were instructed not to move around in the “physical world” (see Fig. 3), to remain within the confines of the virtual environment to maintain their sense of presence. Virtual content and participant actions within the Google Earth application were displayed on an external monitor (Fig. 3). This was done to prevent participants from exiting Google Earth or attempting to deceive during the training session. At the conclusion of this first training session, participants were asked to reflect on the strategy they used (i.e. summarising), to improve it in the subsequent learning session. A major break of 15 min was then given to give the opportunity to participants to refresh and rest. Obviously, the participants were not permitted to discuss with each other.
- E) A second learning session identical to the first one.
- F) After a brief break of 5 min, all the participants completed a digital version of the NASA-TLX test to evaluate cognitive load. It was emphasized that the objective was to assess the cognitive

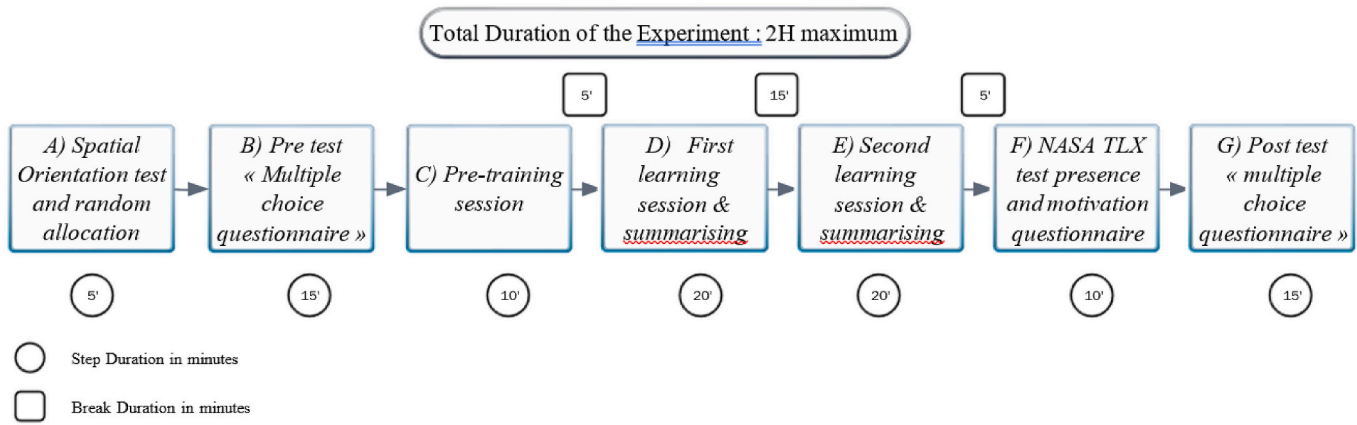


Fig. 2. Flowchart of the experimental process, including resting periods and time duration.

load allocated to the learning session. Additionally, a questionnaire on the sense of presence and motivation was completed to compare learning experiences between an HMD and a computer desktop.

- G) Finally, after completing these tests, the participants completed the multiple-choice questionnaire again. The post-test conditions were like the baseline test, and although the order differed, both the baseline and post-test questions were identical.

3.5. Data analysis

Data entries were collected in several Excel files containing participants' identifiers. These files were compiled into one.xlsx data set for analysis. Statistical analysis was performed using Python 3 (ipykernel) in the Jupyter notebook. The imported libraries for data analysis are Pandas, NumPy, Matplotlib and Stats from Scipy. Normality and equality of variances were tested prior to statistical analysis. Independent t-tests were used to compare sense of presence, motivation (H1), overall cognitive load (H2) and progress on the general geography knowledge variable (H3) between the two groups. Correlation tests were conducted to examine the impact of overall cognitive load on general geography knowledge progress (H2), the impact of the sense of presence on general geography knowledge progress (H3), and the effect of spatial ability on general geography knowledge progress (H4). The statistical significance threshold was set up for a type 1 error rate of 5%. Finally, cybersickness symptoms were collected in a descriptive way.

3.6. Groups description and assumption tests

Non-immersive group counted 30 participants (11 men and 19 women, mean age = 19.63 SD = 1.326) while the IVR group counted 31 participants (14 men and 17 women, mean age = 19.32 SD = 1.194). Assumption tests revealed homogeneity of groups.

4. Results

4.1. General geography knowledge learning

Student t-tests were conducted to assess baseline homogeneity and progress. Only the correct answers on the questionnaire were considered for analysis. Assumption tests indicated that the dataset was homogeneous and normally distributed. Student t-tests confirmed baseline homogeneity with no significant differences between the NIVR group ($M = 16.60$, $SD = 6.48$) and the IVR group ($M = 15.39$, $SD = 5.9$) before training ($p = 0.39$). Both groups showed significant improvement, respectively from ($M = 16.60$; $SD = 6.48$) at baseline to ($M = 22.20$; $SD = 7.07$) ($p < 0.001$) for the NIVR group on the post-test and from ($M =$

15.39 ; $SD = 5.9$) at baseline to ($M = 20.65$; $SD = 4.48$) ($p < 0.001$) on the post-test for the IVR group. No significant difference was found between them after training ($p = 0.30$).

4.2. Spatial orientation

The NIVR group ($M = 98.53^\circ$, $SD = 108.18^\circ$) and the IVR group ($M = 64.48^\circ$, $SD = 51.47^\circ$) showed no significant differences ($p = 0.121$). No correlation was found between spatial ability and geography knowledge progress.

4.3. Cognitive load

Overall cognitive load scores were similar between the NIVR ($M = 54.24$, $SD = 12.04$) and IVR groups ($M = 54.91$, $SD = 12.50$) with no significant difference ($p = 0.832$). However, significant differences were noted in mental demand ($p = 0.047$), physical demand ($p = 0.0462$), and frustration ($p = 0.0462$), with the IVR group showing higher mental ($M = 17.62$; $SD = 7.60$ vs $M = 13.19$; $SD = 9.37$) and physical demands ($M = 1.63$; $SD = 3.21$ vs $M = 0.87$; $SD = 1.17$) but lower frustration ($M = 8.01$; $SD = 8.54$ vs $M = 14.41$; $SD = 8.18$) (Fig. 4). A final analysis was conducted to determine which dimensions significantly contributed more to the overall cognitive load among the groups. The ANOVA demonstrated significant differences in the dimensions' contribution between both groups. The results were ($p < 0.001$) for both groups. Subsequent pairwise comparisons were performed, and the findings are presented in Tables 1 and 2 below. No correlation was found between cognitive load and geography knowledge progress.

4.4. Sense of presence

The IVR group ($M = 26.35$, $SD = 6.606$) scored significantly higher than the NIVR group ($M = 14.00$, $SD = 5.558$) on sense of presence ($p < 0.001$), though no correlation with knowledge progress was observed.

4.5. Motivation

The IVR group ($M = 46.74$, $SD = 6.303$) reported higher motivation than the NIVR group ($M = 41.90$, $SD = 9.792$) with a significant difference ($p = 0.025$).

4.6. Cybersickness

In the IVR group, 28 out of 31 participants experienced some level of cybersickness, with eye strain and headache being the most reported symptom. Specifically, thirteen participants experienced 1 symptom, 5 participants experienced 2 symptoms, 8 participants experienced 3

A



B

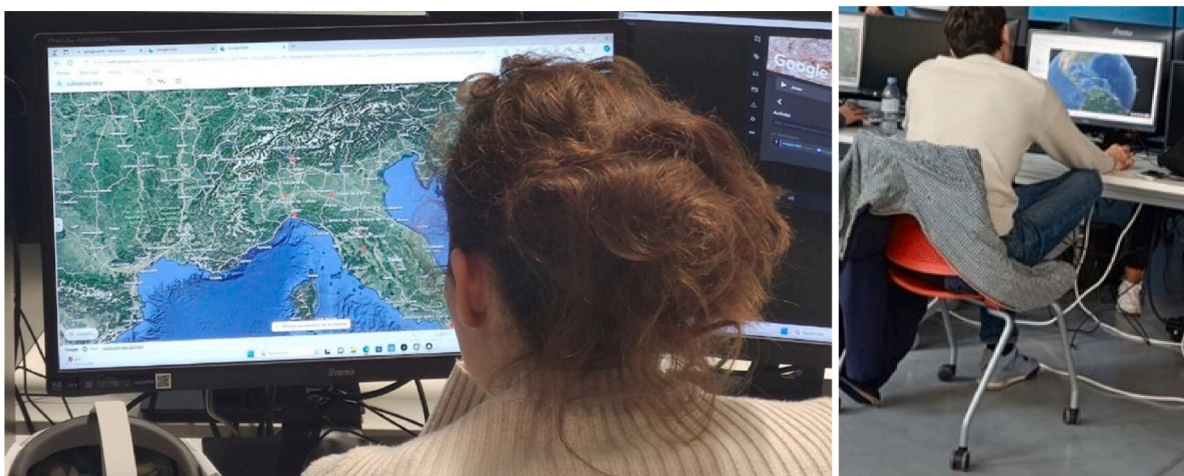


Fig. 3. A) Pictures of the IVR condition with participants' wearing head mounted displays and with the content displayed on an external monitor; B) picture of the NIVR condition.

symptoms, 1 participant experienced 4 symptoms, and 1 participant experienced 5 symptoms. More descriptive information can be found in [Table 3](#) below.

5. Discussion

5.1. Study objectives

This study had two objectives. Firstly, it aimed to investigate the effectiveness of IVR in acquiring factual knowledge, specifically general geography knowledge, compared to NIVR. The participants' motivation for learning in IVR and NIVR was also assessed. Secondly, the study also measured levels of presence, spatial abilities, and cognitive load to evaluate their impact on the learning process. The aim of this work is to contribute to the comprehension of the instructional design for learning, specifically factual knowledge, using IVR.

In relation to the scores on the general geography knowledge questionnaire, it can be observed that both groups significantly improved their performance after the learning sessions. This demonstrates the effectiveness of both tools, the HMD, and the desktop computer, for the learning of factual knowledge.

5.2. Research hypotheses

Based on the four research hypotheses proposed in this study, the following conclusions can be drawn.

5.2.1. Hypothesis 1

The IVR group reported significantly higher levels of "presence" and "motivation" compared to the NIVR group, which supports our initial hypothesis (H1). This outcome corroborates the findings of [Liu et al. \(2022\)](#), who demonstrated that students exhibit greater enthusiasm for IVR tools than for conventional tools for learning purposes. This is a crucial finding because it underscores the pivotal role that the technological components of HMDs play in fostering a more immersive and engaging learning environment. The capacity of HMDs to foster a strong sense of presence thus appears to be a key factor in enhancing their motivation and overall engagement with the learning material. This observation aligns with existing research that suggests that presence is closely linked to positive learning outcomes ([Makransky & Petersen, 2021](#)). By physically feeling part of the virtual scene, students are more likely to be motivated to interact with and retain the material presented to them.

5.2.2. Hypothesis 2

The statistical analysis did not reveal any significant differences in the overall cognitive load levels between the two groups. Additionally, overall cognitive load levels did not significantly impact the learning process, leading to the rejection of Hypothesis 2. Interestingly, although the scores for the NIVR group ($M = 54.24$, $SD = 12.04$) and the IVR group ($M = 54.91$, $SD = 12.50$) are classified as high on the cognitive load range (see [Table 4](#)) ([Pasaribu et al., 2022](#)), it did not impede learning. This indicates that the instructional design, specifically, integrating pre-training and segmentation principles efficiently mastered

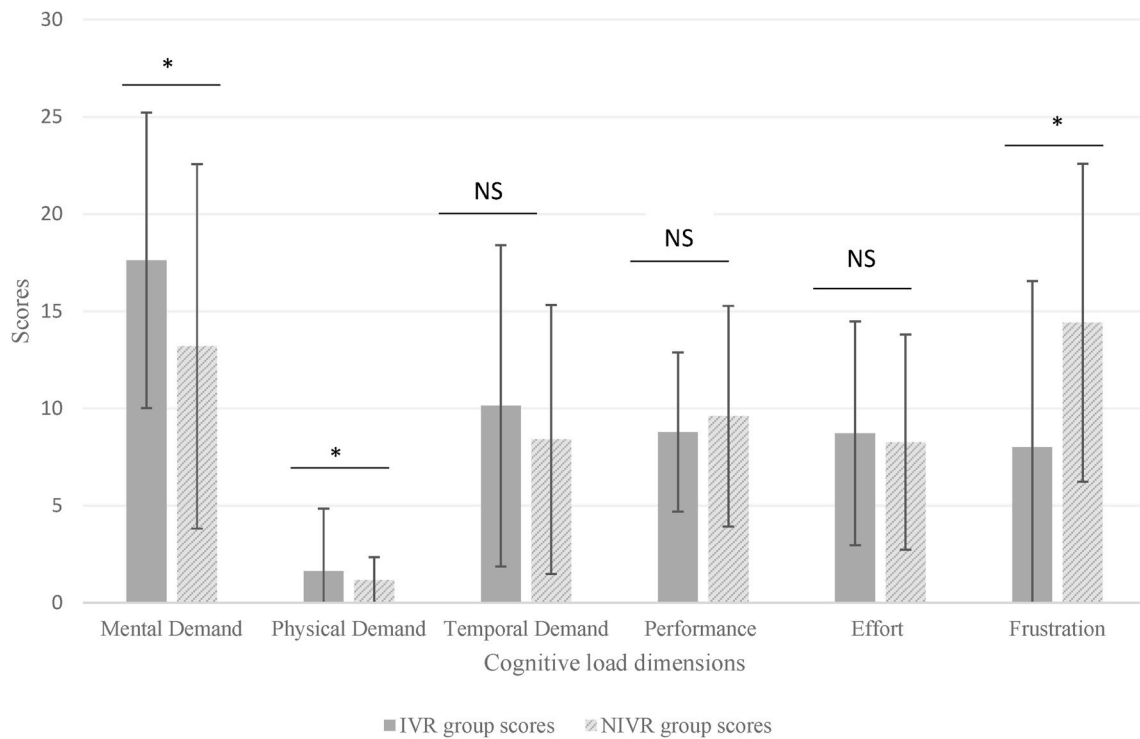


Fig. 4. Compared Cognitive load dimension scores of the IVR and NIVR group
 P values signification * < 0.05; NS = non significant.

Table 1

Pairwise comparisons of the total mental workload dimensions for the immersive virtual reality group (IVR).

IVR GROUP	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
Mental Demand		$p < 0.001^{***}$	$p < 0.001^{***}$	$p < 0.001^{***}$	$p < 0.001^{***}$	$p < 0.001^{***}$
Physical Demand			$p < 0.001^{***}$	$p < 0.001^{***}$	$p < 0.001^{***}$	$p = 0.000252^{***}$
Temporal Demand				$p = 0.416447$	$p = 0.435780$	$p = 0.322394$
Performance					$p = 0.959629$	$p = 0.650533$
Effort						$p = 0.702618$
Frustration						

Table 2

Pairwise comparisons of the total mental workload dimensions for the non-immersive virtual reality group (NIVR).

NIVR GROUP	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
Mental Demand		$p < 0.001^{***}$	$p = 0.028466^*$	$p = 0.077798$	$p = 0.016121^*$	$p = 0.592304$
Physical Demand			$p < 0.001^{***}$	$p < 0.001^{***}$	$p < 0.001^{***}$	$p < 0.001^{***}$
Temporal Demand				$p = 0.469752$	$p = 0.929167$	$p = 0.003277^{**}$
Performance					$p = 0.360417$	$p = 0.010398^*$
Effort						$p = 0.001191^{**}$
Frustration						

Table 3

Symptoms reported by the participants after the IVR training.

Symptoms	Nausea	Vomiting Feeling	Headache	Stomach pain	Dizziness	Eye strain
Number of participants reporting the symptom	5	4	16	1	8	22

Table 4
Total cognitive workload score category.

Number	Range	Total mental workload category
1	0–9	Very Low
2	10–29	Low
3	30–49	Average
4	50–79	High
5	80–100	Very High

the associated cognitive load and enabled learning for the two groups, as suggested by the literature (Han et al., 2021; Mayer & Moreno, 2003). Moreover, to enable a complete understanding it is recommended to examine the dimensions that make up the overall cognitive load (Galy et al., 2018). This provides valuable information on the dimensions which are contributing to the total load and can give indications on what remains to be improved. For example, it can be observed that the dimensions with the greatest impact on the overall scores differ between the two groups (see Fig. 5). For the IVR group, the mental demand is significantly higher than all the other dimensions (Table 1). Additionally, the mental demand dimension is also significantly higher in the IVR group when compared to the NIVR group (Fig. 4). This indicates that using an HMD for learning is more mentally demanding than using a desktop computer. When compared to the literature, the use of HMD for learning general geographic knowledge is for instance generating more mental demand, than during medical simulation learning (i.e. respectively 17 vs 15) (Bhandary et al., 2016). This is an indication that the level of mental demand of the IVR group in this study is particularly high. To gain a better understanding, it is important to note, that learning factual knowledge falls under the first level (i.e. easiest) of the revised Bloom’s Taxonomy, which involves recalling facts and basic concepts (Anderson & Krathwohl, 2021; Krathwohl, 2002). The high levels of mental demand observed in this study for the IVR group were thus certainly not caused by the complexity of the learning content. An alternative explanation is therefore possible. A frequently cited rationale is that elevated cognitive load when utilising HMDs is attributable to the richness and complexity of the virtual environment (Concannon et al., 2019; Parong & Mayer, 2021; Rogers, 2019). However, as both groups were exposed to the same virtual environment in our study, this indicates that the HMD technological features, while enabling a sense of presence also might have elevated the associated mental load. Indeed, as the sense of presence, is a psychological construct, it has been theorised that it requires mental activity, thereby increasing the associated cognitive load levels (Makransky & Petersen, 2021; Vettehen et al., 2019). Nevertheless, as other studies indicate the opposite (Jeong et al., 2011; Parong et al., 2020), the relationship between cognitive load,

specifically mental load, and sense of presence remains unclear. Further investigation is therefore required to gain a deeper understanding of this relationship. Similarly, another potential explanation is that the reported cybersickness symptoms may have influenced the mental load. Indeed, Breves and Stein (2023) demonstrated that cybersickness can act as a mediator between technological immersiveness and cognitive load. In other words, cybersickness symptoms contribute to an increase in cognitive load. Therefore, it is reasonable to assume that in our study, the reported symptoms (e.g. eye strain, headache, dizziness) affected the participants’ mental well-being. This, in turn, contributed to the feeling of high mental load in the IVR group. Regarding the first assumption, the relationship between mental load and cybersickness needs to be further investigated, as mental load, cybersickness and presence are important cognitive characteristics of IVR. Secondly, the results of the frustration dimension are also noteworthy. The findings indicate that using a computer desktop for learning factual knowledge, is more frustrating than using an HMD (Fig. 4). This result aligns with the motivation questionnaire scores, which indicate that IVR tools are more motivating for learning than NIVR tools. This is also illustrated by the overall dimensions’ weight showing that frustration is classified as the penultimate dimension for the IVR group and the second for the NIVR group (Fig. 5). Frustration is considered as a negative activating emotion, that can arise from various factors, such as reduced social interactions (Novak et al., 2023) or the absence of direction and guidance provided by an instructor (Hove & Corcoran, 2008). Frustration can lead to boredom and decrease learning if left unaddressed (Novak et al., 2023). The latter suggests that over a longer period, HMDs are more effective for learning than conventional digital tools, particularly in isolated locations or during pandemics.

5.2.3. Hypothesis 3

Although both groups demonstrated significant progress in general geography knowledge after the learning phase, the IVR group did not exhibit significantly better progress than the NIVR group. Indeed, correlation tests indicate, that the sense of presence did not enhance learning. Therefore hypothesis 3 is rejected. This finding is in line with a recent study by Ochs and Sonderegger (2022). This confirms that presence is an affordance that facilitates learning; however, the level of presence does not necessarily influence the level of learning. In other words, a heightened sense of presence in a virtual environment does not automatically lead to improved learning outcomes. This finding suggests that while presence can enhance the learning experience by making it more engaging and immersive, it is not a direct predictor of learning success. The sense of presence may contribute to a more immersive and realistic experience, but it does not ensure that students will retain more

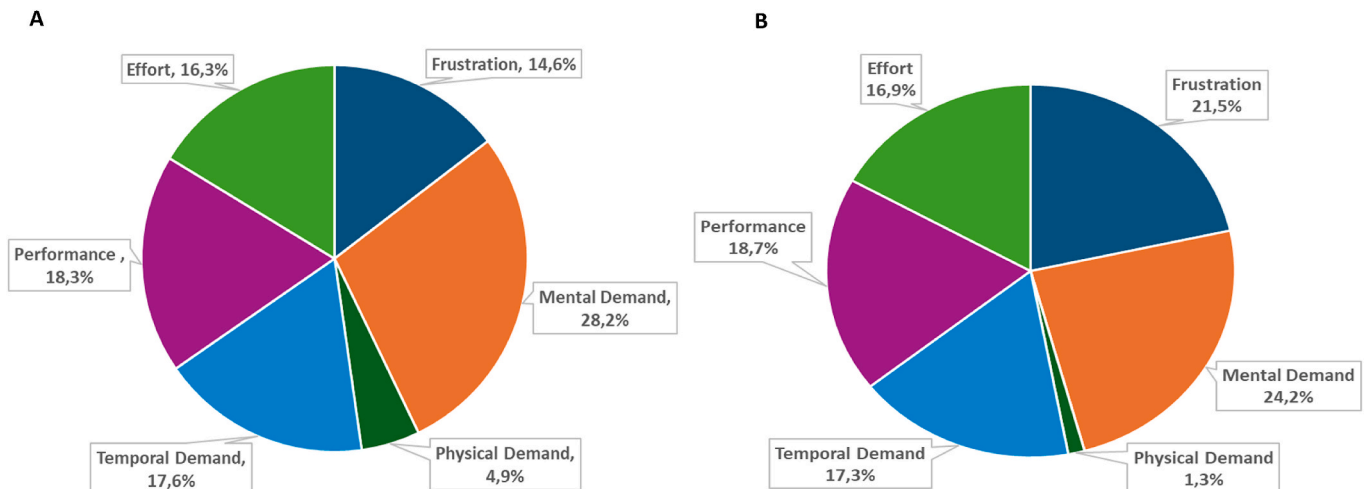


Fig. 5. Cognitive load dimension weight distribution in percentages, respectively for the IVR group (A) and the NIVR group (B).

information or achieve higher levels of understanding solely based on this factor. The implication is that instructional design within virtual environments should not focus exclusively on maximizing presence. Instead, it should ensure that the content is pedagogically sound and that other critical factors, such as cognitive load management, interactivity, and alignment with learning objectives are adequately addressed. It may be the case that enhancing the sense of agency could prove an effective method of improving the efficacy of learning in IVR (Makransky & Petersen, 2021). Indeed, in our study, participants were given the opportunity to travel around the world virtually, but no interaction with the environment was possible. Interacting with the environment is important for eliciting feedback (e.g. haptic, cognitive), which has been shown to be essential in the learning process (Bajaj et al., 2018; Pereira et al., 2016). The latter certainly explains why IVR studies focusing on procedural knowledge, are more effective than conventional tools in a learning process (Wu et al., 2020). Indeed, in procedural knowledge, each step of the procedure depends on the successful completion of the previous one, creating a feedback loop. This assumption is reinforced by the active learning theory, showing that for instance, increased physical activity or interaction can physically or cognitively involve the learner which reinforces the learning process (Hartikainen et al., 2019). This perspective encourages a balanced approach to designing immersive learning experiences, where presence is considered one of several components that contribute to effective learning, rather than the sole determinant.

5.2.4. Hypothesis 4

Finally, no significant correlation has been established between spatial abilities, specifically spatial orientation, and progress. Therefore Hypothesis 4 (a) is rejected and 4 (b) supported. This contrasts with previous findings where positive correlations have been established between spatial abilities, such as mental rotation and NIVR or multimedia learning situations (Andrade et al., 2012; Höffler, 2010). Other studies also showed that using virtual and augmented reality tools enhanced spatial abilities (Heo & Toomey, 2020; Molina-Carmona et al., 2018). The lack of correlation in this study may be due to the difficulty in identifying the spatial ability and determining which test to use, as explained by Yilmaz (2009). More likely, it seems that the learning situation did not require participants to transform or manipulate images, as it provided the opportunity to virtually turn around the content or change the perspective of view. In other words, the features provided by Google Earth, the HMD or the desktop computer during the learning sessions did not require users to activate this ability. As a result, users no longer need to physically turn the content as the display enables them to easily move around it. Joining thereby the ability-as-compensator hypothesis (Berney et al., 2015; Höffler & Leutner, 2011). Taken together this suggests, that virtual reality tools can facilitate the learning of geographical knowledge regardless of the individual's spatial orientation abilities. However, to our knowledge, studies evaluating the influence of baseline levels of spatial abilities, specifically spatial orientation on the learning process in IVR are scarce. This outcome needs thus to be confirmed by future investigations.

5.3. Conclusion and practical implications

In conclusion, this study contributes answering the question whether IVR tools, more specifically HMDs can be used to learn factual knowledge and with which modalities. Therefore, several noteworthy points can be emphasized. Firstly, the results demonstrate that HMDs are reliable tools for enhancing general geographic knowledge. Secondly, regarding the sense of presence, higher levels did not correlate with better learning. This suggests that while presence is significant, its association with for instance, interactions (i.e. agency) could facilitate feedback and enhance its influence on the learning process. This is in line with recent findings which demonstrate that virtual environments that facilitate active learner engagement lead to enhanced knowledge

acquisition (Conrad et al., 2024). Thirdly, thanks to principles such as segmentation, or pre-training (Mayer & Moreno, 2003; Meyer et al., 2019) the overall cognitive load has been efficiently managed to enable learning with both tools. However, in IVR, mental load levels are particularly significant compared to the other dimensions, even though the complexity of the learning content was low. In order to elucidate the levels of mental load associated with the IVR group in our study, assumptions were posited regarding the mental load linked to the sense of presence or cybersickness symptoms. Frustration levels also suggest that IVR tools would be more suitable for distance learning than NIVR tools. These findings corroborate the necessity for a comprehensive analysis of the overall cognitive load dimensions, as they offer invaluable insights into the interplay between the cognitive characteristics of IVR utilisation. Fourthly, results are showing that spatial abilities, specifically, individual levels of spatial orientation were not needed during the learning process. As levels of spatial abilities are heterogeneous among people (Hegarty & Waller, 2005; Jones & Burnett, 2008), this means that users having low spatial abilities will not be disadvantaged when using HMDs for learning geographic knowledge. Finally, cybersickness symptoms occurred by users although there were only two learning sessions of 15 min. Indeed, in this study, only three participants in the IVR group did not experience any symptoms. However, this did not prevent participants from completing the learning sessions in this study. As previously suggested, it is possible that the presence of cybersickness symptoms may have diminished the learning effect observed in our study for the IVR group, as evidenced by the findings of prior research (Chen et al., 2019).

5.4. Future directions and limitations

Future investigations should thus evaluate the impact of cognitive load, (more specifically mental load) on learning more complex knowledge (i.e. conceptual) (Anderson & Krathwohl, 2021) in IVR while integrating interactions. Furthermore, it would be beneficial to investigate the relationship between mental load, cybersickness and presence while learning with IVR in greater depth. Additionally, the impact of baseline levels of other spatial abilities, should be investigated. It would also be valuable to examine the impact of cybersickness symptoms and cognitive load on longer learning sessions and over a longer period of time. According to Wu et al. (2020), most studies do not report learning sessions of more than an hour, which is much shorter than current class durations, although the purpose may be to use the HMD occasionally.

Finally, this study has limitations. Specifically, the use of only two learning sessions restricts the ability to extrapolate the findings to quarter or semester-long courses. Furthermore, learning may have been constrained in our setup due to the limited capacity of working memory. According to Cowan (2008), working memory capacity is limited to remembering and retaining only five plus or minus two items. In our study, cues were found to aid participants (e.g. recall of questionnaire topics) during their learning sessions. However, it is possible that some questions were forgotten leading to a lack of retention and learning. Additionally, the original design included a retention test to assess long-term learning. However, despite several attempts to have participants fill out the questionnaire for a third time, fewer than a half of the sample responded, rendering the scores irrelevant. These limitations need to be considered for future investigations.

CRedit authorship contribution statement

Benjamin De Witte: Writing – review & editing, Writing – original draft, Supervision, Investigation, Formal analysis, Conceptualization. **Vincent Reynaert:** Writing – review & editing, Software, Investigation, Conceptualization. **Jerome Hutain:** Writing – review & editing, Software, Investigation. **Danny Kieken:** Software, Investigation. **Joseph Jabbour:** Investigation, Formal analysis. **Jalal Possik:** Resources, Funding acquisition.

Declaration of competing interest

None.

Acknowledgements

The authors acknowledge the French National Agency for Research for funding the “DemoES” project. This research is part of the “DemoES” project (ANR-DMES-0014), which investigates the effectiveness of immersive learning tools in education. Catherine Demarey, Vice-Rector in charge of academic affairs, leads this project at Lille Catholic University.

References

- Al-Rukban, M. O. (2006). Guidelines for the construction of multiple choice questions tests. *Journal of family & community medicine*, 13(3), 125.
- Alem, L., Hansen, S., & Li, J. (2006). Evaluating clinicians' experience in a telemedicine application: A presence perspective. *Proceedings of the 18th Australia conference on computer-human interaction: Design: Activities, artefacts and environments* (pp. 47–54).
- Anderson, L. W., & Krathwohl, D. R. (2021). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Andrade, A. D., Cifuentes, P., Mintzer, M. J., Roos, B. A., Anam, R., & Ruiz, J. G. (2012). Simulating geriatric home safety assessments in a three-dimensional virtual world. *Gerontology & Geriatrics Education*, 33(3), 233–252. <https://doi.org/10.1080/02701960.2011.611553>
- Baghaei, P., & Amrahi, N. (2011). The effects of the number of options on the psychometric characteristics of multiple choice items. *Psychological Test and Assessment Modeling*, 53(2), 192–211.
- Bajaj, J. K., Kaur, K., Arora, R., & Singh, S. J. (2018). Introduction of feedback for better learning. *Journal of Clinical and Diagnostic Research*, 12, 12.
- Berkowitz, M., Gerber, A., Thurn, C. M., Emo, B., Hoelscher, C., & Stern, E. (2021). Spatial abilities for architecture: Cross sectional and longitudinal assessment with novel and existing spatial ability tests. *Frontiers in Psychology*, 11, Article 609363.
- Berney, S., Bétrancourt, M., Molinari, G., & Hoyek, N. (2015). How spatial abilities and dynamic visualizations interplay when learning functional anatomy with 3D anatomical models. *Anatomical Sciences Education*, 8(5), 452–462.
- Bhandary, S., Lipps, J., Winfield, S. R., Abdel-Rasoul, M., Stoicea, N., Pappada, S. M., & Papadimos, T. J. (2016). NASA Task Load Index scale to evaluate the cognitive workload during cardiac anesthesia based simulation scenarios. *International Journal of Anesthesiology & Research*, 4(8), 300–304.
- Breves, P., & Stein, J.-P. (2023). Cognitive load in immersive media settings: The role of spatial presence and cybersickness. *Virtual Reality*, 27(2), 1077–1089. <https://doi.org/10.1007/s10055-022-00697-5>
- Bryan, S. J., Campbell, A., & Mangina, E. (2018). Scenic spheres-an AR/VR educational game. *2018 IEEE games, entertainment, media conference (GEM)*.
- Buttussi, F., & Chittaro, L. (2017). Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Transactions on Visualization and Computer Graphics*, 24(2), 1063–1076.
- Cao, Y., Ng, G.-W., & Ye, S.-S. (2023). Design and evaluation for immersive virtual reality learning environment: A systematic literature review. *Sustainability*, 15(3), 1964.
- Çeken, B., & Taşkın, N. (2022). Multimedia learning principles in different learning environments: A systematic review. *Smart Learning Environments*, 9(1), 1–22.
- Chen, X., Chen, Z., Li, Y., He, T., Hou, J., Liu, S., & He, Y. (2019). ImmerTai: Immersive motion learning in VR environments. *Journal of Visual Communication and Image Representation*, 58, 416–427.
- Concannon, B. J., Esmail, S., & Roduta Roberts, M. (2019). Head-mounted display virtual reality in post-secondary education and skill training. *Frontiers in Education*, 4, 80.
- Conrad, M., Kablitz, D., & Schumann, S. (2024). Learning effectiveness of immersive virtual reality in education and training: A systematic review of findings. *Computers & Education: X Reality*, 4, Article 100053.
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323–338.
- Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 19(2), 272–309.
- Di, X., & Zheng, X. (2022). A meta-analysis of the impact of virtual technologies on students' spatial ability. *Educational Technology Research & Development*, 70(1), 73–98.
- Galy, E., Paxion, J., & Berthelon, C. (2018). Measuring mental workload with the NASA-TLX needs to examine each dimension rather than relying on the global score: An example with driving. *Ergonomics*, 61(4), 517–527. <https://doi.org/10.1080/00140139.2017.1369583>
- Grassini, S., & Laumann, K. (2020). Questionnaire measures and physiological correlates of presence: A systematic review. *Frontiers in Psychology*, 11, 349.
- Haanstra, F. (1996). Effects of art education on visual-spatial ability and aesthetic perception: A quantitative review. *Studies in Art Education*, 37(4), 197–209.
- Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, 8(1), 1–32.
- Han, J., Zheng, Q., & Ding, Y. (2021). Lost in virtual reality? Cognitive load in high immersive VR environments. *JAIT*, 12(4).
- Harlen, W., Crick, R. D., Broadfoot, P., Daugherty, R., Gardner, J., James, M., & Stobart, G. (2002). In *A systematic review of the impact of summative assessment and tests on students' motivation for learning*.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139–183). Elsevier.
- Hartikainen, S., Rintala, H., Pylväs, L., & Nokelainen, P. (2019). The concept of active learning and the measurement of learning outcomes: A review of research in engineering higher education. *Educational Sciences*, 9(4), 276.
- Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32(2), 175–191. <https://doi.org/10.1016/j.intell.2003.12.001>
- Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. *The Cambridge handbook of visuospatial thinking* (pp. 121–169).
- Heo, M., & Toomey, N. (2020). Learning with multimedia: The effects of gender, type of multimedia learning resources, and spatial ability. *Computers & Education*, 146, Article 103747.
- Höfler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—a meta-analytic review. *Educational Psychology Review*, 22, 245–269.
- Höfler, T. N., & Leutner, D. (2011). The role of spatial ability in learning from instructional animations—Evidence for an ability-as-compensator hypothesis. *Computers in Human Behavior*, 27(1), 209–216.
- Hove, M. C., & Corcoran, K. J. (2008). Educational technologies: Impact on learning and frustration. *Teaching of Psychology*, 35(2), 121–125. <https://doi.org/10.1080/00986280802004578>
- Isen, A. M., & Reeve, J. (2005). The influence of positive affect on intrinsic and extrinsic motivation: Facilitating enjoyment of play, responsible work behavior, and self-control. *Motivation and Emotion*, 29, 295–323.
- Jeong, E. J., Bohil, C. J., & Biocca, F. A. (2011). Brand logo placements in violent games. *Journal of Advertising*, 40(3), 59–72. <https://doi.org/10.2753/JOA0091-3367400305>
- Jicol, C., Clarke, C., Tor, E., Dakin, R. M., Lancaster, T. C., Chang, S. T., Petrini, K., O'Neill, E., Proulx, M. J., & Lutteroth, C. (2023). Realism and field of view affect presence in VR but not the way you think. *Proceedings of the 2023 CHI conference on human factors in computing systems*.
- Jones, S., & Burnett, G. (2008). Spatial ability and learning to program. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, 4(1), 47–61. <https://jyx.jyu.fi/handle/123456789/20219>.
- Karpicke, J. D., & Roediger III, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966–968.
- Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3), 203–220. https://doi.org/10.1207/s15327108ijap0303_3
- Klingenberg, S., Fischer, R., Zettler, I., & Makransky, G. (2023). Facilitating learning in immersive virtual reality: Segmentation, summarizing, both or none? *Journal of Computer Assisted Learning*, 39(1), 218–230.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory and Practice*, 41(4), 212–218.
- Kühl, T., Fehring, B. C., & Münzer, S. (2022). Unifying the ability-as-compensator and ability-as-enhancer hypotheses. *Educational Psychology Review*, 34(2), 1063–1095.
- Liu, R., Wang, L., Koszalka, T. A., & Wan, K. (2022). Effects of immersive virtual reality classrooms on students' academic achievement, motivation and cognitive load in science lessons. *Journal of Computer Assisted Learning*, 38(5), 1422–1433.
- Lohman, D. F. (2013). Spatial ability and g. In *Human abilities* (pp. 97–116). Psychology Press.
- Madden, J. H., Won, A. S., Schuldt, J. P., Kim, B., Pandita, S., Sun, Y., Stone, T. J., & Holmes, N. G. (2018). Virtual reality as a teaching tool for moon phases and beyond. *arXiv preprint arXiv:1807.11179*.
- Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning*, 35(6), 691–707.
- Makransky, G., & Mayer, R. E. (2022). Benefits of taking a virtual field trip in immersive virtual reality: Evidence for the immersion principle in multimedia learning. *Educational Psychology Review*, 34(3), 1771–1798.
- Makransky, G., & Petersen, G. B. (2021). 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.
- 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.
- Maresky, H. S., Oikonomou, A., Ali, I., Ditkofsky, N., Pakkal, M., & Bally, B. (2019). Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education. *Clinical Anatomy*, 32(2), 238–243.
- Mayer, R. E., Makransky, G., & Parong, J. (2022). The promise and pitfalls of learning in immersive virtual reality. *International Journal of Human-Computer Interaction*, 0(0), 1–10. <https://doi.org/10.1080/10447318.2022.2108563>
- Mayer, R. E., & Moreno, R. (1998). A cognitive theory of multimedia learning: Implications for design principles. *Journal of Educational Psychology*, 91(2), 358–368.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52.

- Mayer, R. E., & Pilegard, C. (2005). Principles for managing essential processing in multimedia learning : Segmenting, pretraining, and modality principles. *The Cambridge handbook of multimedia learning*.
- Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video : A media and methods experiment. *Computers & Education*, *140*, Article 103603.
- Molina-Carmona, R., Pertegal-Felices, M. L., Jimeno-Morenila, A., & Mora-Mora, H. (2018). Virtual reality learning activities for multimedia students to enhance spatial ability. *Sustainability*, *10*(4), 1074.
- Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease : A review of cue integration approaches. *Consciousness and Cognition*, *21*(1), 59–68. <https://doi.org/10.1016/j.concog.2011.08.010>
- Novak, E., McDaniel, K., & Li, J. (2023). Factors that impact student frustration in digital learning environments. *Computers and Education Open*, *5*, Article 100153.
- Ochs, C., & Sonderegger, A. (2022). The interplay between presence and learning. *Frontiers in Virtual Reality*, *3*, Article 742509.
- Olmos, E., Cavalcanti, J. F., Soler, J.-L., Contero, M., & Alcañiz, M. (2018). Mobile virtual reality : A promising technology to change the way we learn and teach. *Mobile and ubiquitous learning: An international handbook*.
- Palmer, E., & Devitt, P. (2006). Constructing multiple choice questions as a method for learning. *Annals Academy of Medicine Singapore*, *35*(9), 604.
- Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785.
- 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>
- Parong, J., Pollard, K. A., Files, B. T., Oiknine, A. H., Sinatra, A. M., Moss, J. D., Passaro, A., & Khooshabeh, P. (2020). The mediating role of presence differs across types of spatial learning in immersive technologies. *Computers in Human Behavior*, *107*, Article 106290.
- Pasaribu, I. A., Vera, Y., & Sari, R. A. (2022). NASA-TLX and SOFI methods as a tool for measuring STUDENTS' MENTAL load and failure during online learning basic neurology physiotherapy courses. *Jurnal EduHealth*, *13*(1), 436–445.
- 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*(3), 835–861.
- Pereira, D., Flores, M. A., Simão, A. M. V., & Barros, A. (2016). Effectiveness and relevance of feedback in higher education : A study of undergraduate students. *Studies In Educational Evaluation*, *49*, 7–14.
- Petersen, G. B., Klingenberg, S., Mayer, R. E., & Makransky, G. (2020). The virtual field trip : Investigating how to optimize immersive virtual learning in climate change education. *British Journal of Educational Technology*, *51*(6), 2099–2115.
- 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*, Article 103778.
- Raja, R., & Nagasubramani, P. C. (2018). Impact of modern technology in education. *Journal of Applied and Advanced Research*, *3*(1), 33–35.
- Rogers, J. A. M. (2019). Minimising extraneous cognitive load in immersive virtual environments : Evaluating an immersive virtual reality educational platform against the principles of cognitive load theory. *PhD Thesis*. <https://hekyll.services.adelaide.edu.au/dspace/handle/2440/120342>.
- Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload : A comparison of swat, NASA-TLX, and workload profile methods. *Applied Psychology*, *53*(1), 61–86. <https://doi.org/10.1111/j.1464-0597.2004.00161.x>
- Ryan, G. V., Callaghan, S., Rafferty, A., Higgins, M. F., Mangina, E., & McAuliffe, F. (2022). Learning outcomes of immersive technologies in health care student education : Systematic review of the literature. *Journal of Medical Internet Research*, *24*(2), Article e30082.
- Scheiter, K., Schüler, A., & Eitel, A. (2017). Learning from multimedia : Cognitive processes and instructional support. In S. Schwan, & U. Cress (Eds.), *The psychology of digital learning* (pp. 1–19). Springer International Publishing. https://doi.org/10.1007/978-3-319-49077-9_1.
- Selzer, M. N., Gazcon, N. F., & Larrea, M. L. (2019). Effects of virtual presence and learning outcome using low-end virtual reality systems. *Displays*, *59*, 9–15.
- Slater, M., & Steed, A. (2000). A virtual presence counter. *Presence*, *9*(5), 413–434.
- Srivastava, P., Rimzhim, A., Vijay, P., Singh, S., & Chandra, S. (2019). Desktop VR is better than non-ambulatory HMD VR for spatial learning. *Frontiers in Robotics and AI*, *6*, 50.
- Stieff, M., & Uttal, D. (2015). How much can spatial training improve STEM achievement? *Educational Psychology Review*, *27*, 607–615.
- Stone III, W. B. (2017). *Psychometric evaluation of the simulator sickness questionnaire as a measure of cybersickness*. Iowa State University [PhD Thesis].
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). [No title found]. *Educational Psychology Review*, *10*(3), 251–296. <https://doi.org/10.1023/A:1022193728205>
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design : 20 Years later. *Educational Psychology Review*, *31*(2), 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Ummihusna, A., & Zairul, M. (2022). Investigating immersive learning technology intervention in architecture education : A systematic literature review. *Journal of Applied Research in Higher Education*, *14*(1), 264–281.
- Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using presence questionnaires in reality. *Presence*, *9*(5), 497–503.
- Vettehen, P. H., Wiltink, D., Huiskamp, M., Schaap, G., & Ketelaar, P. (2019). Taking the full view : How viewers respond to 360-degree video news. *Computers in Human Behavior*, *91*, 24–32.
- Wanzel, K. R., Hamstra, S. J., Anastakis, D. J., Matsumoto, E. D., & Cusimano, M. D. (2002). Effect of visual-spatial ability on learning of spatially-complex surgical skills. *The Lancet*, *359*(9302), 230–231.
- Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, *24*(6), 1319–1333.
- Wiklund-Hörnqvist, C., Jonsson, B., & Nyberg, L. (2014). Strengthening concept learning by repeated testing. *Scandinavian Journal of Psychology*, *55*(1), 10–16.
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments : A presence questionnaire. *Presence*, *7*(3), 225–240.
- 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.
- Xin, B., Chen, G., Wang, Y., Bai, G., Gao, X., Chu, J., Xiao, J., & Liu, T. (2019). The efficacy of immersive virtual reality surgical simulator training for pedicle screw placement : A randomized double-blind controlled trial. *World neurosurgery*, *124*, e324–e330.
- Yilmaz, H. B. (2009). On the development and measurement of spatial ability. *International Electronic Journal of Environmental Education*, *1*(2), 83–96.